ISSN 2415-7740 (Print)
ISSN 2415-7074 (Online)

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Открытые семантические технологии проектирования интеллектуальных систем

Open Semantic Technologies for Intelligent Systems

МАТЕРИАЛЫ
МЕЖДУНАРОДНОЙ
НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ

(Республика Беларусь, Минск, 15–17 февраля 2018)

Основан в 2017 году

Выпуск 2

Минск
БГУИР 2018

УДК 004.822+004.89-027.31

Сборник включает прошедшие рецензирование статьи VIII Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем».

Предназначен для преподавателей высших учебных заведений, научных сотрудников, студентов, аспирантов, магистрантов, а также для специалистов предприятий в сфере проектирования интеллектуальных систем.

Материалы сборника одобрены Программным комитетом конференции OSTIS-2018 и печатаются в виде, представленном авторами.

Организаторы конференции:

Министерство образования Республики Беларусь
Министерство связи и информатизации Республики Беларусь
Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»
Российская ассоциация искусственного интеллекта
Государственное учреждение «Администрация Парка высоких технологий» (Республика Беларусь)
Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси»
Учреждение образования «Брестский государственный технический университет»
ОАО «Савушкин продукт»
ФГБНУ «Институт управления образованием Российской академии образования»
Кафедра молодежной политики и социокультурных коммуникаций РИВШ
ЗАО «Qulix Systems»
ОДО «ВирусБлокАда»
ООО «АйтиРекс Групп Бел»
ООО «Альторос Девелопмент»
ООО «Лайтпоинт»
ООО «Октонион технолоджи»
ООО «РСЛ»
ООО «Софт-Перспектива»
ООО «ФордэКонсалтинг»

Техническая и информационная поддержка:

Научный журнал «Информатика»
Научный журнал «Онтология проектирования»
Научно-технический журнал «Автоматизация процессов управления»
Международный журнал «Программные продукты и системы»
Научно-производственный журнал «Веснік сувязі»

Издание осуществлено по заказу государственного учреждения «Администрация Парка высоких технологий»

Сборник включён в Российский индекс научного цитирования.

Сайт сборника: **http://proc.ostis.net**

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Батура Михаил Павлович
председатель, ректор Белорусского государственного университета информатики и радиоэлектроники

Голенков Владимир Васильевич
сопредседатель, заведующий кафедрой интеллектуальных информационных технологий Белорусского государственного университета информатики и радиоэлектроники

Живицкая Елена Николаевна
проректор по учебной работе Белорусского государственного университета информатики и радиоэлектроники

Осипов Анатолий Николаевич
проректор по научной работе Белорусского государственного университета информатики и радиоэлектроники

Никульшин Борис Викторович
проректор по учебной работе и информатизации Белорусского государственного университета информатики и радиоэлектроники

Шилин Леонид Юрьевич
декан факультета информационных технологий и управления Белорусского государственного университета информатики и радиоэлектроники

Крейнина Ольга Валерьевна
начальник центра международного сотрудничества Белорусского государственного университета информатики и радиоэлектроники

Боярко Алла Викторовна
начальник пресс-службы Белорусского государственного университета информатики и радиоэлектроники

РАБОЧАЯ ГРУППА ОРГКОМИТЕТА КОНФЕРЕНЦИИ

Гулякина Наталья Анатольевна
руководитель группы, заместитель заведующего кафедрой ИИТ Белорусского государственного университета информатики и радиоэлектроники

Гракова Наталья Викторовна
заместитель заведующего кафедрой ИИТ по воспитательной работе Белорусского государственного университета информатики и радиоэлектроники

Губаревич Анастасия Владимировна
инженер кафедры ИИТ Белорусского государственного университета информатики и радиоэлектроники

Лучина Татьяна Леонидовна
инженер кафедры ИИТ Белорусского государственного университета информатики и радиоэлектроники

Русецкий Кирилл Валерьевич
инженер-программист кафедры ИИТ Белорусского государственного университета информатики и радиоэлектроники

ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

д.т.н., проф., академик РАЕН

Кузнецов Олег Петрович

Председатель, председатель Научного совета Российской ассоциации искусственного интеллекта, заведующий лабораторией, Институт проблем управления РАН, г. Москва, Россия

д.т.н., проф.

Голенков Владимир Васильевич

Сопредседатель, заведующий кафедрой интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники, г. Минск, Беларусь

д.т.н., проф., академик НАН Беларуси

Абламейко Сергей Владимирович

Профессор механико-математического факультета, Белорусский государственный университет, г. Минск, Беларусь

к.т.н., доц.

Боргест Николай Михайлович

Профессор кафедры конструкции и проектирования летательных аппаратов, Самарский государственный аэрокосмический университет имени Академика С.П.Королёва, г. Самара, Россия

д.ф.-м.н., проф.,академик РАН

Васильев Станислав Николаевич

Председатель Научного совета по теории управляемых процессов и автоматизации; главный редактор «Автоматика и телемеханика» РАН; заместитель главного редактора «Доклады Академии наук» РАН, г. Москва, Россия

д.т.н., проф.

Гаврилова Татьяна Альбертовна

Заведующая кафедрой информационных технологий в менеджменте, Высшая Школа менеджмента СПбГУ, г. Санкт-Петербург, Россия

д.т.н., проф.

Глоба Лариса Сергеевна

Заведующая кафедрой информационно-телекоммуникационных сетей, Национальный технический университет Украины «Киевский политехнический институт им И. Сикорского», г. Киев, Украина

д.т.н., проф.

Головко Владимир Адамович

Заведующий кафедрой интеллектуальных информационных технологий, Брестский государственный технический университет, г. Брест, Беларусь

д.фил.н., проф.

Гордей Александр Николаевич

Директор, Республиканский институт китаеведения имени Конфуция БГУ, г. Минск, Беларусь

д.т.н., с.н.с.

Грибова Валерия Викторовна

Заместитель директора по научной работе, Федеральное государственное бюджетное учреждение науки Институт автоматики и процессов управления Дальневосточного отделения Российской академии наук, г. Владивосток, Россия

к.ф.-м.н., доц.

Гулякина Наталья Анатольевна

Заместитель заведующего кафедрой интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники г. Минск, Беларусь

д.т.н., проф.

Додонов Александр Георгиевич

Заместитель директора по научной работе, Институт проблем регистрации информации НАН Украины, г. Киев, Украина

д.т.н., проф., академик МАИ

Еремеев Александр Павлович

Заведующий кафедрой прикладной математики, Национальный исследовательский университет «МЭИ», г. Москва, Россия

к.фил.н., доц., д.ист.н.

Ефименко Ирина Владимировна

основатель, исполнительный директор ООО "Semantic Hub", г. Москва, Россия

д.т.н., проф.

Заболеева-Зотова Алла Викторовна

Начальник Управления региональных и межгосударственных программ РФФИ,
г. Москва, Россия

к.т.н., с.н.с.

Загорулько Юрий Алексеевич

Заведующий лабораторией, Институт систем информатики им. А.П. Ершова СО РАН,
г. Новосибирск, Россия

д.ф.-м.н., проф.

Клещев Александр Сергеевич

Главный научный сотрудник, ФГБУИН Институт автоматики и процессов управления Дальневосточного отделения РАН,
г. Владивосток, Россия

д.мед.н., проф.

Кобринский Борис Аркадьевич

Главный научный сотрудник, Институт системного анализа Федерального исследовательского центра «Информатика и управление» РАН, профессор, Российский национальный исследовательский медицинский университет им. Н.И. Пирогова Минздрава России, зав. лабораторией систем поддержки принятия клинических решений
г. Москва, Россия

д.п.н., проф.

Козлов Олег Александрович

Заместитель директора института, Учреждение Российской Академии Образования «Институт Информатизации Образования»,
г. Москва, Россия

д.м.н., проф., академик НАН Беларуси

Кульчицкий Владимир Адамович

Заместитель директора по научной работе, член-корреспондент, Национальная академия наук Беларуси,
г. Минск, Беларусь

д.т.н., проф.

Курейчик Виктор Михайлович

Заместитель руководителя по научной и инновационной деятельности, Таганрогский кампус ЮФУ,
г. Таганрог, Россия

д.т.н., с.н.с

Ландэ Дмитрий Владимирович

Заведующий отделом, Институт проблем регистрации информации НАН Украины,
г. Киев, Украина

д.т.н., доц..

Лобанов Борис Мефодьевич

Главный научный сотрудник, Объединённый институт проблем информатики НАН Беларуси,
г. Минск, Беларусь

д.т.н.

Лукашевич Наталья Валентиновна

Ведущий научный сотрудник, Московский государственный университет имени М.В.Ломоносова,
г. Москва, Россия

д.т.н., проф.

Массель Людмила Васильевна

Главный научный сотрудник, заведующая лабораторией, Институт систем энергетики им. Л.А. Мелентьева СО РАН,
г. Иркутск, Россия

к.т.н., доц.

Невзорова Ольга Авенировна

Заместитель директора по научной работе, Научно-исследовательский институт «Прикладная семиотика» АН РТ,
г. Казань, Россия

д.ф.-м.н., проф.

Осипов Геннадий Семенович

Президент РАИИ, заместитель директора по научной работе, Федеральный исследовательский центр «Информатика и управление» РАН,
г. Москва, Россия

д.т.н., проф.

Палюх Борис Васильевич

проректор по информационным технологиям, зав. кафедрой «Информационные системы», Тверской государственный технический университет,
г. Тверь, Россия

д.т.н., проф.

Петровский Александр Александрович

Заведующий кафедрой ЭВС, Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Беларусь

д.т.н., проф.

Петровский Алексей Борисович

Заведующий отделом теории и методов принятия решений, ФГБУН Институт системного анализа Российской академии наук,
г. Москва, Россия

к.ф.-м.н., проф.

Плесневич Геральд Станиславович

Профессор кафедры «Информационные технологии», Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
«МАТИ – Российский государственный технологический университет имени К.Э. Циолковского», г. Москва, Россия

академик РАО, д.пед.н., проф.

Роберт Ирэна Веньяминовна

руководитель Центра информатизации образования ФГБНУ «Институт управления образованием Российской академии образования»,
г. Москва, Россия

д.т.н., проф.

Рыбина Галина Валентиновна

Профессор, Национальный исследовательский ядерный университет МИФИ,
г. Москва, Россия

д.т.н., проф.

Сидоркина Ирина Геннадьевна

Декан факультета информатики и вычислительной техники, Поволжский государственный технологический университет,
г. Йошкар-Ола, Республика Марий Эл, Россия

д.т.н.,проф.

Смирнов Сергей Викторович

Заместитель директора по научной работе, Институт проблем управления сложными системами Российской академии наук,
г. Самара, Россия

д.ф.-м.н., проф.

Соловьёв Сергей Юрьевич

Профессор, Московский государственный университет имени М. В. Ломоносова,
г. Москва, Россия

д.т.н., проф.

Соснин Петр Иванович

Заведующий кафедрой «Вычислительная техника», Ульяновский государственный технический университет,
г. Ульяновск, Россия

д.т.н., проф.

Стефанюк Вадим Львович

Главный научный сотрудник, Институт проблем передачи информации РАН,
г. Москва, Россия

д.т.н., Академик АН РТ, проф.

Сулейманов Джавдет Шевкетович

Профессор кафедры теоретической кибернетики КГУ, заведующий Совместной научно-исследовательской лабораторией АНТ и КГУ «Проблемы ИИ»,
г. Казань, Россия

к.т.н., доц.

Тарасов Валерий Борисович

Доцент кафедры компьютерных систем автоматизации производства, Московский государственный технический университет имени Н. Э. Баумана,
г. Москва, Россия

д.э.н., проф.

Тельнов Юрий Филиппович

Заведующий кафедрой, Московский государственный университет экономики, статистики и информатики,
г. Москва, Россия

д.ф.-м.н., проф.

Тузиков Александр Васильевич

Генеральный директор, Объединенный институт проблем информатики НАН Беларуси,
г. Минск, Беларусь

д.т.н.,с.н.с.

Харламов Александр Александрович

Старший научный сотрудник, Институт высшей нервной деятельности и нейрофизиологии РАН,
г. Москва, Россия

д.т.н., проф.

Хорошевский Владимир Федорович

Профессор, базовая кафедра «Интеллектуальные технологии системного анализа и управления» ФИЦ «Информатика и управление» РАН, г. Москва, Россия

д.т.н., проф., Академик НАН Беларуси

Чернявский Александр Фёдорович

Профессор кафедры интеллектуальных систем, Белорусский государственный университет,
г. Минск, Беларусь

д.т.н., проф.

Шарипбай Алтынбек Амирович

Директор научного центра «Искусственный интеллект», Евразийский национальный университет имени Л. Н. Гумилева,
г. Астана, Казахстан

д.т.н., проф.

Янковская Анна Ефимовна

Профессор, Томский государственный архитектурно-строительный университет,
г. Томск, Россия

д.т.н., проф.

Ярушкина Надежда Глебовна

Зав. кафедрой «Информационные системы», Ульяновский государственный технический университет,
г. Ульяновск, Россия

ORGANIZATIONAL COMMITTEE

Mikhail Batura

Chairman, Rector, Belarusian State University of Informatics and Radio Electronics

Vladimir Golenkov

Co-chairman, Head of Informational Intelligent Technologies Department, Belarusian State University of Informatics and Radio Electronics

Elena Zhivitskaya

Vice-Rector for Academic Work, Belarusian State University of Informatics and Radio Electronics

Anatoly Osipov

Vice-Rector for Research and Development, Belarusian State University of Informatics and Radio Electronics

Boris Nikulshin

Vice Rector for Education and Information, Belarusian State University of Informatics and Radio Electronics

Leonid Shilin

Dean of the Faculty of Information Technologies and Control, Belarusian State University of Informatics and Radio Electronics

Volha Kreinina

Head of International Office, Belarusian State University of Informatics and Radio Electronics

Alla Boyarko

Head of Press Service, Belarusian State University of Informatics and Radio Electronics

CONFERENCE ORGANIZATIONAL COMMITTEE WORKGROUP

Natalia Guliakina

Group leader, assistant head of IIT Department at Belarusian State University of Informatics and Radioelectronics

Natalia Grakova

assistant head of IIT department at Belarusian State University of Informatics and Radioelectronics

Anastasia Hubarevich

engineer at IIT department of Belarusian State University of Informatics and Radioelectronics

Tatiana Lutchina

engineer at IIT department of Belarusian State University of Informatics and Radioelectronics

Kirill Rusetski

software engineer at IIT department of Belarusian State University of Informatics and Radioelectronics

PROGRAM COMMITTEE

Ablameyko S.	d. of t.s., RB proff., NAS Belarus academician
Borgest N.	c. of t.s., RF ass. Proff.
Chernyavsky A.	d. of t.s., RB proff., NAS Belarus academician
Dodonov A.G.	d. of t.s., Ukraine proff.
Efimenko I.	c. of phyl.s., RF ass. Proff.
Eremeev A.	d. of t.s., RF proff.
Gavrilova T.	d. of t.s., RF proff.
Globa L.	d. of t.s., Ukraine proff.
Golenkov V.	d. of t.s., RB proff.
Golovko V.	d. of t.s., RB proff.
Gribova V.	d. of t.s., RF
Guliakina N.	c. of ph.-m.s., RB ass. Proff.
Hardzei A.	d. of phyl.s., RB proff.
Kharlamov A.	d. of t.s., RF
Khoroshevsky V.	d. of t.s., RF proff.
Kleschev A.	d. of t.s., RF proff.
Kobrinskiy B.	d. of med.s., RF
Komartsova L.	d. of t.s., RF
Kozlov O.	d. of teach.s., RF proff.
Kulchitskii V.	d. m.s., RB proff.
Kureychik V.	d. of t.s., RF proff.
Kuznetsov O.	d. of t.s., RF proff.
Lande D.	d. of t.s., Ukraine
Lobanov B.	d. of t.s., RB proff.
Loukachevitch N.	d. of t.s., RF

Massel L.	d. of t.s., RF proff.
Nevzorova O.	c. of t.s., RF ass. Proff.
Osipov G.	d. of ph.-m.s., RF proff.
Palyukh B.	d. of t.s., RF proff.
Petrovsky A.	d. of t.s., RB proff.
Petrovsky A.	c. of t.s., d. of t.s., RF proff.
Plesniewicz G.	c. of ph.-m.s., RF proff.
Robert I.	Academician of REA, d. of teach.s., RF proff.
Rybina G.V.	d. of t.s., RF proff.
Sharipbay A.	d. of t.s., Kazakhstan proff.
Sidorkina I.	d. of t.s., RF proff.
Smirnov S.	d. of t.s., RF proff.
Soloviev S.	d. of ph.-m.s., RF proff.
Sosnin P.	d. of t.s., RF proff.
Stefanuk V.	d. of t.s., RF proff.
Suleymanov D.	AS Tatarstan academician, KF
Tarasov V.	c. of t.s., RF ass. Proff.
Telnov Yu.	d. of e.s., RF proff.
Tuzikov A.	d. of ph.-m.s., RB proff.
Vasilyev S.	academician of RAS, d. of ph.-m.s., RB proff.
Yarushkina N.G.	d. of t.s., RF proff.
Yankovskaya A.E.	d. of t.s., RF proff.
Zaboleeva-Zotova A.	d. of t.s., RF proff
Zagorulko Yu.	c. of t.s., RF ass. Proff.

СОДЕРЖАНИЕ

TABLE OF CONTENTS

ПРЕДИСЛОВИЕ

Основным практическим результатом исследований в области искусственного интеллекта является разработка не только интеллектуальных систем, но и технологий, обеспечивающих быстрое и качественное построение таких систем. Создание указанных технологий требует решения следующих задач:

- чёткого выделения логико-семантического уровня интеллектуальных систем, который абстрагируется от всевозможных вариантов технической реализации этих систем (в том числе и от использования принципиально новых компьютеров, ориентированных на аппаратную поддержку интеллектуальных систем);
- разработки онтологии проектирования интеллектуальных систем и унификации описания логико-семантических моделей интеллектуальных систем;
- обеспечения платформенно независимого характера логического проектирования интеллектуальных систем, результатом которого является унифицированное описание логико-семантических моделей проектируемых интеллектуальных систем;
- использования методики компонентного проектирования интеллектуальных систем, в основе которой лежит постоянно пополняемая библиотека многократно используемых компонентов интеллектуальных систем (многократно используемых подсистем, компонентов баз знаний, агентов обработки знаний, компонентов пользовательских интерфейсов);
- обеспечения семантической совместимости многократно используемых компонентов интеллектуальных систем и семантической совместимости самих интеллектуальных систем и технологий их проектирования.

Основной целью ежегодных международных научно-технических конференций OSTIS (Open Semantic Technology for Intelligent Systems) является создание условий для расширения сотрудничества различных научных школ, вузов и коммерческих организаций, направленного на разработку и применение массовых и постоянно совершенствуемых технологий компонентного проектирования интеллектуальных систем.

Основной темой VIII конференции OSTIS являются проблемы интеграции различных моделей представления знаний и моделей решения задач в прикладных интеллектуальных системах.

VIII конференция OSTIS посвящается памяти профессора Лотфи Аскера Заде – основоположника ряда крупных научных направлений в искусственном интеллекте и профессора Юрия Роландовича Валькмана – заведующего отделом распределенных интеллектуальных систем Международного научно-учебного центра информационных технологий и систем Национальной академии наук Украины и Министерства образования и науки Украины.

Председатель Программного комитета конференции OSTIS
Председатель Совета Российской ассоциации искусственного интеллекта
Кузнецов Олег Петрович

FOREWORD

Development of not only intelligent systems but also technologies that ensure fast and efficient construction of intelligent systems is the main practical result of research in the artificial intelligence. Development of these technologies requires the following tasks solution:

- precise separation of logical-semantic level of intelligent systems, which is abstracted from various variants of the technical implementation of these systems (including the use of innovative computer-based hardware support for intelligent Systems);
- development of the ontology of intelligent systems design and unification of intelligent systems logical-semantic models description;
- providing of a platform independent nature of the logical design of intelligent systems, which result is a unified description of logical-semantic models of intelligent systems;
- use of a component design methodology of intelligent systems, which is based on permanently expanding library of reusable components of intelligent systems (reusable subsystems, knowledge bases components, knowledge processing agents, user interfaces components);
- ensuring of semantic compatibility of reusable components of intelligent systems and semantic compatibility of these intelligent systems and technologies of such systems design.

Creating the conditions for the expansion of cooperation between different scientific schools, universities and business organizations, aimed on the development and using of mass and continuously improved component design technologies for intelligent systems is the main purpose of annual international scientific and technical conferences OSTIS (Open Semantic Technology for Intelligent Systems).

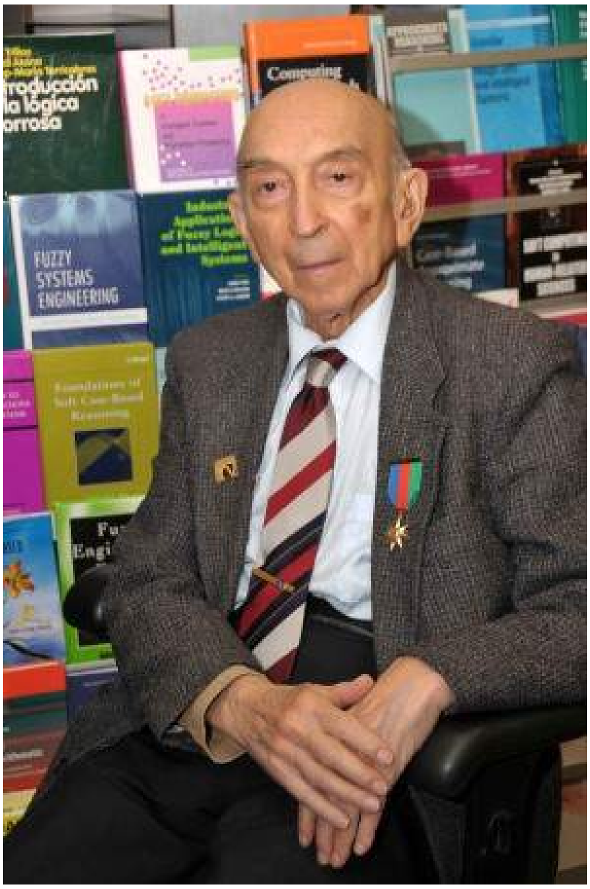
The main topic of the VIII OSTIS conference is problems of various knowledge representation models and problem solving models integration in applied intelligent systems.

The VIII OSTIS conference is dedicated to the memory of Professor Lotfi Askar Zadeh, the founder of a number of major scientific trends in artificial intelligence and Professor Yuri Rolandovich Valkman, head of the distributed intelligent systems department of the International Scientific and Educational Center for Information Technologies and Systems of the National Academy of Sciences of Ukraine and the Ministry of Education and Science of Ukraine.

Programme Committee Chair.
Chairman of the Council of the Russian Association for Artificial Intelligence
Kuznetsov Oleg Petrovich

Remembering Lotfi Zadeh

Pospelov D.A, Stefanuk V.L., Averkin A.N., Batyrshin I.Z., Tarassov V.B., Yarushkina N.G., Yazenin A.V.
Russian Association for Artificial Intelligence
Russian Association for Fuzzy Systems and Soft Computing



Lotfi Zadeh
February 4, 1921 – September 6, 2017

Recently, world science has suffered a serious loss. The great scientist of our time, the founder of a number of major scientific trends in the theory of control, applied mathematics electrical engineering, computer science, artificial intelligence, the Father of Fuzzy Logic, Professor Lotfi Zadeh passed away on 6 September 2017. He was part of a cohort of very few pioneering scientists who generate new, original scientific ideas and form the basic scientific paradigms that are changing our world.

Professor L.Zadeh was the creator of fuzzy set theory and linguistic variables, fuzzy logic and approximate reasoning, possibility theory and soft computing, information granulation and generalized uncertainty theory, z-numbers and generalized constraints, etc. His ideas and theories not only opened a new era in the development of science free from restrictions of narrow disciplines to enable interdisciplinary synergies. They highly contributed to the emergence of new information and cognitive technologies, brought about the arrival of effective industrial applications, such as fuzzy controllers and computers, fuzzy chips and networks, fuzzy recognition and clustering systems, and so on [1].

Zadeh's origin and amazing personal destiny made him a convinced internationalist. In his interviews he said that «the question really is not whether I'm American, Russian, Iranian, Azerbaidjani or anything else. I've been shaped by all these people and I feel quite comfortable among all of them» [1]. He called himself «an American, mathematically oriented, electrical engineer of Iranian descent, born in Russia» [2]. This phrase can seem strange for young reader. Zadeh's birthplace is Novkhany near Baku. Why Russia? Where began the «Russian trace» of his biography?

The youth of Zadeh's parents, his birth and childhood came at the turning points of our history: disintegration of the Russian Empire and the emergence of the USSR. From 1859 to 1917 Baku was the center of Baku province of the Russian Empire. At the crossroad of the XIX-XXth centuries it was considered as the fastest growing city of all Russia. The main reason for the rapid growth of economic activity was a large oil production. At the beginning of the XXth century more than half of the world oil output came from Baku region. It worthily enjoyed the glory of «Russian Texas». Engineers and specialists from the most developed countries, including Germany, United Kingdom, USA, France, Sweden, worked at Baku province and introduced the most advanced industrial technologies of those times at the Baku plants.

So a multinational city emerged, the center of different cultures and religions that rapidly became an economic capital of Russian South. During the «oil fever» period, it grew rapidly, attracting rich people, businessmen and simple adventurers from all over the world. The Rothschild and Nobel companies prospered. . . This «oil fever» weakened, but did not stop even after the Great Russian revolution. In this wonderful city, «Paris of the East», Zadeh's parents met each other. It happened already in the years of great social cataclysms, Civil War in Russia and USSR formation. Between 1918 and 1920 there was chaos in Baku, one power (musavatists) succeeded another one (commissars), Turkish and German troops occupied the city. In April 1920, the Red Army broke into Baku, and the Azerbaijan Soviet Socialist Republic was organized.

Lotfi Asker Zadeh was born in **Novkhany**, **Baku Region**, Soviet Azerbaijan, on **February 4, 1921** as Lotfi Aliaskerzadeh [3]. His father, Rahim Aliaskerzade, an Iranian Azeri from **Ardabil**, was a journalist, the foreign correspondent for the newspaper Iran in Baku. Rahim also had a good own business – wholesale matches. His Russian Jewish mother, Feiga (Fanya) Korenman, from **Odessa**, studied medicine and became a pediatrician. After the wedding she also obtained the status of Iranian citizen.

So his mother tongue was Russian. In 1928 Lotfi was enrolled in the Russian elementary school No 16 in Baku. Even 80 years after he kept fond memories about his school [4]: «at dawn of the Soviet era, what was extolled was science, scientists and engineers». Under this influence, Lotfi decided at an early age to become an engineer. This decision was the core of his outlook on life.

In Baku Lotfi completed three years of elementary school. It is not surprising that Russian culture and literature highly influenced the young boy (Figure 1).

At the beginning of 1930's ethnic Azerbaijanis with foreign passports were faced with a choice – obtain Soviet citizenship or leave the USSR. In 1931 Zadeh's parents decided to move to Iran, his father's homeland, taking him with them. Ten-year's Lotfi was enrolled in American Alborz College in Tehran, which was a Presbyterian missionary school.

Figure 1. Lotfi Zadeh in his youth

During his student days at the Albortz college, Lotfi met Fay, his future wife, who was the student at the women's branch of Alborz college. She also had Russian roots (her family was from Dvinsk, an old city on the Western Dvina, nowadays Latvian Daugavpils). According to her words, in Tehran period of Zadeh's life «in his room on the shelves along the wall a library of about 2000 volumes in Russian was collected» [5]. Young Lotfi read Tolstoy and Dostoevsky, Turgenev and even Shakespeare (translated into Russian).

After graduating from American college, Lotfi Zadeh passed the exams to the University of Tehran and placed third in the entire country. In 1942, he graduated from the University of Tehran with a degree in electrical engineering. Then Lotfi worked with his father, who did business with American military commanders in Tehran as a supplier of construction materials. Despite a rather high income, Lotfi decided to leave behind a comfortable life in Tehran and immigrate to the United States to fulfill his dream of a career in the academic world (see Figure 2).

Lotfi Zadeh left Tehran early in 1944 traveling to the United States by air and sea. He arrived in New York in July 1944 and moved to Cambridge after spending the summer months working at the International Electronics Corporation.

MIT: 1944-1946 MS	Columbia University: 1946-1949 PhD 1949-1959	UC Berkeley: 1959-2017, 1963-1968

Figure 2. A chronological representation of Zadeh's educational and academic trajectory in the USA

Zadeh entered the Massachusetts Institute of Technology in 1944 as a graduate student and received his MS degree in electrical engineering from MIT in February 1946 (Figure 3).

Figure 3. Newly-made master Lotfi Zadeh with his parents

His parents came to the United States a little later and settled in New York. Lotfi did not want to be away from them and decided to move to Columbia University. There he was lucky to find the position of instructor in electrical engineering. After spending three years as instructor, he obtained his PhD degree in 1949 under the supervision of Professor J. Ragazzini. The thesis was concerned with the frequency analysis of time-varying networks.

In 1950's Zadeh's scientific interests shifted from classical electrical engineering to systems analysis and information science. Already in 1950 he published a significant paper «An extension of Wiener's theory of prediction», co-authored with Professor Ragazzini. This work found application in designing finite-memory filters; today it is considered classical. In 1952 Zadeh again together with Ragazzini proposed the z-transform method for discrete systems. Nowadays this method is also viewed as classical one; it is widely used in digital signal processing. In 1953 developed a new approach to non-linear filtration and constructed a hierarchy of non-linear systems, which was based on the Volterra-Wiener presentation. Thus, the fundamentals of optimal non-linear processors to detect useful signals in noise were formulated.

In 1950's L.A.Zadeh became very interested in probability theory and its application to decision analysis. He met H.Robbins, a brilliant mathematician, and R.Bellman, the father of dynamic programming, who later became his close friends. In 1956-1957 he was a visiting member of the Institute for Advanced Study in Princeton, New Jersey. There he was inspired by a course of logic taught by S.Kleene.

In 1954 Lotfi Zadeh was promoted to the rank of Associate Professor and he received a full professor rank in 1957.

Zadeh taught for ten years at the Columbia University. In January 1959 Professor J.Whinnery, the Chair of Electrical Engineering Department at the University of California, proposed him to move to Berkeley. There were pros and cons. After weighing it, in July 1959 the 38-year-old Lotfi Zadeh with his family (Figure 4) started a long journey by car from New York to Berkeley.

Figure 4. Lotfi Zadeh with his mother Feiga and wife Faina (Fay). San Francisco, 1960

Professor Lotfi Zadeh joined the Department of Electrical Engineering at the University of California, Berkeley, in 1959 and served as its chairman from 1963 to 1968. In 1963 he published an important book [6], co-authored by Prof. Desoer, where a new state-based approach to linear system theory was described. This book has survived 4 editions. Its ideas and results were the sources of various modern approaches in systems analysis and automatic control. Nowadays the state space approach is widely used in system engineering ranging from industrial robots to space guidance control.

Thus, in mid-sixties, Professor L.A.Zadeh had already become a leading scientist in the field of systems theory, automatic control theory, and their applications. However, an innovation spirit proper to Lotfi Zadeh did not allow him to rest on his laurels. In 1965 44-year-old Prof. L.A.Zadeh published in Information and Control a main scientific work of his life – the pioneering paper «Fuzzy Sets» [7]. This work is of great historical significance. It opened a new scientific area that induced a powerful resonance all over the world and generated an enormous flow of publications. This flow is not exhausted up to now.

A basic Zadeh's idea was simple: real human reasoning based on natural language cannot be adequately modeled in the framework of classical mathematical methods. The introduction of *fuzzy set* – a class with vague boundaries, described by membership function, provided a suitable basis for developing more flexible approach to reasoning, decision-making and modeling of complex humanistic systems. The behavior of such systems is characterized rather by linguistic variables than ordinary numeric variables.

In the above mentioned paper «Fuzzy Sets» L.Zadeh defined the concept of fuzzy set and its ordinary level sets, suggested various ways of specifying intersection and union operations, introduced pseudo-complementation operation, and new unary operations – concentration and dilatation? Fuzzy relations, their composition and projections were defined. The extension principle was formulated, and mappings of fuzzy sets were considered. Fuzzy sets with fuzzy membership functions called fuzzy sets of type 2 were introduced. Fuzzy restrictions and translations rules for fuzzy propositions were proposed.

The development of fuzzy models for complex systems to bridge the gap between classical logic and intuition, creation of innovative formal approaches allowing adaptation of strict mathematics to real human ways of everyday thinking and communication – there were novel keynote scientific problems formulated by Lotfi Zadeh.

In fact, such an unexpected and sharp turn over of Zadeh's scientific interests from «honorable strict science» to non-classical «vague science» had been a risky step that put in danger his further scientific career. Primarily, the ideas of fuzzy sets and fuzzy logic were rather coldly received by Western scientists, including Americans. For instance, one of the most brilliant Zadeh's ex-students, R.Kalman, the author of the well-known filter, had to say [4]: «Fuzzification is a kind of scientific permissiveness, it tends to result in socially appealing slogans unaccompanied by the discipline of hard scientific work and patient observation… These proposals could be severely, even brutally criticized from a technical point of view».

This primary negative reaction to fuzzy logic and linguistic variables had a rational explication. At first, from ancient times, Western philosophy and logic was based on the principle of clear boundaries and the law of excluded middle. Yes or no, true or false, good or bad, who is not our friend, is our enemy, and other similar expressions of hard opposition are the children of classical binary logic. The adoption of intermediate truth values and grades of truth by many-valued and fuzzy logics was a break-down for this perfect, certain world. At second, during the ages it was a deep-seated tradition in science to accord much more respect to numbers than to words. The introduction and use of linguistic variables put this tradition in question.

The first two papers on fuzzy sets by Lotfi Zadeh, «Fuzzy Sets» [7] and «Shadows of Fuzzy Sets» [8] were published almost at the same time in the USA and Soviet Union. The first Zadeh's translator into Russian V.L.Stefanuk confirms that Lotfi himself selected adequate words for Russian translation. He wished that his new ideas on fuzzy sets and fuzzy logics were known both in the West and in the East.

To perform it in the period of «iron curtain», Professor Zadeh decided to disseminate his works in the East, particularly, in Soviet Union and Japan.

In this fortunate for him 1965, Lotfi and Fay Zadeh visited Soviet Union two times [5]: the first one they were in Moscow in May by invitation of Popov Radio-Electronics and Automatics Society and the second time he was invited by the Soviet Academy of Sciences to attend in September a six-days congress on automatic control. After the opening ceremony in September 20, 1965 at Odessa Opera House, the congress was held on board of Admiral Nakhimov's cruise ship (Figure 5) on the route Odessa – Batumi and back.

Figure 5. Professor L.Zadeh with the participants of the Congress on board of Admiral Nakhimov's ship. Black Sea, September, 1965

This congress was attended by more than a thousand Soviet scientists and about 60 foreign guests, including Lotfi Zadeh and his wife Fay. On the board of «Admiral Nakhimov» L.A.Zadeh delivered his first talk on fuzzy set theory in Russian.

In the memoirs of Academician N.N.Moisseev [9], it was mentioned that this Zadeh's talk provoked a great interest and was favorably received by the audience, although in the remarks of pure mathematicians some condescending notes were guessed due to rather informal mode of presentation and non-usual terminology. So in 1960-1970's the concept of fuzziness in applied mathematics was better received by Soviet scientists than their American colleagues.

It was after these trips to the USSR that L.Zadeh found good personal acquaintances and kind friends among Soviet scientists. In particular, Academicians N.N.Moisseev, R.V.Gamkrelidze, V.A.Ilyin, G.S.Pospelov, V.I.Siforov, Ya.Z.Tsypkin became his friends. Nikita Moisseev himself remembered with pleasure the International School on Optimal Control in Dubrovnik two years after the first Zadeh's visit in Moscow [9]. He delivered lectures as an invited professor and shared a picturesque cottage over the Adriatic Sea with the families of two other lecturers – Lotfi Zadeh and Richard Bellman. Every evening the lecturers were sitting together over the sea and discussed on mathematics. Their mutual sympathy and close scientific views served an additional stimulus for friendship.

Just Moisseev and Pospelov were the first outstanding Soviet scientists, who highly appreciated and promoted Zadeh's works. The Russian translation of the book «Linear System Theory: The State Space Approach» was published in 1970 under the editorship of G.S.Pospelov.

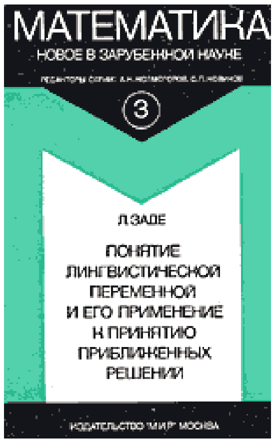


Figure 6. The first Zadeh's book on linguistic variables and fuzzy sets was in Russian

In 1974 an important volume «Mathematics Today» appeared (edited by N.N.Moisseev); it contained the Russian translation of Zadeh's foundational work «Outline of a New Approach to the Analysis of Complex Systems and Decision Processes». Also Academician Nikita Moisseev initiated the Russian translation of Zadeh's papers on linguistic variables and suggested their publication in the form of monograph in the book series «Mathematics: New Trends in Foreign Science». This book «The Concept of a Linguistic Variable and Its Application to Approximate Reasoning». edited by N.Moisseev and S.Orlovsky, was published in 1976 by the famous publishing house Mir [10].

For a long time, this thin book in Russian (Figure 6) remained the only Zadeh's monograph on fuzzy sets and linguistic variables.

Also in 1976, a new volume (collection of foreign papers) «Issues of Decision Analysis and Decision-Making Procedures» was published in Russian with a preface by G.S. Pospelov. The translated version of the paper by R.E.Bellman, L.A.Zadeh «Decision-Making in a Fuzzy Environment» was included into it.

In his turn, in 1978 Lotfi Zadeh invited N.N. Moisseev and M.A.Aizerman to join the editorial board of newly formed International Journal of Fuzzy Sets and Systems.

A funny coincidence: a well-known Soviet scientist from the Institute for Control Studies of the Soviet Academy of Sciences, Prof. M. Aizerman was also born in Dvinsk like Fay Zadeh.

In 1966 Lotfi and Fay Zadeh again visited the USSR. The future «Father of Fuzzy Logic» participated at the XVth International Congress of Mathematicians in August 16-26. The first time this congress was held in Moscow. The opening ceremony took place in the Grand Kremlin Palace, and regular sessions were held in high-rise building of Moscow State University.

This year, together with R.Bellman and R.Kalaba, L.Zadeh published a paper [11] on the use of fuzzy sets in abstraction and pattern classification in prestigious Journal of Mathematical Analysis and Applications.

In late 1960's Lotfi Zadeh published such papers as «Probability Measures of Fuzzy Events» [12], «Fuzzy Algorithms» [13], «Note on Fuzzy Languages» [14]. He was always interested in probability theory and searched for natural ways of its extension. In [12] Zadeh introduced the notion of a fuzzy event. Usually an event is seen as a precisely specified collection of points in the sample space. By contrast, in everyday experience one frequently encounters situations in which an «event» is ill-defined or fuzzy. Zadeh cited as examples of ill-defined events: «It is a warm day», «x is approximately equal to 5», «In twenty tosses of a coin there are several more heads than tails». These expressions are fuzzy because of the imprecision of the meaning of the underlined words. He generalized the mathematical expressions for mean, variance and entropy in probability theory in case of fuzzy events. In his opinion, there are many concepts and results in probability theory, information science and related fields which admit of such generalization.

A new conceptual framework for decision-making in the case of fuzzy goals and fuzzy constraints was proposed in the earlier mentioned paper «Decision-Making in a Fuzzy Environment» [15]. The most important feature of this framework is its symmetry with respect to goals and constraints – a symmetry that erases the differences between them and makes it possible the specification of fuzzy goals and constraints in the same set of alternatives. Here fuzzy decision is obtained as a convolution of fuzzy goals and fuzzy constraints.

This confluence principle was detailed by considering three cases: intersection, product and convex combination of fuzzy goals and constraints. It was also shown that the case where the goals and the constraints were defined as fuzzy sets in different spaces could be easily reduced to the previous case as they would be defined in the same space. Furthermore, the authors illustrated the new decision-making framework by examples of multi-stage decision processes, stochastic systems in a fuzzy environment and systems with implicitly defined termination time.

In «Similarity Relations and Fuzzy Orderings» [16] two basic kinds of fuzzy relations were defined. The degrees of similarity and preference were introduced. A fuzzy similarity as generalization of the notion of equivalence is a reflexive, symmetric and transitive fuzzy relation. A fuzzy ordering is a fuzzy relation which is transitive. In particular, a fuzzy partial ordering is a fuzzy ordering which is anti-symmetric and reflexive. At last, fuzzy linear ordering meets the extended condition of linearity: for any two alternatives x, y either x is preferred to y with a degree $\mu > 0$ or inversely y is preferred to x with a degree $\mu > 0$. Various properties of fuzzy similarity and fuzzy ordering relations were investigated and, as an illustration, an extended version of Szpilrajn's theorem was proved.

Zadeh was deeply interested in the problems of natural and artificial languages that stimulated his studies on semantics. His main semantic question was «Can the fuzziness of meaning be treated quantitatively, at least in principle? In the paper «Quantitative Fuzzy Semantics» [17] he gave an affirmative answer to this question. In the section «Meaning» of this paper he formulated the basics: «We consider two spaces: a universe of discourse U and a set of terms T, which play the roles of names of subsets of U. Let the generic elements of T and U be denoted by x and y, respectively. Then the meaning $M(x)$ of a term x is given a fuzzy subset of U characterized by a membership function $\mu(y \mid x)$ which is conditioned on x. For instance, if we take a color palette, then the meaning of «red» $M(red)$, is a fuzzy subset of U».

In the following section «Language», Zadeh defined a language L a fuzzy binary relation in $T \times U$ that is characterized by the membership function $\mu L : T \times U \to [0, 1]$.

Another semantic-oriented paper which was appeared in 1972, concerned the concept of linguistic hedge. A basic idea suggested in [18] was that a linguistic hedge such as «very», «more», «more or less», «much», «essentially», «slightly» etc. may be viewed as a nonlinear operator which acts on the fuzzy set representing the meaning of its operand.

Among Zadeh's works in 1970's, four seminal papers are of special concern: the already mentioned «Outline of a New Approach to the Analysis of Complex Systems and Decision Processes» [19], «The Concept of Linguistic Variable and Its Application to Approximate Reasoning» [20], as well as «Fuzzy Logic and Approximate Reasoning» [21] and «Fuzzy Sets as a Basis for a Theory of Possibility» [22].

In [19] Zadeh's *Principle of Incompatibility* was formulated: «As the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics». And further: «the key elements in human thinking are not numbers, but labels of fuzzy sets, that is, classes of objects in which the transition from membership to membership is gradual rather than abrupt. Indeed, the pervasiveness of fuzziness in human thought processes suggests that much of the logic behind human reasoning is not the traditional two-valued or even multi-valued logic, but the logic with fuzzy truths, fuzzy connectives, and fuzzy rules of inference».

Three main features of the proposed new approach were noticed:

1) use of linguistic variables in place of or in addition to numerical variables;
2) characterization of simple relations between fuzzy variables by conditional statements;
3) characterization of complex relations by fuzzy algorithms.

In particular, if x and y are linguistic variables, the conditional statements describing the dependence of y on x can be written in the form: «If x is small then y is very large», «If x is not small and not large then y is not very large», and so on. A fuzzy algorithm [13] is an ordered sequence of instructions (like a computer program) in which some of the instructions may contain labels of fuzzy sets, e.g. «Reduce x slightly if y is large», «If x is small then stop; otherwise increase x by 2».

Besides, a compositional rule of inference was proposed and the notion of «Computation of the Meaning of Values for a Linguistic Variable» was introduced. Fuzzy relational and behavior algorithms, in particular, algorithm Behavior, algorithm Oval, algorithm Intersection, algorithm Obstacle and others were constructed.

«The Outline of a New Approach...» was really a landmark paper. It served as a foundation of fuzzy control: on its basis E.Mamdani developed the first fuzzy controller.

Another keynote paper [20] contained a basic definition of linguistic variables: «by a linguistic variable we mean a variable whose values are words or sentences in a natural or artificial language». In more specific terms, a linguistic variable LV is characterized by a quintuple $\langle L, T(L), U, G, M\rangle$, where L is the name of the variable, $T(L)$ is the term-set of L, that is the collection of its linguistic values; U is a universe of discourse, G is a syntactic rule which generates the terms in $T(L)$ and M is a semantic rule which associates with each linguistic value X its meaning $M(X)$. Here $M(X)$ denotes a fuzzy subset of U. The meaning of a linguistic value X is characterized by a compatibility function $c : U \to [0,1]$, which associates with each u in U its compatibility with X.

In this paper, the examples of term-sets were specified for *Age, Appearance, Truth, Probability*, etc. «The specification of *Truth* as a linguistic variable with values such as *true, very true, completely true, not very true, untrue*, etc., leads to what is called *fuzzy logic*. By providing a basis for approximate reasoning, that is, a mode of reasoning which is neither exact nor very inexact, such logic may offer a more realistic framework for human reasoning than the traditional two-valued logic». Basic logical connectives for fuzzy logic were specified. An example of approximate Modus Ponens rule was given

It was shown that probabilities, too, can be treated as linguistic variables with values such as *likely, very likely, unlikely*, etc.

The paper entitled «Fuzzy Logic and Approximate Reasoning» [21] was published in 1975 and was the first Zadeh's great publication with reflection on fuzziness in logic (the short paper of 1974 [23] can be mentioned only in a historical retrospective).

The term «fuzzy logic» is used in this paper to describe an imprecise logical system, in which the truth-values are fuzzy subsets of the unit interval with linguistic labels such as *true, false, not true, very true, more or less true, quite false, very false*, etc. Linguistic truth values are not allowed in traditional logical systems, but are routinely used by humans in everyday discourse. The truth-value set is assumed to be generated by a context-free grammar, with a semantic rule providing a means of computing the meaning of each linguistic truth-value as a fuzzy subset of [0,1]. Since is not closed under the operations of negation, conjunction, disjunction and implication, the result of an operation on truth-values requires, in general, a linguistic approximation. As a consequence, the truth tables and the rules of inference in fuzzy logic are inexact and depend on the meaning associated with the primary truth-value *true* as well as the modifiers *very*, *quite*, *more* or *less*.

In [21] L.Zadeh summarized: «Perhaps the simplest way of characterizing fuzzy logic is to say that it is a logic of approximate reasoning. As such, it is a logic whose distinguishing features are:

i fuzzy truth-values expressed in linguistic terms with modifiers;
ii imprecise truth tables;
iii rules of inference whose validity is approximate rather than exact.

In these respects, fuzzy logic differs significantly from standard logical systems ranging from the classical Aristotelian logic to inductive logics and many-valued logic with set-valued truth-values».

In the paper «Local and Fuzzy Logics» [24] the authors emphasized that «Fuzzy logic is local, i.e. both the truth values and their conjunctions such as "AND", "OR" and "IF-THEN" have variable rather than fixed meanings. This is the reason why fuzzy logic can be viewed as a local logic. Hence, the inference process has a semantic character rather than a syntactic one: in FL, the conclusion depends on the meaning assigned to the fuzzy sets that appear in the set.

Consequently, fuzzy logic is the result of a double weakening of the basic laws of classical logic. On the one hand, the principle of bivalence and the law of excluded middle are rejected, that gives rise to a multi-valued logic and, finally, to a membership function that allows us to interpret the predicates. On the other hand, the variability of the meaning related both to truth values and connectives, makes the logical inference imprecise.

Later on, in 1994 L.Zadeh already noticed that «The term fuzzy logic is actually used in two different senses. In a narrow sense, fuzzy logic is a logical system which is an extension of multi-valued logic and is intended to serve as a logic of approximate reasoning. But in a wider sense, fuzzy logic is more or less synonymous with the theory of fuzzy sets» [25].

Nowadays, a broad concept of fuzzy logic includes fuzzy sets and linguistic variables (specifically, linguistic truth values), fuzzy relations and approximate reasoning, fuzzy rules and fuzzy constraints, test-score (experience-based) semantics and generalized uncertainty theory, etc. It encompasses a variety of soft formal methods and tools for fuzzy control, pattern recognition, natural language processing, and so on. In some sense, it implements the engineering approach to logical modeling.

The really close acquaintance of Russian specialists in Artificial Intelligence, including some authors of this paper, with Professor L.Zadeh happened at the International Workshop on AI in Repino near Leningrad in 1977. This scientific meeting took place in a small resort on the shore of the Gulf of Finland and was in its own way unique. The leading experts in AI in Western Europe and North America, mainly its founding fathers such as J.McCarthy, M.A.Arbib, J.E.Hayes, C.Hewitt, D.B.Lenat, D.Michie, N.J.Nilsson, R.S.Shank, L.A.Zadeh, joined forces with their Soviet counterparts to give answers to new scientific challenges in AI (Figure 7). At this workshop Professor Zadeh gave a presentation on approximate reasoning. In 1960-1970's L.A.Zadeh was sending to Moisseev and G.S.Pospelov the preprints of his papers from Berkeley (the University of California memos). It considerably contributed to the formation of Soviet specialists in the field of fuzzy sets and their applications. When reading them, it seemed that a large team was working on a new topic.

Lotfi Zadeh always treated Soviet scientists very well and willingly talked with them in Russian. In an informal conversation with him, we noted the importance of fuzzy reasoning for intelligent systems, thanked him for these preprints and congratulated with good disciples. How great was our surprise, when Lotfi said that he had no post-graduate students in the field of fuzzy sets for a long time. He had post-graduate students only in the areas of system science and automatic control. The reason was that fuzzy sets and systems not only did not find financial support, but were under an unofficial ban. Then no one knew that everything would change in a few years, not even years, but months later, and fuzzy technologies, especially, fuzzy controllers, will begin their triumphal march through Japan and the whole world.

The paper by Lotfi Zadeh «A Theory of Approximate Reasoning», as other contributions of this Workshop, was published in the volume «Machine Intelligence 9» [26], preface by Academician A.P.Ershov, edited by J.E.Hayes, D.Michie and L.I.Mikulich.

After this wonderful meeting, some scientific trips of professors from the Academy of Sciences of the USSR to Berkeley were organized (Figure 8).

A seminal paper «Fuzzy Sets as a Basis for a Theory of Possibility» [22] was published in 1978 in the first issue of Fuzzy Sets and Systems. Here the main Zadeh's thesis was as follows: «When our main concern is with the meaning of information rather than with its measure, the proper framework for information analysis is possibilistic in nature than probabilistic one. What is needed for such an analysis is not probability theory but an analogous and yet different. This theory might be called the theory of possibility. The mathematical apparatus of the theory of fuzzy sets provides a natural basis for the theory of possibility, playing a role which is similar to that of measure theory in relation to the theory of probability».

In [22] Zadeh introduced the concept of possibility distribution function via membership function of fuzzy set and considered it as an interpretation of fuzzy restriction. The possibility measure can be easily constructed by possibility distribution with using max-normalization. To a large extent, possibility theory is comparable to probability theory because

Figure 7. The International Workshop on Artificial Intelligence in Repino near Leningrad (April 18-24, 1977). From right to left: L.Zadeh is participating in a discussion, J.McCarthy, computer scientist known as the father of AI, V.I. Varshavsky, Soviet classic in the field of collective behavior of automata, D.A.Pospelov, founder of AI in Soviet Union

Figure 8. The Soviet scientists from the Academy of Sciences of the USSR after the workshop in Berkeley (1977). From right to left: L.Zadeh, Ya.Z. Tsypkin, M.A.Aizerman, E.Jury, A.A.Dorofeuk

it is based on set-functions. However, the possibility measure is a modification of probability measure: the two first axioms of classical measure–boundary condition and monotonousness axiom– are preserved, but the additivity axiom is replaced by max-axiom (the either-or condition).

The possibility-probability consistency principle was proposed to explain a weak connection between possibility and probability.

Marginal possibility distributions and conditional possibility distributions were also studied in [22]. Moreover, the possibility distributions of composite and qualified propositions were introduced. Conditional translation rules of type I and type II were proposed.

In brief, possibility theory is an uncertainty theory aimed at handling of incomplete information. It differs from the probability theory by the use of a pair of dual set functions (possibility and necessity measures) instead of only one. Besides, it is not additive and makes sense on ordinal structures.

According to [27], possibility theory is one of the most promising off-springs of fuzzy sets that can bridge the gap between artificial intelligence and statistics. Possibility theory clarifies the role of fuzzy sets in uncertainty management and explains why probability degrees, viewed as frequency or betting rates, can be used to derive membership functions.

The concepts of possibility theory were successfully applied in PRUF (Possibilistic Relational Universal Fuzzy) – a meaning representation language for natural languages [28]. In addition to approximate reasoning, PRUF can be employed as a language for the representation of imprecise knowledge and as a means of precisiation of fuzzy propositions expressed in a natural language.

In the opinion of B.Turksen [29], by the late 1970's, Lotfi Zadeh and his followers essentially developed the foundation of applied fuzzy mathematics.

Many Zadeh's papers in 1980's were dedicated to applications of fuzzy sets, fuzzy logic, possibility theory in artificial intelligence. In[30] it was stressed that «Management of uncertainty is an intrinsically important issue in the design of expert systems because much of the information in the knowledge base of a typical expert system is imprecise, incomplete or not totally reliable. Fuzzy logic subsumes both predicate logic and probability theory, and makes it possible to deal with different types of uncertainty within a single conceptual framework. In fuzzy logic, the deduction of a conclusion from a set of premises is reduced, in general, to the solution of a nonlinear program through the application of projection and extension principles. This approach to deduction leads to various basic syllogisms which may be used as rules of combination of evidence in expert systems. Among them are the intersection/product syllogism, the generalized modus ponens, the consequent conjunction syllogism, and the major-premise reversibility rule».

In [31] this fuzzy syllogistic topic was continued: «A fuzzy syllogism in fuzzy logic is defined as an inference schema in which the major premise, the minor premise, and the conclusion are propositions containing fuzzy quantifiers. A basic fuzzy syllogism in fuzzy logic is the intersection/product syllogism. Furthermore, it is noticed that syllogistic reasoning in fuzzy logic provides a basis for reasoning with dispositions, that is, with propositions that are preponderantly but not necessarily always true. It is also shown that the concept of dispositionality is closely related to the notion of usuality and serves as a gateway to what might be called a theory of usuality, a theory that may eventually provide a computational framework for commonsense reasoning».

The theory of disposition was outlined in [32, 33]. The basic idea was that «commonsense knowledge may be viewed as a collection of dispositions, that is, propositions with implied fuzzy quantifiers». Typical examples of dispositions are: *Icy roads are slippery*. *Tall men are not very agile*. *What is rare is expensive*, etc. It is understood that, upon restoration of fuzzy quantifiers, a disposition is converted into a proposition with explicit fuzzy quantifiers, e.g., *Tall men are not very agile* $\rightarrow$ *Most tall men are not very agile*.

Since traditional logical systems do not provide methods for representing the meaning of propositions containing fuzzy quantifiers, such systems are unsuitable for dealing with commonsense knowledge. An appropriate computational framework for dealing with commonsense knowledge is provided by fuzzy logic, which is the logic underlying approximate reasoning. A summary of the basic concepts and techniques underlying the application of fuzzy logic to knowledge representation was given in [34].

In 1990's Professor L.A.Zadeh perceived a new burst of creative energy that was resulted in opening innovative scientific areas–Soft Computing and Information Granulation. In 1994 two pioneering papers on Soft Computing [35,36] appeared. This Zadeh's initiative is closely related to the emergence of hybrid systems in computer science and AI.

In biology, hybridization is considered as the most powerful form of integration, when in one organism the various hereditary features are merged. By analogy, a hybrid system in Computer Science includes two or more heterogeneous subsystems, integrated by a shared goal or joint actions, although these subsystems can have both different nature and specification languages. In brief, hybrid computer systems use two or more various computer technologies

According to Zadeh [35,36], the basis of soft computing is that unlike the traditional hard computing, soft computing is aimed at an accommodation with pervasive imprecision of real world. The guiding principle of soft computing is to exploit the tolerance for imprecision and uncertainty to achieve tractability, robustness, low solution cost and better rapport with reality. The role model for soft computing is the human mind.

Soft computing is not a single methodology. Moreover, it is not a simple collection of methodologies, but their partnership. The principal partners in this juncture are fuzzy logic, neurocomputing, genetic algorithms, probabilistic reasoning and chaos theory. Here fuzzy logic is mainly concerned with imprecision and approximate reasoning, probabilistic reasoning – with uncertainty and propagation of beliefs, neural network – with learning, genetic algorithm – with search and optimization, and chaos theory – with nonlinear dynamics.

In essence, fuzzy logic, neurocomputing, genetic algorithms, etc. are complementary and synergistic rather than competitive technologies. For this reason, it is advantageous to use them in combination.

The 1st EUFIT (European Congress on Fuzzy and Intelligent Techniques) took place in September 7-10, 1993 in Aachen (Germany). The arrival of soft computing was the reason to rename it, beginning from the 2nd EUFIT (the abbreviation remained the same) into European Congress on Intelligent Techniques and Soft Computing. After the creation of Soviet Association for Fuzzy Systems in 1991 (from 1993 it became known as Russian Association for Fuzzy Systems), we regularly participated at both IFSA and EUFIT congresses and rather often met L.A.Zadeh. In particular, we have good memories of the 6th and 7th International Fuzzy Systems Association World Congresses in Rio de Janeiro (Figure 9) and Prague in 1995 and 1997 respectively.

Figure 9. The IFSA-95 Conference in Rio de Janeiro. From right to left: L.A.Zadeh, A.N.Averkin, D.A.Pospelov, A.P.Ryjov, G.Klir, T. Yamakawa

In Rio Professor L.Zadeh made a presentation «New Frontiers of Fuzzy Logic» and in Prague he was awarded the B.Bolzano Medal by the Academy of Sciences of the Czech Republic «For outstanding achievements in fuzzy mathematics». By the way, he became a Foreign Member of Russian Academy of Natural Sciences (Computer Sciences and Cybernetics Section) in 1992.

Computing with Words (CW) is a methodology in which words are used in place of numbers for computing and reasoning. In [37] it was argued that «fuzzy logic plays a pivotal role in CW and vice-versa. Thus, as an approximation, fuzzy logic may be equated to CW».

In two first decades of XXIst century such initiatives of the Father of Fuzzy Logic as an Information Granulation Theory and a non-traditional Granular Mathematics program seem to be of primary concern. In 1979 his work entitled «Fuzzy Sets and Information Granularity» was published, where information granules were introduced [38].For some time this work remained imperceptible. The situation changed in 1997 when L.Zadeh formulated some fundamentals of the Theory of Fuzzy Information Granulation (TFIG) in his seminal paper «Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic» [39].

The term «granule» is originated from Latin word *granum* that means grain, to denote a small particle in the real world. In [39] L.Zadeh specified *granule* as «a collection of objects which are drawn together by indistinguishability, similarity, proximity or functionality». There are various classifications of granules: crisp and fuzzy granules, information and knowledge granules, time and space granules, etc.

Specifically, the following Zadeh's granulation principle [40] is worth mentioning: «To exploit the tolerance for imprecision, employ the coarsest level of granulation, which is consistent to allowable level of imprecision».

Granulation is a basic property of human cognition. There are three basic concepts that underlie human cognition: granulation, organization and causation. Informally, granulation involves decomposition of whole into parts; organization involves integration of parts into whole; and causation involves association of causes with effects.

The TFIG is inspired by the ways in which humans granulate information and reason with it. However, the foundations of TFIG and its methodology are mathematical in nature. The point of departure in TFIG is the concept of a generalized constraint. A granule is characterized by a generalized constraint which defines it. The principal types of granules are: possibilistic, veristic and probabilistic.

In Zadeh's opinion, «the fuzzy information granulation may be viewed as a mode of generalization which may be applied to any concept, method or theory. Related to fuzzy granule, there are the following principal modes of generalization:

1) *Fuzzification* (f-generalization). In this mode of generalization a crisp set is replaced by a fuzzy set.
2) *Granulation* (g-generalization). In this case, a set is partitioned into granules.
3) *Randomization* (r-generalization). In this case, a variable is replaced by a random variable.
4) *Usualization* (u-generalization). In this case, a proposition expressed as X is A is replaced with: usually (X is A)».

The TFIG provides a reliable basis for computing with words [37]. The point of departure in CW is the observation that in a natural language words play the role of labels of fuzzy granules. In computing with words, a proposition is viewed as an implicit fuzzy constraint on an implicit variable. The meaning of a proposition is the constraint which it represents.

In CW, the initial data set (IDS) is assumed to consist of a collection of propositions expressed in a natural language. The result of computation, referred to as the terminal data set (TDS), is likewise a collection of propositions expressed in a natural language. To infer TDS from IDS the rules of inference in fuzzy logic are used for constraint propagation from premises to conclusions.

There are two main rationales for CW. First, computing with words is a necessity when the available information is not precise enough to justify the use of numbers. And second, computing with words is advantageous when there is a tolerance for imprecision, uncertainty and partial truth that can be exploited to achieve tractability, robustness, low solution cost and better rapport with reality.

The 1st International Conference on Soft Computing and Computing with Words in System Analysis, Decision and Control was held in Antalya, Turkey, on June 6-8, 2001. It was dedicated to the 80th anniversary of Professor L.A.Zadeh. The jubilee himself gave a lecture «A critical view of the foundations of control and decision». He pointed out that «CW opens the door to a potentially radical enlargement of the role of natural language in science and, in particular, in information processing, decision and control».

Even in 2000's after his 80th anniversary Professor L.A.Zadeh continued to work actively. He extended CW to computing with perceptions [40-42] developed the theory of generalized constraints [39] and the generalized uncertainty theory [43], introduced the concept of fuzzy validity [44], proposed an extended fuzzy logic based on fuzzy geometry and fuzzy transforms [45,46], as well as the restriction-centered theory of truth [47], with truth values interpreted as fuzzy restrictions, the theory of Z-numbers [48], i.e. two-fold numbers, where the first component is a restriction, and the second component is a measure of reliability (certainty) of the first one.

In 2000-2010's Lotfi Zadeh was a Professor of Graduate School and served as the director of BISC (Berkeley Initiative in Soft Computing). The BISC program of UC Berkeley was positioned as a leading organization of fundamental and applied investigations on soft computing. An important part of its activities was to hold meetings and conferences.

The BISCSE'05 International Conference was held in the University of California, Berkeley, in November 2-5, 2005. Here SE stands for Special Event in Honor of Professor Zadeh, which was organized in occasion of the 40th anniversary of his seminal paper on fuzzy sets. The event was well attended by most of the pioneers in fuzzy logic, and by prominent researchers and practitioners around the world who were interested in expanding the knowledge frontiers by the use of fuzzy logic and soft computing methods.

This event was a good occasion to congratulate the Father of Fuzzy Logic and to give diplomas of honorary doctor of Ulyanovsk State Technical University and Tver State University (Figure 10).

Figure 10. Professor L.A.Zadeh receives the honorary doctorat from Ulyanovsk State Technical University. Three Presidents of Russian Association for Fuzzy Systems and Soft Computing (in different years) – Ildar Batyrshin, Nadezhda Yarushkina and Alexander Yazenin congratulate the Father of Fuzzy Logic on the 40th anniversary of Fuzzy Sets. Berkeley. November 2, 2005.

At the World Congress of the The International Fuzzy Systems Association in June 2007, Cancun, Mexico, 86-year old L.A.Zadeh gave the hour-long plenary talk «Fuzzy Logic as the Logic of Natural Languages». As always, he was glad to talk on the sidelines with participants from Russia (Figure 11).

Figure 11. Lotfi Zadeh and vice-president of Russian Association for Fuzzy Systems and Soft Computing Valery Tarassov at IFSA-2007 World Congress in Cancun, June 19, 2007

In a few months we talked again to Professor L.Zadeh, already in Europe, at the 5th EUSFLAT Conference in Ostrava.

The 6th World Conference on Soft Computing dedicated to 95th anniversary of L.A,Zadeh took place in Berkeley in May

22-25, 2016. Ildar Baryrshin and Vadim Stefanuk attended this conference as advisory co-chairs and made their presentations. It was the last meeting with Lotfi Zadeh.

Figure 12. Vadim Stefanuk, a pioneer of AI in the USSR, at Zadeh's office in Berkeley: the last meeting. May, 2015

The scientific strength of the theory is largely determined by the possibilities of its further evolution and extension, the resonance it causes in the scientific community. Fuzzy set theory has not been the only model that has been introduced to deal with imprecise and uncertain information. During the past five decades a lot of new models have been proposed to mathematically tackle incomplete information. Some models are extensions of fuzzy set theory and others use a different path.

The following set theories were created after Zadeh's paper on fuzzy sets [7]: Type-2 fuzzy set theory and L-fuzzy set theory, Flou set theory and L-flou set theory, Interval-valued fuzzy set theory and Probabilistic set theory, Intuitionistic fuzzy set and Twofold fuzzy set theory, Grey set theory and Soft set theory, Toll set theory and Bipolar fuzzy set theory, Vague set theory and Theory of imprecise probabilities, Rough set theory, Fuzzy rough set theory and Rough fuzzy set theory, etc. (see [49,50]).

The role of Lotfi Zadeh in the modern world is not limited only to scientific achievements that opened the age of innovative fuzzy technologies. His extraordinary destiny, his own graded membership to different nations and cultures, comfortably carried through the whole life, that had directed his international activities, made a valuable contribution to the formation of planetary scientific community of XXI century. It accelerated the emergence of a new synergistic scientific vision that supposes a symbiosis of eastern and western traditions.

And yet, it is great that last travel of Lotfi Asker Zadeh was his way back to Baku. He returned 86 years later to his hometown of a new country which is proud of a great man.

REFERENCES

[1] Lotfi Zadeh Available at: https://en.wikipedia.org/wiki/Lotfi_A._Zadeh

[2] B.Blair, "Interview with Lotfi Zadeh, Creator of Fuzzy Logic Azerbaijan International, vol. 2, No 4 (Winter 1994), p.46.

[3] Thomson Gale, Lotfi Asker Zadeh Biography, World of Computer Science.

[4] L.Zadeh, "My Life and Work: a Retrospective View", Applied and Computational Mathematics, Special Issue on Fuzzy Set Theory and Applications, vol.10, No 16 pp.4-9.

[5] F.Zadeh, My Life and Travels with the Father of Fuzzy Logic. TSI Press, Albuquerque, New Mexico, USA, 1998.

[6] L.Zadeh, Ch.Desoer, Linear System Theory: The State Space Approach. New York: McGraw-Hill Book Company Inc., 1963.

[7] L.Zadeh, "Fuzzy Sets", Information and Control, vol.8, 1965, pp.338-353.

[8] L.Zadeh, "Shadows of Fuzzy Sets" Problems Inform. Transmission, 1966, vol. 2, pp. 29–34.

[9] Moiseev NN, How far to tomorrow. Free thinking. 1917-1993. Moscow: MNEPU Publishing House, 1997.

[10] L.Zadeh. The concept of a linguistic variable and its application to the adoption of approximate solutions. Moscow, The World, 1976.

[11] R.E.Bellman, R.Kalaba, L,Zadeh, "Abstraction and Pattern Ckassification", Journal af Mathematical Analysis and Applications, vol.13, 1966, pp.1-7. Reprinted in: Fuzzy Models for Pattern Recognition. J.C.Bezdek, S.K.Pal (Eds.). New York: IEEE Press, 1992, pp.231-235.

[12] L.A.Zadeh, "Probability Measures of Fuzzy Events", Journal of Mathematical Analysis and Applications, vol.23, 1968, pp.421-427.

[13] L.A.Zadeh, "Fuzzy Algorithms", Information and Control, vol.12, 1968, pp.94-102.

[14] E.T. Lee, L.A.Zadeh, "Note on Fuzzy Languages",Information Sciences, vol.1, p.421-434.

[15] R.E.Bellman, L.A.Zadeh, "Decision-Making in a Fuzzy Environment", Management Sciences, vol.17, No 4, 1970, pp.141-164.

[16] L.A.Zadeh, "Similarity Relations and Fuzzy Orderings", Information Sciences, vol.3, No 2, 1971, pp.177-200.

[17] L.A.Zadeh, "Quantitative Fuzzy Semantics", vol.3, No 2, 1971, pp.159-176.

[18] L.A.Zadeh,"A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges", Journal of Cybernetics, vol.2, 1972, pp.4-34.

[19] L.A.Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes", IEEE Transactions on Systems, Man and Cybernetics, vol.SMC-3, No 1, 1973, pp.28-44.

[20] L.A. Zadeh, "The Concept of Linguistic Variable and its Application to Approximate Reasoning". Parts 1 and 2. Information Sciences, vol.8, 1975, pp.199-249, 301-357.

[21] L.A. Zadeh, "Fuzzy Logic and Approximate Reasoning", Synthese, vol.30, 1975, pp.407-428.

[22] L.A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility", Fuzzy Sets and Systems, vol.1, No 1, 1978, pp.3-28.

[23] L.A.Zadeh, "Fuzzy Logic and Its Application to Approximate Reasoning", Proceedings of 1974 IFIP Congress on Information Processing. Amsterdam: North-Holland, vol.3, 1974, pp.591-594.

[24] R.E. Bellman, L.A. Zadeh, "Local and Fuzzy Logics", Modern Uses of Multiple-Valued Logics, J.M.Dunn and G.Epstein (Eds.). Dordrecht: D.Reidel, 1977, pp.105-165.

[25] L.A. Zadeh, "Fuzzy Logic and Soft Computing: Issues, Contentions and Perspectives", Proceedings of the 3rd International Conference on Fuzzy Logic, Neural Nets and Soft Computing [Iizuka, Japan, March 1994), pp. 1-2, 77-84.

[26] L.A.Zadeh, "Theory of Approximate Reasoning", Machine Intelligence, J. Hayes, D. Michie and L.I. Mikulich (Eds.). New York: Halstead Press, 1979, pp.149-194.

[27] D.Dubois, H.Prade, "Practical Methods for Constructing Possibility Distributions", Internationa Journal of Intelligent Systems, vol. 31, No 3, 2015, pp. 215-239.

[28] L.A.Zadeh, “PRUF: a Meaning Representation Language for Natural Languages”, International Journal of Man-Machine Studies, vol.10, 1978, pp.395-460.
[29] B.Turksen, “Scientific and Philosophical Contribution of L.A.Zadeh”, Proceedings of the 1st International Conference on Soft Computing and Computing with Words in System Analysis, Decision and Control (Antalya, Turkey, June 6-8, 2001). Kaufering:B Quadrat Verlag, 2001, pp.17-23.
[30] L.A.Zadeh,“The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems”, Fuzzy Sets and Systems, vol.11, 1983, pp.199-227.
[31] L.A.Zadeh, “Syllogistic Reasoning in Fuzzy Logic and its Application to Usuality and Reasoning with Dispositions, IEEE Transactions on Systems, Man and Cybernetics”, vol.SMC-15, 1985, pp.754-763.
[32] L.A.Zadeh, “A Theory of Common-Sense Knowledge”, Aspects of Vagueness, H.J.Skala, S.Termini, E.Trillas (Eds.). Dordrecht: D.Reidel, 1984, pp.257-295.
[33] L.A.Zadeh, “A Computational Theory of Dispositions”, International Journal of Intelligent Systems, vol.2, 1987, pp.39-63.
[34] L.A.Zadeh, “Knowledge Representation in Fuzzy Logic”, IEEE Transactions on Knowledge and Data Engineering, vol.1, No 1, 1989, pp.89-100.
[35] L.A.Zadeh, “Fuzzy Logic, Neural Network and Soft Computing”, Communications of the ACM, vol.37, No 3, 1994, pp.77-84.
[36] L.A.Zadeh, “Soft Computing and Fuzzy Logic”, IEEE Software, vol.11, No 6, 1994, pp.48-58.
[37] L.A.Zadeh, “Fuzzy Logic=Computing With Words”, IEEE Transactions on Fuzzy Systems, vol.4, 1996, pp.103-111.
[38] L.A. Zadeh, “Fuzzy Sets and Information Granularity”. Advances in Fuzzy Sets Theory and Applications, M.M.Gupta, R.K.Ragade, R.R.Yager (Eds.). Amsterdam: North-Holland,1979, pp.3-20.
[39] L.A. Zadeh, “Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic”, Fuzzy Sets and Systems, vol.90, 1997, pp.111-127.
[40] L.A. Zadeh, “Toward a Logic of Perception Based on Fuzzy Logic”, Discovering the World with Fuzzy Logic. V.Novak and I.Perfilieva (Eds.). Heidelberg: Physica-Verlag, 2000, pp.4-28.
[41] L.A.Zadeh, “A New Direction in AI: Toward a Computational Theory of Perceptions”, AI Magazine, vol.22, No 1, 2001, pp.73-84.
[42] L.A.Zadeh,“From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions”, Computing With Words, P.P.Wang (Ed.). – New York: Wiley and Sons, 2001, pp.35-68.
[43] L.A.Zadeh, “Toward a Generalized Theory of Uncertainty (GTU) – An Outline”, Information Sciences, vol.172, 2005, pp.1-40.
[44] L.A.Zadeh, ‘Fuzzy Logic –a New Direction. The Concept of f-Validity”, New Dimensions in Fuzzy Logic and Related Technologies, Proceedings of the 5th EUSFLAT Conference (Osrava, Czech Republic, September 11-14, 2007), vol.1, p.29.
[45] L.A.Zadeh, “Toward Extended Fuzzy Logic – A First Step”, Fuzzy Sets and Systems, vol.160, 2009, pp.3175-3181.
[46] L.A.Zadeh, “Fuzzy Logic: a Personal Perspective”, Fuzzy Sets and Systems, vol.281, 2015, pp.4-20.
[47] L.A.Zadeh, “Toward a Restriction-Centered Theory of Truth and Meaning”. Information Sciences, vol.248, 2013, pp.1-14.
[48] L.A.Zadeh, “A Note on Z-numbers”, Information Sciences, vol.181, 2011, pp.2923-2932.
[49] A.N.Averkin, I.Z. Batyrshin, A.F. Blishun, V.B. Silov, V.B. Tarasov, Fuzzy sets in control models and artificial intelligence. Moscow, Nauka, 1986.
[50] E.Kerre, “The Impact of Fuzzy Set Theory on Contemporary Mathematics”. International Journal of Applied and Computational Mathematics, vol.10, No 1, 2011, pp.20-34.

ВСПОМИНАЯ ЛОТФИ ЗАДЕ

Поспелов Д.А, Стефанюк В.Л, Аверкин А.Н., Батыршин И.З., Тарасов В.Б., Ярушкина Н.Г., Язенин А.В.

Статья коллектива авторов посвящена памяти профессора Лотфи Заде – выдающегося ученого, внесшего крупный вклад в развитие теории управления и прикладной математики, информатики и искусственного интеллекта.

Лотфи Заде входил в когорту немногочисленных ученых-первооткрывателей, которые генерируют новые, оригинальные научные идеи и формируют базовые научные парадигмы, изменяющие наш мир. Он был основоположником теории нечетких множеств и лингвистических переменных, «отцом» нечёткой логики и приближённых рассуждений, родоначальником теории возможности и обобщенной теории неопределённости, создателем концепций гранулярных и мягких вычислений, вычислений со словами и перцептивными оценками, автором теории Z-чисел.

Его идеи и теории не только открыли новую эпоху в развитии научной мысли, свободную от ограничений узких научных направлений и способствующую их синергизму. Они внесли весомый вклад в развитие новых информационных и когнитивных технологий, привели к созданию эффективных промышленных технологий, таких как нечёткие компьютеры и процессоры, нечёткие регуляторы, нечёткие системы кластеризации и распознавания, и многие другие.

Роль Л.Заде в современном мире не ограничивается его достижениями в области науки и технологий. Его необычная собственная биография, вся творческая жизнь и международная деятельность внесли неоценимый вклад в формирование единого общепланетарного научного сообщества в эру «сетевого интеллекта», способствовали становлению нового научно-технического мировоззрения, предполагающего симбиоз восточных и западных культур.

Авторы были лично знакомы и нередко встречались с «Отцом нечеткой логики». Поэтому они стремились не только описать основные этапы жизненного пути и научного творчества Лотфи Заде, но и рассказать о многих обстоятельствах и эпизодах его жизни, связанных с Советским Союзом и Россией, а также поделиться воспоминаниями о личных встречах с великим ученым и замечательным человеком.

Памяти Юрия Роландовича Валькмана

Юрий Роландович Валькман
(26 марта 1948 г. – 16 апреля 2017 г.)

16 апреля 2017г. в Киеве на 70-м году ушел из жизни наш друг и коллега, известный советский и украинский ученый в области информатики и искусственного интеллекта (ИИ), заведующий отделом распределенных интеллектуальных систем Международного научно-учебного центра информационных технологий и систем Национальной академии наук Украины и Министерства образования и науки Украины, профессор кафедры математических методов системного анализа Национального технического университета Украины «Киевский политехнический институт (НТУУ КПИ) им. Игоря Сикорского», член программного комитета минских конференций OSTIS, доктор технических наук, профессор Юрий Роландович Валькман.

Он был ведущим специалистом Украины в области интеллектуальных систем и технологий,вице-президентом Международной ассоциации создателей и пользователей интеллектуальных систем (АСПИС), членом Советской, а затем Российской ассоциации искусственного интеллекта (РАИИ) и Международной ассоциации когнитивных исследований.

Юра Валькман родился 26 марта 1948г. в небольшом городе Пайде в Центральной части Эстонии. Окончил в 1971 г. мехмат Киевского государственного университета им. Т.Г.Шевченко. После службы в армии был принят на работу в Институт кибернетики им. В.М.Глушкова в отдел, которым руководил заместитель директора Института по науке, д.т.н. (будущий академик) Владимир Ильич Скурихин. Этот отдел занимался проблематикой разработки систем автоматизированного проектирования сложных объектов и автоматизированной обработки экспериментальных данных. Соответственно, первой научной темой Юрия Валькмана стала обработка результатов испытаний сложных объектов.

В 1986г. Ю.Р.Валькман защитил кандидатскую диссертацию по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин и систем, а в 1997 году в Твери – докторскую диссертацию на тему «Исследование сложных объектов в модельно-параметрическом пространстве» (по специальности 05.13.16).

В 1980-1990-е годы принимал участие в разработке крупных программно-информационных комплексов систем автоматизации исследований и проектирования сложных объектов новой техники: ТЕМП – система обработки результатов лётных испытаний (г. Жуковский), ГЕЛИОГРАФ – автоматизированная система научных исследований гидрофизических объектов (г.Севастополь), ЧЕРТЁЖ – система автоматизации исследовательского проектирования сложных изделий кораблестроения (г.Ленинград), ИЗИН – система управления базами данных, ориентированная на обработку результатов исследований (г.Тверь).

В Институте кибернетики Ю.Р.Валькман проработал около 25 лет вплоть до его структурной реорганизации в 1990-е годы и переориентации прикладной тематики. Вместе с академиком В.И.Скурихиным он перешел в образованный в 1997г. МНУЦ информационных технологий и систем НАН Украины и МОН Украины, где получил должность заведующего новым Отделом распределенных интеллектуальных систем. Работал в составе редколлегий ведущих научных журналов: «Программные продукты и системы», «Интеллектуальные системы и технологии», «Онтология проектирования».

Автор более 300 научных публикаций, в том числе 5 монографий, 9 учебно-методических пособий. Кратко опишем области его исследовательских интересов и основные компоненты научного вклада. Среди его лучших книг: «Информационные технологии в испытаниях сложных объектов: методы и средства» [1*]; «Интеллектуальные технологии исследовательского проектирования: формальные системы и семиотические модели» [2*], «Модельно-параметрическое пространство: теория и применение» [3*].

Пожалуй, лучшая монография Ю.В.Валькмана [2*] посвящена проблемам создания новых компьютерных технологий проектирования сложных объектов. Путем интеграции формальных систем и семиотических моделей разработаны методологические и теоретические основы создания баз знаний проектировщиков, конструкторов исследователей сложных изделий новой техники. Особое внимание уделено формальным конструкциям параметров и моделей для разработки «исчисления обликов сложных объектов». Рассмотрены проблемы, принципы и методы отчуждения трудноформализуемых знаний проектантов, предложены варианты использования идей совмещенной разработки (Concurrent Engineering) в исследовательском проектировании [7]. Выделены основные НЕ-факторы (в смысле А.С.Нариньяни) на ранних этапах проектирования и описаны математические средства их моделирования.

В книге [3*] с целью построения мультимодельной, гетерогенной, многомерной, сложноструктурированной, семантически насыщенной вычислительной среды введен и исследован новый формальный аппарат, названный модельно-параметрическим пространством. Определены меры совместимости и согласованности моделей знаний в этом пространстве.

Юрий Роландович всегда был инициатором и сторонником тесного сотрудничества между российскими и украинскими учёными. Он активно участвовал в научных мероприятиях РАИИ, в том числе почти во всех Национальных конференциях по ИИ: от КИИ-1992 до КИИ-2016 (см. его работы [5],[7],[13],[14],[17],[20],[21],[22],[29],[40],[46],[47],[49],[50]), делал пленарные доклады на международных научно-практических конференциях «Интегрированные модели и мягкие вычисления» в Коломне (рис.1) [15],[26],[37],[38],[48],[51], выступал на крупных российских конференциях по компьютерной лингвистике и интеллектуальным технологиям «Диалог» [34], а также был одним из главных организаторов международных конференций «Интеллектуальный анализ информации» им. Т.А.Таран (рис.2), проводившихся в Киеве в НТУУ КПИ (см. работы [16],[18],[39],[41],[44], [46], [47], [52]).

Юра (именно так, без всяких церемоний, его всегда звали друзья и просто знакомые) пришёл в ИИ как сложившийся специалист в области автоматизации исследовательского проектирования и испытаний сложных технических объектов, таких как корабли [1],[10] и самолеты [2]. Поэтому примерно до середины 1990 гг. большинство его публикаций носили преимущественно прикладной характер: им велись поиск и разработка новых компьютерных методов и информационных технологий в целях интеллектуализации проектирования, испытаний и других этапов жизненного сложных изделий.

Научные интересы Ю.Р.Валькмана прошли несколько этапов развития от автоматизированных систем обработки результатов испытаний сложных объектов [11],[12], средств интеллектуализации САПР [3],[4],[9] и АСНТИ [12] до проблем моделирования образного мышления и интуиции, когнитивной графики [29], [30] и когнитивной семиотики [46], [48].

Знакомство в 1990-х годах с трудами Д.А.Поспелова по прикладной семиотике [1d],[2d], монографией А.А.Зенкина по когнитивной графике [3d], идеями НЕ-факторов А.С.Нариньяни [4d], обсуждение тенденций развития ИИ как в кулуарах КИИ, так и с крупнейшими украинскими учеными, работавшими в этой области или смежных дисциплинах – А.А.Летичевским, В.П. Гладуном, З.Л.Рабиновичем, Т.А.Таран, и др. – привели к зарождению у Юры глубокого интереса к ряду теоретических областей ИИ, группируемых вокруг проблемы моделирования образного мышления и интуиции человека. Особое место в этом ряду занял научный семинар «Отражение образного мышления и интуиции специалиста в системах искусственного интеллекта» и доклад Д.А.Поспелова [5d], который стал решающим толчком.

Примерно через полгода после КИИ-2014 (рис.3), 27-28 марта 2015г., в Казани прошла научная конференция в области прикладной семиотики и моделирования в сфере гуманитарных наук. Она была посвящена 60-летию академика АН РТ Д.Ш. Сулейманова и 5-летию Института «Прикладная семиотика» АН РТ. На ней Ю.Р. Валькман выступил с пленарным докладом. По итогам конференции была опубликована коллективная монография с его участием [4*].

Так у Юры сформировались новые научные интересы в области ИИ, которые указаны на обложках его книг и статей конца 1990-х годов и 2000-х годов: когнитивная графика и прикладная семиотика, языки образного мышления и когнитивное моделирование, НЕ-факторы знаний [26],[27],[28]и когнитивная семиотика[4*],[46],[48], одним из основателей которой он является.

В 2004г. вышел в свет журнал РАИИ «Новости искусственного интеллекта», №2, содержавший рубрику «Моделирование НЕ-факторов–ключевое направление ИИ в начале XXI-го века». После работы А.С. Нариньяни «НЕ-факторы: краткое введение» была опубликована замечательная статья Ю.В.Валькмана [27], где были рассмотрены методы моделирования НЕ-факторов, показана их роль в образном мышления, предложена классификация НЕ-факторов в систем координат «НЕ-факторы – методы – объекты». В заключении содержится «основной вопрос» искусственного интеллекта начала XXI-го века: Может ли искусственная система называться интеллектуальной, если она не моделирует какие-либо НЕ-факторы?

Уже будучи тяжело больным, он организовал Круглый стол по когнитивной семиотике на КИИ-2016 в Смоленске (рис. 4) и принял участие в съезде РАИИ (рис. 5).

Свою первую работу по когнитивной семиотике Юра опубликовал в 2012г. [46]. Когнитивная семиотика – это наука о методах получения знаний из знаковых струк-

Рис. 1. Общее фото перед прогулкой на теплоходе по Оке. Ю.Р.Валькман среди участников научно-практической конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте». Коломна, 28 мая 2009 г.

Рис. 2. После конференции «Интеллектуальный анализ информации–2010. Ю.Р.Валькман и В.Б.Тарасов в музее под открытым небом «Киев в миниатюре». Киев, 19 мая 2010г.

Рис. 3. На конференции КИИ-2014 в Казани: Ю.В.Валькман, Д.Ш.Сулейманов, А.П.Еремеев.

тур и, наоборот, методах синтеза знаковых конструкций, представляющих структуры знаний. В основе этих методов лежит идея взаимодействия трех миров (согласно эволюционной эпистемологии К.Поппера) – реального, мира знаков и ментального мира. В рамках когнитивной семиотики может рассматриваться весь спектр проблем синтеза и анализа смыслов, определения значений разных знаковых структур, понимания ситуативного поведения людей и их коллективов, познания мира в целом.

Профессор Ю.Р.Валькман был ярким, замечательным педагогом, воспитавшим многие поколения студентов в НТУУ КПИ, Киевском национальном университете им. Тараса Шевченко, НУ «Киево-Могилянская академия». Его лекционная нагрузка была просто колоссальной. Даже в последние месяцы жизни он продолжал читать свои лекции в удаленном режиме по интернету и завоевал у ничего не подозревавших о его тяжелом состоянии студентов репутацию «продвинуто-

Рис. 4. Юрий Роландович Валькман выступает на Круглом столе по когнитивной семиотике на КИИ-2016. Смоленск, 7 октября 2016г.

го» профессора.

Юрий Роландович Валькман был, наверное, главным историографом РАИИ, на протяжении многих лет неустанно фиксировавшим на видео все наши основные события.

Рис. 5. На выборах исполкома и научного совета РАИИ Юрий Роландович почти всегда был «главным» – председателем счетной комиссии. Смоленск, 5 октября 2016г.

Мы хорошо помним его, сидящим с видеокамерой (рис. 6) на многих конференциях РАИИ. Благодаря ему, мы можем найти сегодня на YouTube и других ресурсах доклады и лекции ведущих учёных по искусственному интеллекту.

Последний приезд Юры в Минск состоялся в августе 2017г. Мы предложили ему принять участие в совместном проекте, и он без колебаний согласился. Потом была поездка в Несвижский замок, прогулка

Рис. 6. Юрий Роландович Валькман отснял десятки видеокассет с лекциями и докладами ведущих ученых РАИИ

к знаменитому памятнику собаке, спасшей хозяину замка жизнь во время охоты.

Юра был добрым, увлекающимся и жизнерадостным человеком, имевшим множество знакомых в разных «тусовках». В период советской юности он занимался горным туризмом, был инструктором и исходил весь Крым. Женившись на Лилии Исмагиловой – талантливой женщине, хорошо известной любителям бардовской песни, он быстро стал своим в их особой музыкальной среде. В зрелом возрасте он увлекся путешествиями и изъездил полмира. Уже в 2000-е годы он вместе с любимой женой «заболел» удивительными красотами островов Таиланда, настоящими фанатами которых они оказались. Но, главным в его жизни, была наука и общение с коллегами на конференциях, которое вдохновляло его продолжать столь интенсивную, но любимую работу.

Мы глубоко скорбим о смерти известного ученого, замечательного педагога, светлого и хорошего человека. Нам будет очень не хватать тебя, Юрий Роландович!

Научные труды Ю.Р. Валькмана

I. Monographs

1* Skurikhin, V.I., Kvachev, V.G., Val'kman, Yu.R. Informatsionnye tekhnologii v ispytaniyakh slozhnykh ob"ektov: metody i sredstva [Information Technologies in complex objects testing: methods and means], Kiev, Naukova dumka, 1990. 316 p.

2* Val'kman Yu.R. Intellektual'nye tekhnologii issledovatel'skogo proektirovaniya: formal'nye sistemy i semioticheskie modeli [Intelligent technologies of research design: formal systems and semiotic models], Kiev, Port-Royal, 1998. 250 p.

3* Val'kman Yu.R., Gritsenko, V.I., Rykhalsky A.Yu. Model'no-parametricheskoe prostranstvo: teoriya i primenenie [Model-parameter space: theory and applications], Kiev, Naukova dumka, 2012. 190 p.

4* Kognitivno-semioticheskie aspekty modelirovaniya v gumanitarnoi sfere. Kollektivnaya monografiya [Cognitive and semiotic aspects of modeling in humanitarian area. Collective monograph] V.L. Stefanyuk, E.A. Taisina (Eds.), Kazan, Izdatelstvo AN RT, 2017. 346 p.

II. Automated processing of test results. Research design. Intelligent CADs

1 Val'kman, Yu.R. Informatsionnoe obespechenie avtomatizirovannykh sistem obrabotki rezul'tatov kompleksnykh ispytanii novykh tekhnicheskikh ob"ektov [Infoware of automated

systems for processing complex testing results of the new technical objects]. Voprosy sudostroeniya. Ser. Matematicheskie metody, programmirovanie, ekspluatatsiya EVM [Shipbuilding problems. Ser. Mathematical methods, programming, computer usage], 1982, no. 28, pp. 54-68.

2 Val'kman, Yu.R. IZIN – sistema upravleniya bazami dannykh, orientirovannaya na obrabotku rezul'tatov ispytanii [IZIN - a database management system for test result processing]. SUBD i pakety okruzheniya: problemy razrabotki i primeneniya [DBMS and environment packages: development and applications], 1985, pp. 86-92.

3 Val'kman, Yu.R. Ischislenie modelei – osnova intellektualizatsii protsessov issledovatel'skogo proektirovaniya [Model calculus as the basis for intellectualizing of the research design processes]. Programmnye produkty i sistemy [Software and systems], 1995, vol. 1, no. 4, pp. 18-23.

4 Val'kman, Yu.R. Model'no-parametricheskoe prostranstvo v issledovatel'skom proektirovanii: tseli postroenii, opredeleniya, struktura i svoistva [Model-parametric space in research design: construction goals, definitions, structure and properties]. Voprosy kognitivno-informatsionnoi podderzhki postanovki i resheniya novykh nauchnykh problem [Questions of cognitive and informational support for the formulation and solution of new scientific problems], 1995, pp. 103-115.

5 Val'kman, Yu.R. Videoobrazy v operatsiyakh issledovatel'skogo proektirovaniya: otnosheniya mezhdu abstraktnym i konkretnym, logichnym i metaforichnym, ob"ektivnym i sub"ektivnym, informatsionnym i kognitivnym [Video images in research design operations: relations between abstract and concrete, logical and metaphorical, objective and subjective, informational and cognitive]. Natsional'naya konferentsiya po II [National AI Conference], 1996, vol.3, pp. 389-393.

6 Val'kman, Yu.R. O probleme «otchuzhdeniya modelei issleduemykh ob"ektov ot sozdatelei» v proektirovanii slozhnykh izdelii [On the problem of "alienation of models of studied objects from creators"in the design of complex products]. Teoriya i sistemy upravleniya [Control systems and theory], 1996, no.3, pp.116-152.

7 Val'kman, Yu.R. Ot modelei zhiznennykh tsiklov k S-tekhnologiyam [From lifecycle models to S-technologies]. Natsional'naya konferentsiya po II [National AI Conference], 1998, vol.2, pp. 671-676.

8 Val'kman, Yu.R., Rykhalsky A.Yu. Printsipy postroeniya ischisleniya oblikov slozhnykh izdelii [Principles of complex shapes appearance calculus]. Mezhdunarodnyi kongress «Iskusstvennyi intellekt v XXI-m veke» [International Congress "Artificial Intelligence in the 21st Century"], 2001, vol. 1, pp. 432-440.

9 Val'kman, Yu.R., Rykhalsky A.Yu. «Intellektual'nyi datchik» v issledovatel'skom proektirovanii slozhnykh ob"ektov ["Intelligent sensor"in the research design of complex objects].

10 Val'kman, Yu.R., Solomakha O.N., Suvorov A.I. Informatsionnye tekhnologii edinogo kompleksa issledovanii v voennom korablestroenii [Information technologies of a single complex of research in military shipbuilding]. Programmnye produkty i sistemy [Software and systems], 1993, no. 4, pp. 10-19.

11 Val'kman, Yu.R., Stepashko V.S. Printsipy postroeniya ekspertnykh sistem matematicheskogo modelirovaniya dlya ispytanii slozhnykh ob"ektov [Principles of constructing expert systems of mathematical modeling for testing complex objects]. Prikladnaya informatika [Applied informatics], 1990, no. 16, pp. 129-143.

12 Val'kman, Yu.R., Suvorov A.I. Problemy intellektualizatsii informatsionnykh tekhnologii nauchnykh issledovanii slozhnykh ob"ektov [Problems of Intellectualization of Information Technologies for Scientific Research of Complex Objects], Kiev, Znanie, 1990.

III. Knowledge bases, ontologies and models of understanding

13 Val'kman, Yu.R. Printsipy postroeniya algebry i logiki tekstov i kontekstov matematicheskikh modelei [Principles of constructing algebra and logic of texts and contexts of mathematical models]. Natsional'naya konferentsiya s mezhdunarodnym uchastiem «Iskusstvennyi intellekt-92» [National conference with international participation "Artificial Intelligence-92"], 1992, vol. 1, pp. 48-53.

14 Val'kman, Yu.R. Model'no-parametricheskoe prostranstvo: predstavlenie znanii ob issleduemykh protsessakh i ob"ektakh [Model-Parametric Space: Representation of Knowledge of Processes and Objects Being Researched]. Natsional'naya konferentsiya po II [National AI Conference], 1996, vol. 2, pp. 299-303.

15 Val'kman, Yu.R. Mul'timodal'nost', mul'timediinost', mul'tikontekstnost', mul'timodel'nost' predstavleniya ponyatii v bazakh znanii [Multimodality, multimedia, multi-context, multi-model representation of concepts in knowledge bases]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2007, vol. 1, pp. 76-83.

16 Val'kman, Yu.R., Val'kman, R.Yu. Biznes-intellekt i upravlenie znaniyami: ponyatiya i tekhnologii [Business Intelligence and Knowledge Management: Concepts and Technologies]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2010, pp. 28-35.

17 Val'kman, Yu.R., Val'kman, R.Yu. Ontologii v modelirovanii obraznogo myshleniya [Ontologies in the modeling of figurative thinking]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National Conference on Artificial Intelligence with International Participation], 2012, vol.1, pp. 310-317.

18 Val'kman, Yu.R., Val'kman, R.Yu. Kontsepty: opredelenie, struktura, klassifikatsiya [Concepts: definition, structure, classification]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2013, pp. 82-94.

19 Val'kman, Yu.R., Ismagilova, L.R. Intellektual'nye sistemy: o modelirovanii ponimaniya [Intelligent systems: on modeling understanding]. «Intellektual'nye sistemy - 2004» i «SAPR-2004» [Intelligent Systems-2004 and CAD-2004], 2004, vol. 1, pp. 160-167.

20 Val'kman, Yu.R., Ismagilova, L.R., Rykhalsky A.Yu. Intellektual'nye sistemy: ot tselenapravlennosti k tseleustremlennosti [Intelligent systems: from focus to purposefulness]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National conference on artificial intelligence with international participation], 2004, vol. 2, pp. 767-775.

21 Val'kman, Yu.R., Lozovoi, V.V. Ot sistematizatsii klassifikatorov k ontologii informatsionno-analiticheskikh sistem statisticheskogo monitoringa [From the classification of classifiers to the ontology of information-analytical systems of statistical monitoring]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National conference on artificial intelligence with international participation], 2002, vol. 2, pp. 606-615.

22 Val'kman, Yu.R., Maistrenko, S.A. Metaznaniya – sredstvo effektivnogo upravleniya raspredelennymi znaniyami intellektual'nykh sistem [Meta-knowledge as a means of effectively managing the distributed knowledge of intelligent systems]// Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National conference on artificial intelligence with international participation], 2006, vol. 3, pp. 1062-1070.

23 Val'kman, Yu.R., Rykhalsky A.Yu. Analiz metodov i sredstv predstavleniya znanii proektirovshchikov slozhnykh tekhnicheskikh ob"ektov [Analysis of methods and means of representing knowledge of designers of complex technical objects]. «Intellektual'nye sistemy - 2004» i «SAPR-2004» [Intelligent Systems-2004 and CAD-2004], 2004, vol. 1, pp. 167-174.

24 Val'kman, Yu.R., Rykhalsky A.Yu. Model'no-parametricheskoe prostranstvo – sredstvo predstavleniya znanii issledovatelei slozhnykh sistem [Model-parametric space - a means of representing the knowledge of researchers of complex systems]. Upravlyayushchie sistemy i mashiny [Control systems and machines], 2009, no. 1, pp. 20-30.

25 Val'kman, Yu.R., Stepashko, P.V. Ob ontologii intellektual'nogo modelirovaniya [On the Ontology of Intellectual Modeling]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2016, pp. 165-170.

IV. Uncertainty modeling. UN-factors

26 Val'kman, Yu.R. NE-faktory – osnova obraznogo myshleniya [UN-factors as the basis of figurative thinking]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2003, pp. 26-33.

27 Val'kman, Yu.R. Modelirovanie NE-faktorov: osnova intellektualizatsii komp'yuternykh tekhnologii [UN-factors modeling as a basis of intellectualization of computer technologies]. Novosti iskusstvennogo intellekta [Artificial intelligence news], 2004, no. 2, pp. 64-81.

28 Val'kman, Yu.R., Gaevoi S.A. NE-faktory v issledovatel'skom proektirovanii slozhnykh ob"ektov [UN-factors in the research design of complex objects]. Mezhdunarodnyi seminar po komp'yuternoi lingvistike i ee prilozheniyam «Dialog» [International seminar on computer linguistics and its applications "Dialog"], 1998, vol. 2, pp. 649-659.

V. Cognitive graphics and graphical images

29 Val'kman, Yu.R. Graficheskaya metafora – osnova kognitivnoi grafiki [Graphical metaphor - the basis of cognitive graphics].

Natsional'naya konferentsiya po II [National AI Conference], 1994, vol. 1, pp. 9-18.
30 Val'kman, Yu.R. Kognitivnye graficheskie metafory: kogda, zachem, pochemu i kak my ikh ispol'zuem [Cognitive graphical metaphors: when, why, why and how we use them]. Mezhdunarodnaya konferentsiya «Znaniya – dialog – reshenie» [International Conference "Knowledge - Dialogue - Solution"], 1997, vol.2, pp. 261-272.
31 Val'kman, Yu.R. Kniga, Yu.N. Analiz ponyatiya «graficheskii obraz» [Analysis of the "graphical image"concept]. Mezhdunarodnyi seminar po komp'yuternoi lingvistike i ee prilozheniyam «Dialog» [International seminar on computer linguistics and its applications "Dialog"], 2002, pp. 41-52.

VI. Modeling of figurative thinking

32 Val'kman, Yu.R. Kategorii «obraz» i «model'» v kognitivnykh protsessakh ["Image"and "model"categories in cognitive processes]. Mezhdunarodnaya konferentsiya «Intellektual'nye sistemy» [International Conference "Intelligent Systems"], 2003, vol. 2, pp. 318-323.
33 Val'kman, Yu.R. Konteksty v protsessakh obraznogo myshleniya: opredeleniya, otnosheniya, operatsii [Contexts in the processes of figurative thinking: definitions, relations, operations]. Rossiiskaya konferentsiya po kognitivnoi nauke [Russian Conference on Cognitive Science], 2004.
34 Val'kman, Yu.R., Ismagilova, L.R. O yazyke obraznogo myshleniya [On the language of figurative thinking]. Komp'yuternaya lingvistika i intellektual'nye tekhnologii. Trudy mezhdunarodnogo seminara Dialog'2004 [Computer linguistics and intellectual technologies. International seminar Dialog'2004], 2004.
35 Val'kman, Yu.R. Analiz ponyatiya «obraz»: otnosheniya «obrazy – ponyatiya» [Analysis of the concept of "image": the relationship "images - concepts"]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National conference on artificial intelligence with international participation], 2008. vol. 1, pp. 369-377.
36 Val'kman, Yu.R. Modelirovanie obraznogo myshleniya: ot obraza k ponyatiyu i ot ponyatiya k obrazu [Modeling of figurative thinking: from image to concept and from concept to image]. Mezhdunarodnye nauchno-tekhnicheskie konferentsii «Intellektual'nye sistemy-2008» i «Intellektual'nye SAPR-2008» [International scientific and technical conferences "Intelligent systems-2008"and "Intelligent CAD-2008"], 2008, vol. 2, pp. 151-159.
37 Val'kman, Yu.R. Obrazy i obraznoe myshlenie: nekotorye otnosheniya i struktury [Images and imaginative thinking: some relationships and structures]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2009, vol. 1, pp. 109-120.
38 Val'kman, Yu.R. Dinamicheskaya struktura obraza: doformal'noe issledovanie [Dynamic image structure: preformal research]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2011. vol. 1, pp. 136-148.
39 Val'kman, Yu.R. O strukture obraza: doformal'noe issledovanie [On the structure of the image: a preformal study]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2010, pp. 17-27.
40 Val'kman, Yu.R. Struktura obraza: doformal'noe issledovanie [Image structure: preformal research]. Natsional'naya konferentsiya po II [National AI Conference], 2010, vol. 1, pp. 344-352.
41 Val'kman, Yu.R. O modelirovanii smysla, ponimaniya, tselostnosti obraza [On modeling the sense, understanding, integrity of the image]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2011, pp. 170-179.
42 Val'kman, Yu.R., Bykov V.S. O modelirovanii obraznogo myshleniya v komp'yuternykh tekhnologiyakh: obshchie zakonomernosti myshleniya [On modeling of figurative thinking in computer technologies: general patterns of thinking]. Mezhdunarodnaya konferentsiya «Znaniya – dialog – reshenie» [International Conference "Knowledge - Dialogue - Solution"], 2005.
43 Val'kman, Yu.R., Bykov V.S. O modelirovanii obraznogo myshleniya v komp'yuternykh tekhnologiyakh: operatsii myshleniya [On modeling of figurative thinking in computer technologies: operations of thinking]. Mezhdunarodnye nauchno-tekhnicheskie konferentsii «Intellektual'nye sistemy-2005» i «Intellektual'nye SAPR-2005» [International scientific and technical conferences "Intelligent systems-2005"and "Intelligent CAD-2005"], 2005, vol. 1.
44 Val'kman, Yu.R., Valkman R.Yu. Dynamic image structure [Dinamicheskaya struktura obraza]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2011, pp. 160-169.
45 Val'kman, Yu.R., Valkman R.Yu. Tselostnost' obrazov: o modelirovanii smysla, ponimaniya, znacheniya [Image integrity: on modeling the sense, understanding, meaning]. Kongress po intellektual'nym sistemam i informatsionnym tekhnologiyam [Congress on Intelligent Systems and Information Technologies], 2011. vol.1, pp. 497-509.

VII. Cognitive semiotics

46 Val'kman, Yu.R. O kognitivnoi semiotike [On cognitive semiotics]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2012. pp. 19-30.
47 Val'kman, Yu.R. Kognitivnost' semiotiki [Cognitivity of semiotics]. Intellektual'nyi analiz informatsii [Intelligent information analysis], 2013. pp. 69-81.
48 Val'kman, Yu.R. Kognitivnaya semiotika: istoki i perspektivy [Cognitive semiotics: origins and perspectives]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2013, vol. 1, pp. 48-61.
49 Val'kman, Yu.R. Kognitivnaya semiotika: semiozisy i geshtal'ty [Cognitive semiotics: semiosis and gestalt]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National Conference on Artificial Intelligence with International Participation], 2014, vol. 1, pp. 180-188.
50 Val'kman, Yu.R. Kognitivnaya semiotika: geshtal'ty i znaki, tselostnost' i struktura [Cognitive semiotics: gestalt and signs, integrity and structure]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National Conference on Artificial Intelligence with International Participation], 2016. T.2. C.250-258.
51 Val'kman, Yu.R. Kognitivnaya semiotika: geshtal'ty i metafory [Cognitive semiotics: gestalt and signs, integrity and structure]. Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2015. vol.1, pp. 54-59.
52 Val'kman, Yu.R. Geshtal'ty i metafory v kognitivnoi semiotike [Gestalts and metaphors in cognitive semiotics]// Intellektual'nyi analiz informatsii [Intelligent information analysis], 2015, pp. 31-39.

VIII. Works on new areas of AI that affected Yu.R. Valkman's scientific interests

1d Pospelov, D.A. Prikladnaya semiotika i iskusstvennyi intellekt [Applied semiotics and artificial intelligence]. Programmnye produkty i sistemy [Software and systems], 1996, no. 3, pp. 10-13.
2d Pospelov, D.A., Osipov, G.S. Prikladnaya semiotika (chast' 1) [Applied semiotics (part 1)]. Novosti iskusstvennogo intellekta [Artificial intelligence news], 1999, no. 1, pp. 9-35.
3d Zenkin, A.A. Kognitivnaya komp'yuternaya grafika [Cognitive computer graphics], Moscow, Nauka, 1991.
4d Narin'yani A.S. NE-faktory i inzheneriya znanii: ot naivnoi formalizatsii k estestvennoi pragmatike [UN-factors and knowledge engineering: from naive formalization to natural pragmatics]. Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem [National Conference on Artificial Intelligence with International Participation], 1994. vol. 1, pp. 9-18.
5d Pospelov, D.A. Metafora, obraz i simvol v poznanii mira [Metaphor, image and symbol in the knowledge of the world]. Novosti iskusstvennogo intellekta [Artificial intelligence news], 1998, no. 1, pp. 94-136.

IN MEMORIAM OF YURI ROLANDOVICH VALKMAN (26.III.1948 – 16.IV.2017)

On April 16th, 2016 in Kiev, our dear friend and colleague, the prominent Soviet and Ukrainian scientist in the areas of information science and artificial intelligence, head of the Distributed intelligent systems department at the International scientific and educational centre of information technologies and systems of the Ukrainian National Academy of Sciences and Ukrainian Ministry of Education, professor of the Mathematical methods of systems analysis department at the Ukrainian National Technical University «Kiev Polytechnical University n.a. Igor Sikorskiy», program committee member of the OSTIS conferences, doctor of technical sciences, professor Yuri Rolandovich Valkman, has passed away in his late sixties.

Information Granulation, Cognitive Logic and Natural Pragmatics for Intelligent Agents

Valery B.Tarassov
CIM Department, Bauman Moscow State Technical University
Moscow, Russia
Vbulbov@yahoo.com

***Abstract*—The paper is devoted to the development of Cognitive Logic in the framework of building intelligent agents. The drawbacks of classical mathematical logic and automated reasoning are discussed. The difference between classical logic and human cognition is shown on simple examples. The concept of cognitive agent, in particular, cognitive robot, is considered, its architecture is presented. Information granulation based on pragmatics is viewed as a principal capacity of cognitive agent. The role of logical pragmatics in cognitive logic is revealed. The emphasis is made on the development of generalized logical values and formation of various logical worlds to construct granular logical semantics and pragmatics. An extended definition of logical world is proposed. In the context of developing cognitive graphics for applied logics, the colored representation of Hasse diagrams is suggested. Possible applications of colored representation of logical worlds are discussed.**

***Keywords*—Artificial Intelligence; Intelligent Agent; Cognitive Logic; Pragmatics; Pragmatic Truth, Logical World, Cognitive Graphics, Color Metaphor**

I. INTRODUCTION

One of the main trends in the development of new generation technologies in the XXI century is the formation of hybrid systems combining advanced information, cognitive and social technologies with biotechnologies and nanotechnologies in the scope of NBICS convergence conception [1]. Up to now, cognitive technologies remain a «bottleneck» in NBICS complex, and the creation of new cognitive microsciences (see[2,3]), such as cognitive graphics, cognitive linguistics, cognitive semantics, cognitive semiotics, cognitive informatics, seems to be a necessary step on the way to autonomous artificial cognitive agents, both individual and collective.

This paper discusses the prospects of Cognitive Logic for Intelligent Agents. It is a new trend in applied logic based on the characteristics of human cognition and developing new logical systems to support cognitive processes in agents. The term «Cognitive Logic» was launched into circulation by Pei Wang [4], who constructed a Non-Axiomatic Reasoning System (NARS) and proposed an experience-grounded semantics. An early precursor of cognitive logics was D.A.Pospelov, the author of pseudo-physical logics [5].

Rather close ideas were proposed by V.K.Finn [6] with his JSM-method, using four-valued argumentation logic, quasi-axiomatic theory and synthesis of various reasoning types, as well as O.M.Anshakov and T.Gergely [7], who introduced the procedures of «cognitive reasoning».

In our paper, the experience is associated with information granulation that is viewed as a crucial cognition and comprehension mechanism. So Zadeh's TFIG (Theory of Fuzzy Information Granulation) [8] and Lin's Granular Computing [9] are seen as a natural basis for Cognitive Logics. The emphasis is made on the development of generalized logical values and formation of various logical worlds to construct granular logical semantics and pragmatics. Our approach is based on D.A.Bochvar's thesis «from logical semantics to logical calculus».

The OSTIS project [10] has been initiated in order to develop open semantic technologies of designing intelligent systems. In this paper we suggest to complement it by open pragmatic technologies for intelligent agents, rising to the ideas of the «Father of Pragmatism» Ch.S. Peirce [11].

The paper is organized in the following way. The reasons for the emergence of Cognitive Logic are revealed in Section II. The structure and operation of classical automated reasoning system is considered in *Subsection A*. The limitations of traditional automated reasoning to compare with everyday human reasoning are shown on many examples in *Subsection B*. The problem of reasoning uncertainty is faced. In this context, the concept of Non-factors is discussed. A classification of Non-factors in knowledge engineering is given. In *Subsection C* the processes of human cognition and their properties are analyzed. Four basic types of cognitons – complex cognitive units, enabling self-organization of agent's activity – are presented.

The fundamentals of Agent Theory are presented in Section III. Firstly, in *Subsection A*, the concept of artificial agent is explained. Some interpretations, classifications and architectures of agents are introduced. The difference between reactive and intelligent agents is discussed. Secondly, in *Subsection B*, the notion of cognitive agent is specified. An example of cognitive robot is considered, specific features of its architecture are pointed out. Finally, an interactive model of robot's dialogue control is suggested.

A main capacity of cognitive agent is a goal-driven information granulation. In *Subsection C* some basic definitions and classifications of granules are given. Two general approaches to constructing granules are analyzed. In particular, non-classical sets are mentioned as rather new and convenient formalisms to create granules.

Section IV is devoted to pragmatics viewed as a keynote attribute of both agent's individual behavior and collective behavior of communicating cognitive agents. Specifically, logical pragmatics is seen as a necessary condition of understanding and applying logics by cognitive agents. Here logical granules are of special concern too. In *Subsection A* the difference between logical semantics and pragmatics is shown on the basis of both meta-logic and communication model. Some intrinsic links between logical pragmatics and pragmatic logics are investigated. Human cognition is based on the unity of descriptions and prescriptions. In this context, a complementary role of Aristotle-Tarski's correspondence truth theory and Peirce's vision of truth as utility (value) is demonstrated.

Subsection B contains some counter-arguments against the universal character of truth in knowledge engineering by cognitive agents. In *Subsection C* some new interpretations of truth values are gathered: from epistemic Dunn's vision of semantics leading to generalized (granular) truth-values to context-dependent considerations like factual truth, concerted truth and measured truth.

In Section II we put the question «Why Cognitive Logic?» and try to justify the relevance of this concept in the modeling of intelligent agents. At last, in Section V we outline a possible answer to the question «How it could be created?» by rethinking the concept of Logical World and applying the ideas of Cognitive Graphics. Some basic definitions of logical worlds and their representative examples are included into *Subsection A*. A visualization of logical worlds through colored logical values in Hasse diagrams is proposed in *Subsection B*. In our opinion, it opens new opportunities in building anthropomorphic interfaces between human and artificial cognitive agents by standardizing the interpretation of logical values used in different applications.

II. Why Cognitive Logic?

A. Classical Mathematical Logic and Automated Reasoning Systems

Reasoning is the ability to make inferences, and automated reasoning supposes the development of computing systems that automate this process. An automated reasoning system usually includes the following basic components [4]: 1) a formal language that represents knowledge; 2) a semantics that defines meaning and truth value in the language; 3) a set of inference rules to derive new knowledge; 4) a memory that stores knowledge;5) a control mechanism that selects premises and rules in each step. Here the former three components are usually related to a logic and form a logical part of reasoning system, and the latter two components responsible for an implementation of this logic are called the control part of the system.

At present, first-order predicate logic remains the basis for the logical part of automated reasoning, and the theory of computability and computational complexity is extensively used in the control part. In fact, these logical theories and tools have been successfully used in many practical domains. However, the continuation of their application in advanced intelligent systems such as cognitive agents and their groups seems very doubtful, due to some fundamental differences between automated reasoning and human cognition.

Classical automated reasoning is based on purely axiomatic systems, certainty conditions and deduction rules of traditional logic, where the truth of the premises guarantees the truth of the conclusion. Contrarily, human cognition and reasoning is deployed under uncertainty by using mainly non-deductive (common-sense) reasoning in semi-axiomatic or non-axiomatic systems. We will give below some examples to clarify the difference between deductive and non-deductive reasoning, as well as a short description of human cognition and its characteristics.

B. From Non-Deductuve to Uncertain Reasoning

Inference rules of classical logic are deduction rules, based in truth preservation and certain conclusion. In a sense, here the information in a conclusion is contained already in the premises, and the inference rule simply makes it explicit. For example, from «Crows are birds» and «Birds have feathers» it is valid to derive «Crows have feathers».

Meanwhile, in everyday life we often use other reasoning types, where the conclusions seem to carry new information not available in the premises. In case of *induction* a broad generalization is made from special cases. Let us use again the previous example. Here we take «Crows are birds» and «Crows have feathers» to derive «Birds have feathers». It is obvious that for inductive reasoning, even if all the premises are true, the conclusion can be false.

Further we consider *abductive reasoning* based on explanations for given case. Example: from «Birds have feathers» and «Crows have feathers» to conclude «Crows are birds».

Finally, *analogical reasoning* is a kind of similarity-based reasoning. Example: «Rooks are similar to crows» and «Crows have feathers», hence «Rooks have feathers».

So both inductive and abductive and analogical inference rules do not guarantee the truth of the conclusion for true premises. Therefore, they are not valid rules in the sense of classical logic. Nevertheless, all these types of inference are widely used in many branches, specifically, in learning and creative design.

Traditional formal theories of reasoning are certain in several aspects, whereas real-world human reasoning is often uncertain in these aspects. Now let us face the problem of reasoning uncertainty or, more generally, Non-factors of reasoning. What are Non-Factors? This is a variety of different factors, which are expressed by the words (linguistic labels) having some negative hints in natural language, remain largely unexplored in traditional mathematics, but are inherent attributes of human knowledge and cognition.

The term «Non-factors» was coined by A.S.Narinyani (see [12]) in early 1980's. He pointed out a universal character of Non-factors: they played a keynote role not only in the structure of real human knowledge, but also in many applications of computational mathematics.

The English counterpart of Non-factors called (Im-In-Un's) was introduced in [13]. Non-factors penetrate all the stages of knowledge engineering: from knowledge acquisition and knowledge representation to knowledge processing and knowledge transfer [14]. Moreover, «the main issue of Artificial Intelligence» (AI) in the first quarter of XXI century should be formulated as follows: «Can a system be considered *intelligent*, if it does not model some Non-factors?» [15].

A classification of Non-factors [16] is shown in Figure 1.

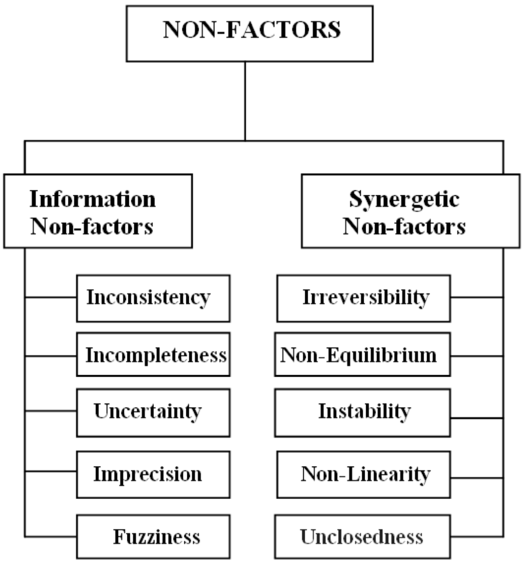


Figure 1. Information and Synergetic Non-Factors

Below we shall specify Non-factors by comparing well-known laws of classical logic (identity. excluded middle, non-contradiction, ex falso quodlibet) with real human logic and everyday reasoning. A more detailed analysis of non-classical logics induced by knowledge Non-factors can be found in [17].

The meaning of a term in mathematical logic is determined according to an interpretation, so it does not change as the system runs. Contrarily, the meaning of a term in human mind often changes according to personal experience and context. Example: What is «truth»?

In classical logic the principle of compositionality is used: the meaning of any complex expression is completely determined by the meanings of its constituent expressions and the combination rules (connectives). On the contrary, the meaning of a compound term in human mind or natural language usually cannot be reduced to that of its components, though is still related to them. Example: «Is really an AI concept «blackboard» a black board?» [4].

In classical logic, a statement is either true or false, but people often take intermediate truth values of statements as between true and false. Such a value can be viewed as «uncertain», «possible», «half true», and so on. The use of such intermediate truth values, the truth graduation makes an appeal to many-valued logics for AI.

Furthermore, classical logic is explosive. It means that from contradiction we can obtain any arbitrary conclusion. However, the existence of a contradiction in a human mind does not interfere common-sense reasoning. Moreover, a detection of technical contradiction is a starting point for Altshuller's algorithm (shortly АРИЗ in Russian) of inventive problem-solving theory (ТРИЗ in Russian) . So paraconsistent logics are in great demand to model human reasoning.

In classical logic, the truth value of a statement does not change over time, it is monotonous. However, people easily revise their beliefs after getting new information. For our through-section example, if we take instead crows some more exotic birds like penguins, then we have to discard an ordinary premise «Birds fly», but can preserve the early used premise «Birds have feathers». Such situations give us good examples of non-monotonous reasoning.

In traditional reasoning systems, inference processes follow strict algorithms, therefore are predictable. On the other hand, human reasoning processes are often unpredictable, and can «jump» on the unexpected side. In is natural for scientific discovery, then a researcher deviates from the research plan and waits for an «inspiration».

In classical logical reasoning, the backtracking procedure is crucial, i.e. how a conclusion is obtained may be accurately explained step by step. Of course, this conclusion can be repeated. Contrarily, the humans are able to generate such conclusions, whose sources and paths contain «blank spots» or cannot be backtracked at all. As an example we cite a typical variant of everyday uncertain reasoning: «I don't know why it will occur. It is only my bad feeling».

Finally, classical reasoning systems meet the Closed World Assumption (CWA) – what is not known to be true must be false. However, the practice of human reasoning shows that Open World Assumption (OWA) – what is not known to be true is simply unknown – is much more realistic.

C. Human Cognition : Processes, Properties and Units

Basically, cognition stands for gaining new information and knowledge by providing the missing knowledge necessary to solve a problem under uncertainty [7]. In other words, cognition may be seen as the ability of intelligent system to find new information, acquire knowledge and reduce its environment uncertainty for the sake of adaptation. It is reached by improving an internal model of this environment.

In psychology, the term «cognition» encompasses various individual mental processes, such as sensation, perception, representation, imagination, cogitation, thinking, memory, learning, attention, explication, comprehension. In particular, cognition can be viewed as a thinking process oriented towards problem-solving; in this sense, it is involved into any human activity. In practice, problem-solving directly connects perception, thinking, memory and learning.

Following T.Gergely [7], let us recall some basic features of cognition, which are of primary concern for developers of artificial cognitive systems. First of all, cognition is an open system based on both available knowledge and current data perception. Secondly, cognition does not provide conclusions, but generates hypotheses, and these hypotheses should be confirmed or denied. Thirdly, cognition is tightly connected with understanding: it leads to knowledge changes and modifies the capacity of information processing. And fourthly, cognition in a purposeful system is intrinsically linked with the organization of action (as information process, local environment change or physical movement).

A suitable way of treating cognition in agent is to divide it into smaller units, called *cognitons* [18,19].These units are open and heterogeneous: they represent from a cognitive angle of view different sides of consciousness – cognition itself, communication, activity regulation. Besides, the notion of «cogniton» is considered here as a basic term to denote principles, mechanisms and models of self-organization in agent from the viewpoint of its cognitive subsystem. The specification of generic classes of cognitons and establishing links between them is the first stage of *cognitive engineering*, extending well-known approaches of knowledge engineering. Good examples of cognitive engineering in creating dynamic mental structures of intelligent agents are BDI-models [20] and WILL-architecture [21]. For instance, the BDI (Belief – Desire – Intention) complex unit shows what an agent thinks to be true, what it would like to achieve and how it expects to do it. A classification of cognitions is given in Figure 2.

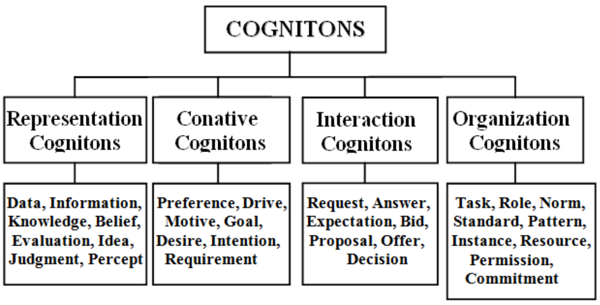


Figure 2. Four Basic Types of Cognitons

Conation is a term that stems from the Latin *conatus*, meaning any natural tendency, impulse, or directed effort. Conative cognitons representing an intentional side of activities are the keystones of agency.

III. Cognitive Agents in Actions

A. What is Agent?

According to Longman Dictionary, *agent* is a person or organization that represents another person or organization and manages their business. From methodological point of view, agent theory is intended to bridge the gap between two poles: active subject and classically passive object [19]. In this context, it is natural to notice two contrary approaches to constructing agent: an *antropomorphic vs programmer's approach*. If we move from «subject pole», then the agent can be seen as a *quasi-subject*, able to substitute his master (owner) and perform necessary task. Here a subject delegates some functions, permissions and rights to his agent. Vice versa, if we start from «object pole», then the agent may be viewed as a sort of *active object* or meta-object capable to manipulate various objects, create or destroy them, and communicate with other agents. In other words, a problem of making the object more active and more intelligent is faced.

Agents are classified into natural and artificial, physical and virtual, static and mobile, reactive and intelligent. For instance, artificial agents can be both physical (autonomous mobile robots, artificial swarms) and virtual (softbots, infobots, mobots). Four basic interpretations of artificial agents are specified [19]: artificial organism, active object, personal assistant, virtual doer. Properties and architectures of artificial agents depend on their definition, interpretation and status.

There are different definitions of agents. S.Russell and P.Norvig [22] gave a very weak definition of agent as an entity that can be viewed as perceiving its environment through sensors, to obtain data about events in this environment, and acting upon it through effectors. In fact, this definition reduces agent to a basic «organism – environment» model by M.G.Gaaze-Rapoport and D.A.Pospelov [23]. In some sense, socially-oriented definition of software agent was given by M.Coen [24]: software agents are programs that engage in dialogs, negotiate and coordinate the transfer of information.

The most popular definition belongs to M.Wooldridge and N.Jennings [25]. They defined artificial agent as an autonomous, reactive, pro-active, communicative system. Let us discuss the components of this minimal «gentleman's set». Here the term «autonomous» means that agents operate without direct intervention of humans and have some kind of control over their actions and internal state. The word «reactive» includes the perception of agent's environment and response in a timely fashion to all the changes in it. «Proactive» means that agents do not simply act in response to their environment, they are able to exhibit goal-oriented behavior by taking the initiative. Communication stands for a social ability, i.e. agents interact with other agents (and possibly humans) via some kind of agent-communication language.

In [19] we proposed the following definition: an agent is an open, active, intentional (goal-directed) system able to generate and perform its proper activity in an uncertain or fuzzy environment.

In our opinion, it is necessary to emphasize an intentional nature of any agent: the reason of agent's activity is the need that is viewed as a difference between desired and current agent's state. The need generates some motivation or forms some preferences, and agent's motive is deployed into its goal – a model of agent's wanted future. Agent's autonomy is ensured by its proper resources; that supposes a periodic resource acquisition from the environment (or other agents).

The behavior of reactive agents is determined by simple impulses and preferences and stimulus-reactive links, whereas the synthesis of intelligent agents supposes the development of internal model for external world, formation of both belief-base and knowledge base, reasoning for planning and performing actions (Figure 3). Besides beliefs and planning, intelligent agents are often equipped with such features as prediction and persistency.

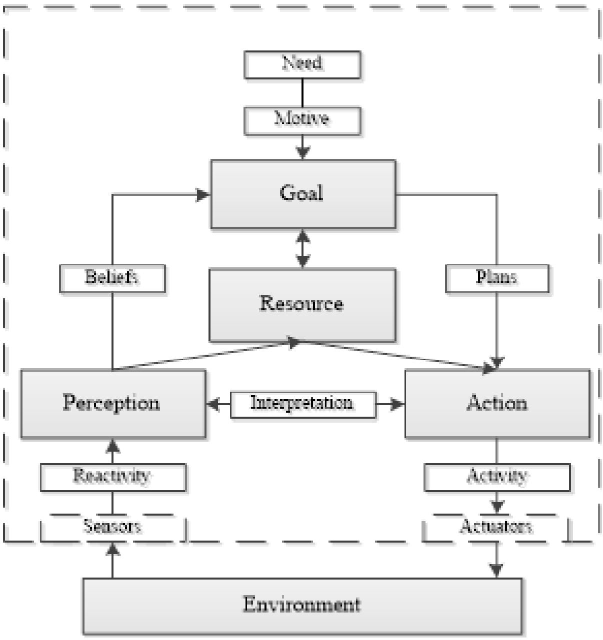


Figure 3. Basic Architecture of Intelligent Agent

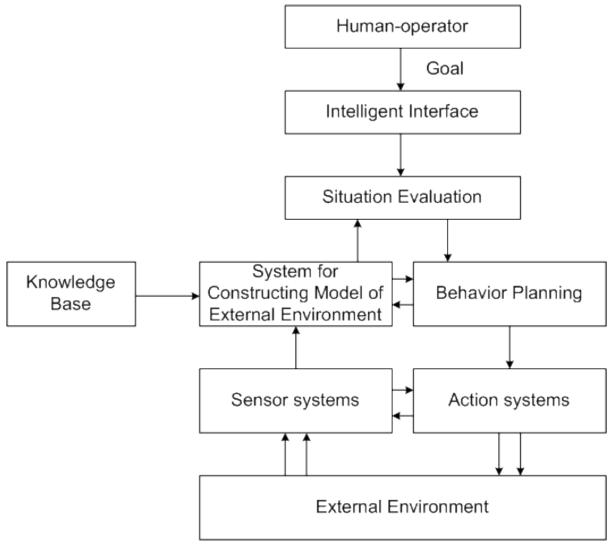


Figure 4. Architecture of Cognitive Agent as Open Semi-Autonomous Goal-Directed System

B. Artificial Cognitive Agents

An artificial cognitive agent possesses a well-developed internal model of its duties, external world, other agents (including human agents in order to understand human needs and queries) and itself. It receives and integrates current information from, at least, three sources: a) its human partner (in the form of goal formulation or adjustment, operating instructions, on-line responses to questions); b) its sensor system; c) its belief/knowledge base (Figure 4). Agent's cognition is the process of acquiring knowledge and understanding through the senses, perception, thought and experience. It opens new possibilities of learning and reasoning about how to behave in order to achieve goals in uncertain or ill-defined environment.

A typical example of physical artificial cognitive agent is cognitive robot (group of cognitive robots). Cognitive Robotics is a new branch of robotics aimed at generating an intelligent behavior in robot by enhancing its cognitive capacities. It studies how cognitive robot obtains and aggregates information on his world, in which form it should be represented and memorized, how this information is transformed into beliefs and knowledge, and how these beliefs govern robot's behavior.

Basic technological problems of cognitive robotics are machine vision, voice recognition, speech synthesis, various types of sensing (proximity sensing, pressure sensing, texture sensing, and so on).

Thus, a central problem of cognitive robotics is data fusion – the integration of multiple data sources to produce more diverse, rich, accurate and useful information, as well as sensor data mining and knowledge discovery.

Cognitive capacities of intelligent robots also include perception processing (specifically, computing with words and perceptions [26]), approximate reasoning, anticipation, attention sharing, ability to learn from mistakes, etc.

Moreover, artificial cognitive agents ought to have the possibility of communicating in a dialogical manner with human agents (users) by applying a restricted natural language (Figure 5). Such a dialogue includes both tasks instructions given by human agent to artificial agent and a feedback from artificial agent (situational information, report about goal achievement or request for additional data)

Both individual and collective behavior of cognitive agents is goal-driven and supposes the study of practical aspects of their acts and actions to obtain useful result. Also communication processes between cognitive agents based on speech theory and conversation rules have situational context, i.e. pragmatic foundations. Therefore, agent-oriented paradigm is closely related to the area of pragmatics.

So the involvement of cognition into action by cognitive agent supposes information granulation [8] or more generally, cognition granulation and aggregation. Below we will consider pragmatic granulation as a basic feature of cognitive agent.

C. Information Granulation by Cognitive Agent

According to Zadeh, *granule* is a collection of objects which are drawn together by the relations of similarity, indistinguishability, functionality or proximity [8]. Generally, information granules are complex dynamic information entities which are formed to achieve some goal. The arrival of information granulation means the transition from ordinary machine-centric to human-centric approach in information gathering and knowledge discovery [27]. The concept of information granulation is closely related to data abstraction and derivation of knowledge from information. By selecting different levels of granulation one can obtain different levels of knowledge.

Granulation theory includes studies in classification, generation, representation, interpretation and use of granules. Typical

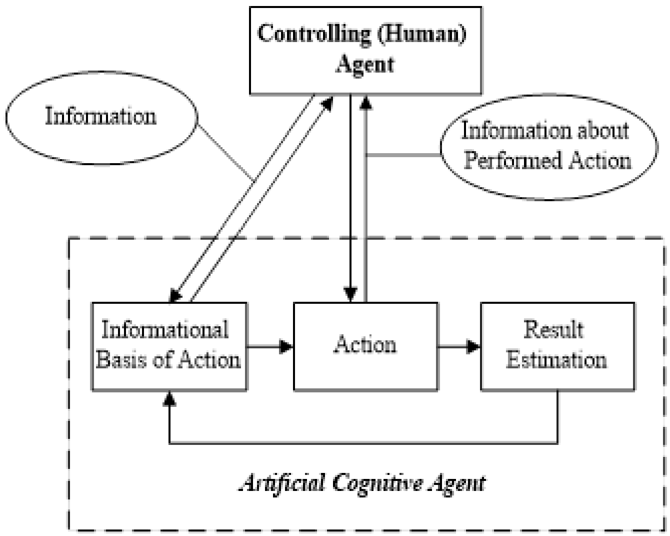


Figure 5. Interactive Model of Dialogical Control

interpretations of granules are: part of the whole, sub-problem of the problem, uncertainty zone, variable constraint.

There are various classifications of granules: physical and conceptual granules, one-dimensional and multidimensional, information and knowledge granules, time and space granules, crisp and fuzzy granules, etc.

Where are two general approaches to generating granules – top-down and bottom-up. A top-down approach is based on a set that is divided into subsets, these subsets – into smaller subsets, etc. Various coverings, partitions, nested sets are typical examples of this approach. Inversely, in case of bottom-up approach we firstly take a point (a singular object) and then construct its neighborhood. As a result, a pre-topology of neighborhood system is obtained. These two approaches show a hierarchical nature of both granules and granulation process itself.

Granules may be obtained by specifying *non-classical sets*. Classical sets have crisp boundaries and additive measure. They satisfy two basic postulates: 1) membership postulate; 2) distinguishability postulate (do not confound with. *Membership postulate* is analogous to excluded middle law in classical logic: Every element of a set must be uniquely specified as belonging to the set or not. According to *distinguishability postulate*, a set is viewed as a collection of different, clearly distinguishable elements which can be enumerated, represented by a list. If either one or both of these postulates are rejected, then we obtain non-classical set theories. Valuable examples of non-classical sets are over-determined and under-determined sets depending on observer's awareness parameter [28]. Another well-known example concerns *rough sets* [29]. These three non-classical variants of sets can be expressed by three-valued characteristic functions.

In Section IV we will focus on logical granules and granulation driven by pragmatics.

IV. Towards Logical Pragmatics: a New Status of Truth Values

A. Logical Pragmatics and Pragmatic Logics

Nowadays, the arrival and intensive development of both Logical Pragmatics and Pragmatic Logics is founded on Ch.S.Peirce's ideas on relationships between information, logics and semiotics. According to Peirce [11], «Logic, in its general sense, is another name for semiotic, a formal doctrine of signs». In information theory, any message can be related to both its author (sender) and its user (recipient): the first relation specifies semantics and the second one – pragmatics.

This pragmatic side of logic was also taken into consideration by N.A.Vasiliev in the context of two-leveled logical hierarchy [30, 31]: «Some logical principles are fixed, unchangeable and absolute, some other principles, such as non-contradiction law and excluded middle law, are relative, changeable and have empirical sources. It means that our human everyday logic is dual, semi-empirical, semi-rational, and we can consider by contrast formal and purely rational discipline, a sort of generalized logic; we call it *meta-logic*».

According to Vasiliev, we ought to make difference between two levels of knowledge: a) empirical level based on real-world's events; b) conceptual level depending on our thinking.

In modern logic, meta-logic means the study of meta-theory of logic, including the construction of logical theories, intrinsic properties of these theories, interpretations of formal systems, etc.

So Pierce's vision of logic encompasses both logical *semantics* and *logical pragmatics*. Semantics is a branch of meta-logics that studies the interpretations of logical calculus. It is worth stressing that these interpretations are context-independent and meet closed-world assumption. Inversely, pragmatics takes into account the dependence of interpretation from context.

Furthermore, Peirce considered logic as a normative science and defined truth as the good of logic [11]. A well-known Peirce's definition of truth as «the concordance of an abstract statement with the ideal limit towards which endless investigation would tend…» [11] and, even, more radical sentence by W.James [32] that «truth is the expedient in the way of our thinking», anticipated modern theories of approximated, partial, gradual, granular truth.

A pragmatic approach gives us a functional (or axiological) interpretation of truth where some proposition or belief is true, if it has some utility (enables us to attain useful practical result).

So *logical pragmatics* are associated with the pragmatic truth theory, whereas *pragmatic logics* suppose an axiological consideration of logical concepts, the specification of pragmatic truth values and the application of effectiveness principle in the form of pragmatic maxim.

To differ from descriptive correspondence theory, here the nature of truth is attributed to the reason of truth and supposes the transition from prescriptive proposition (norm) to reality (Figure 6). Here the opposition «Description-Prescription» clarifies the meaning of the opposition «Truth – Value». Truth is the correspondence between a reality object and a proposition giving its description; inversely, utility is the correspondence between a prescription and the reality object (the usefulness of norm). It is worth noticing that any activity of cognitive agent is based on both descriptions and prescriptions; it supposes a joint use of these two truth theories.

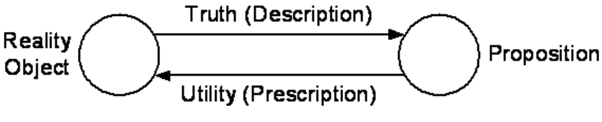


Figure 6. Classical Truth vs Utility: the Opposite Status

The next step on the way from logical semantics to logical pragmatics was made by Polish scientists: K.Ajdukiewich [33], a founder of Pragmatic Logic, and T Kotarbinsky, the author of Praxiology, as well as by Russian logician A.A.Iwin [34], who constructed the logics of values and evaluations. Another Russian logician B.Pyatnitsyn, who specified the class of pragmatic logics, is worth mentioning. Typical cases of pragmatic logics are inductive and probabilistic logics; more recent examples encompass various modal logics such as epistemic, doxastic, deontic, communication, preference, decision logics. All these logics express the relationships between some standards given by modalities and their use in practice.

B. Belief Utility vs Knowledge Truth in Intelligent Agents

In Theaetetus Plato introduced the definition of knowledge which is often translated as «justified true belief». This definition is even today largely accepted by knowledge engineers in Artificial Intelligence. They suppose that knowledge should be strictly true, and the main restriction for knowledge processing is truth preservation. The procedures of knowledge adjunction and correction are oriented to the monotonic increase of «general truth level» of stored knowledge.

However, in cognitive systems beliefs and knowledge are not necessarily based on truth. Instead of truth such criteria as belief value, utility, adequacy, stability can be useful. Moreover, the concept of truth itself undergoes significant changes.

One of the first Russian scientists, who noticed a non-universal character of truth in the context of knowledge engineering was O.P.Kuznetsov. In 1995 in the course of the International Conference «Artificial Intelligence in XXI century», he participated in the Round Table with the talk «About Knowledge Based not on the Truth». According to him, truth is a sort of knowledge values contributing to knowledge stability. However, common-sense knowledge and reasoning is not based on the truth in the sense of classical logic, but it employs suitable knowledge structures (in the sense of Gestalt Psychology) [35].

J Łukasiewicz defined logic is the science of objects of a special kind, namely the science of *logical values* [36]. In this paper the term «logical values» encompasses both truth values, including granular and fuzzy truth values, and axiological, epistemic, doxastic, deontic modal values.

C. Interpretations of Logical Values: From Dunn's Semantics to Natural Pragmatics.

Nowadays we have various novel interpretations of logical values [37], for example: a) values that convey some information on a proposition; b) entities that explain the vagueness of concepts; c) indicators of degree of truth, etc. On the one hand, it is clear that truth values can be used to deal with information and uncertainty, belief and doubt, knowledge and ignorance. On the other hand, these logical values can be gradual and granular.

Let us recall that there exists an epistemic logic that is the logic of knowledge and belief. In the context of knowledge engineering, an epistemic interpretation of truth values is quite natural.

One of the first successful attempts to construct non-standard logical semantics for practical use was performed by J.Dunn [38]. He proposed a new epistemic strategy of constructing logical semantics by rejecting classical principles of Bivalence and Functionality (singularity of both Truth and Falsity). The following three postulates underlie Dunn's approach: 1) information (or knowledge) can be incomplete and/ or inconsistent; 2) some propositions can be neither true nor false; 3) some propositions can be both true and false. This approach means specifying logical values as the set of subsets; it leads to a generalization of the classical notions of the truth value and truth value function. Such generalized truth values are considered as a rational explication of agent's incomplete and inconsistent information states.

Now let us mention some context-based, i.e. pragmatic truth-values. Good examples are: Finn's factual truth, factual falsity and factual contradiction in an argumentation context [6], our concerted truth and concerted falsity in a negotiation context [39], measured truth, measured falsity and measured ambiguity in the framework of cognitive measurement [40], and so on. A natural representation for three-valued and four-valued pragmatics is traffic lights pragmatics.

V. On the Way to Cognitive Logic

The idea of Cognitive Logic (CL) can be interpreted in two ways: a) CL as a logic based on the principles, mechanisms and attributes of human cognition; b) CL as a logical tool for supporting the cognition and understanding processes.

One of the principal mechanisms of human cognition is the construction and use of bipolar (opposition) scales. Bipolarity is referred to as the capacity of human mind to evaluate reason and make decisions on the basis of both positive and negative estimates and affects. Its application in logical investigation brings about the specification of logical worlds.

A. Logical Spaces and Worlds–an Old Vine in New Bottles

The concept of logical space was introduced by L. Wittgenstein in his famous «Tractatus Logico-Philosophicus» [41]. It is based on «possible state of affairs», specified by a proposition. An elementary state of affairs is a point of logical space. If the proposition is not elementary, it corresponds to a region in logical space. Actual states of affairs are called facts. The world is the totality of facts. By taking the condition of logical independence for n states of affairs, we can obtain 2^n combinations of these states. Each combination can be called a possible world.

The term «Logical World» can be defined more generally as a set of logical entities. From a pragmatic standpoint, it can be viewed as any non-empty set of logical values (truth values in [42,43]). According to Ya.Shramko, the elements of logical world meet two basic principles: a) distinguishability (do not confound with distinguishability postulate in set theory); b) designation. Here the word «distinguishability» means that all logical values differ among themselves. Some of them have a particular status, i.e. they are «designated». For instance, classical Frege's world Cl2 is expressed by a pair

$$LW_{C12} = \langle V_2 = \{T, F\}, D_1 = \{T\}\rangle.$$

To specify simple, unidimensional logical worlds, we should give a set of truth values V with its cardinality $\mid V \mid$, a set of designated truth values D with its cardinality $\mid D \mid$, $D \subset V$. For the sake of convenience, we write it in a short form:

$$LW_s = \langle V_i, D_j\rangle, i = 1, 2, \ldots, n, \ldots, \infty, j < i. \quad (1)$$

Generally $D_j = D_j^+ \cup D_j^-$, where D_j^+ is a set of designated values and D_j^- is a set of anti-designated values. These anti-designated values (truth-values, «similar to falsity») are often used in practice, for instance, when we face the problem of fault diagnosis.

A granular logical world is a pair

$$LW_g = \langle 2^{V_i}, 2^{D_j}\rangle, i = 1, 2, \ldots, n, \ldots, \infty, j < i. \quad (2)$$

To differ from [42], we shall consider the basis of logical world LW_b as a bipolar scale with a neutrality (midpoint) M, e.g.

$$LW_b = \langle\{T, F\}, D, M\rangle$$

and complete the principles of distinguishability and designation by the principle of structuration. The specification of world supposes the interpretation of neutral value. Generally, T, F, M are granular truth values, for instance, intervals or distributions. In the case of singletons $T = \{T\}$, $F = \{F\}$, $M = \{M\}$ we obtain a minimal logical world.

For example, Lukasiewicz's minimal logical world

$$LW_{L3} = \langle V_3 = \{T, M, F\}, D_1 = \{T\}, M = "possibility"\rangle.$$

Kleene's minimal paracomplete world

$$LW_{K3} = \langle V_3 = \{T, M, F\}, D_1 = \{T\}, M = "ignorance"\rangle.$$

Bochvar's minimal non-sense world

$$LW_{L3} = \langle V_3 = \{T, M, F\}, D_1 = \{T\}, M = "non-sense"\rangle.$$

Moreover, Vasiliev's paraconsistent world is a triple

$$LW_{Las3} = \langle V_3 = \{T, M, F\}, D_2 = \{T, B\}, M = B - "bothtrueandfalse"\rangle.$$

For Dunn-Belnap's world we have two neutral points

$$M = \{N, B\}, N = "neithertruenorfalse".$$

Finally, Zadeh's logical world

$$LW_Z = \langle V = [0, 1], D = [\alpha, 1], M \approx 0.5\rangle.$$

where $0.5 < \alpha < 1$.

More generally, fuzzy logical worlds can be specified by the set of fuzzy truth values together with fuzzy designated truth values (e.g. Radecky's fuzzy level sets), fuzzy inclusion $D \subset V$ with a grade μ and fuzzy neutralities. We have to make difference between Zadeh's fuzzy world, Atanassov's intuitionistic fuzzy world, Goguen's L-fuzzy world and so on.

Now let us consider two-dimensional logical worlds. According to the *structuration principle*, logical entities, in particular, logical values, form various structures. In other words, various order relations, for instance truth order and knowledge order, win-loss order and consensus order, form various logical structures. So the truth partial order $\leq_T$ generates a truth-value lattice (V, $\leq_T$) defined on a partially ordered set of truth values V (with at least two elements), the knowledge partial order $\leq_K$ underlies a knowledge lattice (V, $\leq_K$), etc. It is suitable to uniformly represent the pairs of lattices above by bilattices [44] with double Hasse diagrams. In Figure 7 two Hasse diagrams are drawn to illustrate dialogical bilattices with the consensus order $\leq_C$ and disputation order $\leq_D$, where N_1 and N_2 stand for uncertainty (the indices 1 and 2 correspond to two agents involved in the dialogue).

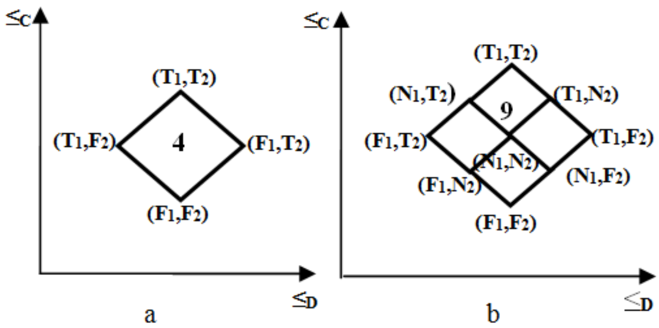


Figure 7. Examples of Dialogical Bilattices: a) the Minimal Dialogical Bilattice 4; b) the Bilattice 9.

A compound logical world is specified by a quadruple

$$LW = \langle V, D, M, R\rangle \quad (3)$$

where V is the universe of logical values v, $v \in V$, D is the set of designated values, M is the set of neutralities with appropriate interpretations, R is the set of relations in V. Let us note that logical worlds expressed by truth values with order relations play the part of logical ontologies.

B. Cognitive Graphics for Understanding Various Logics

A very important property of graphics is its direct impact on human creative thinking that results in information compression and better understanding of faced problem. In particular, cognitive graphics deals with a computer-based visualization of internal content deploying the sense of scientific abstractions. It was introduced by A.A.Zenkin [45] and D.A.Pospelov to intensify human cognitive processes related to problem formulation, searching for solution and scientific innovations.

Cognitive computer graphics was successfully used in pure mathematics to visualize the distribution of prime numbers by Ulam spiral, invariant sets of a generalized Waring problem, generate cognitive images of such transcendental numbers as π and e.

Two basic functions of graphic images are considered: illustrative function and cognitive one. Illustrative function of graphics visualizes already known objects to ensure their image-based recognition and comprehension. On the contrary, cognitive function of graphics is intended to reveal vague or hidden sense and contributes to a new knowledge generation.

There are fuzzy boundaries between illustrative and cognitive graphics: visual representation and conceptual compression of existing knowledge may inspire some new idea or hypothesis, and its confirmation or demonstration supposes the use of images or diagrams.

The use of graphics in logic has a long history: Euler circles, Venn diagrams are the best known examples of a wide use of diagrammatic tools in representing syllogisms, classical predicates, inference rules, etc. Below we will focus on colored representations of logical scales and Hasse diagrams to understand different multi-valued and fuzzy logical pragmatics (see also a practical example of using colored Hasse diagrams in synthesizing cognitive sensors [46]).

The color metaphor seems to be an adequate pragmatic tool to represent different classes of multi-valued and fuzzy logical worlds. The classical logic can be naturally viewed as a model of «black and white world». The transition to three- or four valued worlds puts into operation the traffic light pragmatics: here the truth corresponds to a green color, the falsity – to a red color, the contradiction – to a yellow color and the uncertainty – to a dark blue color. The pragmatics of Heyting's world LW_{H3} is given by the green color for the truth, the light green color for the half-truth and the red color for the falsity. Inversely, the pragmatics of Brower's world LW_{Br3} is presented by the red color for the falsity, the pink color for the half-falsity and the green color for the truth. The representation of Bochvar's world LW_{B3} can use an analogy to the «black hole»: any logical value encountering a non-sense becomes a non-sense too. So an intense black color is welcomed.

In case of two dimensional four-valued representations of modalities by two criteria of modality strength and modality sign (strong positive, weak positive, weak negative, strong negative) we obtain modal lattices. It is suitable to select a dark green color for a strong positive modality («necessary» in alethic logic, «obligatory» in deontic logic, «certain» in doxastic logic, etc.) and light green color for a weak positive modality («possible», «permitted», «hypothetical». On the contrary, we take red color to explain the role of a strong negative modality («impossible», «forbidden», «denied») and pink color for a weak negative modality («contingent», «non-obligatory», «doubtful»).

A finite-valued logic reflects as much «rainbow colors», as the number of logical values it contains, and a fuzzy logic corresponds to continuous spectrum of colors, including all shades between green and yellow, yellow and orange, orange and red, and so on. For instance, the pragmatics of Godel's world LW_{Gn} is illustrated by (n-1) shades of green from dark green to very light green (grades of the truth) and one red value (the falsity).

VI. Conclusion

The concept of Cognitive Logic for intelligent agents has been developed on the basis of information granulation, generalized logical values, logical structures, logical worlds and their visualization by introducing colored Hasse diagrams. In order to create understanding artificial agents with highly organized cognitive capacities, main properties and processes of human cognition have been considered, and the paradigm of Cognitive Engineering by specifying special cognitive units called cognitons and performing their granulation and aggregation has been proposed. Pragmatic roots of agents and multi-agent systems have been traced back, and the concept of Logical Pragmatics has been clarified. The necessity to introduce pragmatic issues into applied logics for cognitive agents is explained by their role of «logic users» (not «logic developers»). It expresses the importance of contextual factors and norms in the modeling of agent's individual behavior and agent's communication.

Two sides of cognitive logic have been analyzed. The intrinsic connections between cognitive logic and pragmatic logics have been shown. Here the principle «First pragmatics, then calculus» has to be satisfied. So our approach to constructing cognitive logic is based on Logical Worlds. New formalisms have been introduced to model both unidimensional and two-dimensional worlds. Some examples of logical worlds have been constructed and visualized by taking colored representation of logical values.

Our further investigation will be associated with establishing links between opposition scales and logical worlds and building a special cognitive logic to support understanding process in agent.

Acknowledgment

This work is supported by Russian Science Foundation, project No 16-11-00018 and Russian Fund for Basic Research, project No 17-07-01374.

References

[1] M.C.Roco, W.S.Bainbridge (Eds.), Converging Technologies for Improving Human Performance. NSF/DOC-sponsored report. – Heidelberg: Springer-Verlag, 2003.

[2] H.E.Gardner, The Mind's New Science: a History of Cognitive Revolution. New York: Basic Books, 1987.

[3] V.B.Tarassov, "Problem of Understanding: the Present and Future of Artificial Intelligence", Open Semantic Technologies for Intelligent Systems. Proceedings of the 5^{th} International Conference, Minsk: BSUIR Editions, 2015, pp. 25-42 (in Russian).

[4] Pei Wang, "Cognitive Logic vs Mathematical Logic", Lecture Notes of the 3^{rd} International Seminar on Logic and Cognition, Guangzhou, China, 2004.

[5] D.A.Pospelov (Ed.), Time and Space Knowledge Representation in Intelligent Systems. Moscow: Nauka, 1989 (in Russian).

[6] V.K.Finn, Artificial Intelligence: Methodology, Applications, Philosophy. Moscow: KRASAND, 2011(in Russian).

[7] O.M. Anshakov, T.Gergely, Cognitive Reasoning: a Formal Approach. Heidelberg: Springer-Verlag, 2010.

[8] L.Zadeh, "Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic", Fuzzy Sets and Systems, vol.90, 1997, pp.111-127.

[9] T.Lin, "Granular Computing: from Rough Sets and Neighborhood Systems to Information Granulation and Computing in Words, Proceedings of the European Congress on Intelligent Techniques and Soft Computing, 1997, pp. 1602-1606.

[10] V.V.Golenkov, N.A.Guliakina, "The Project of Open Semantic Technology for Component Design of Intelligent Systems. Part 1. Principles of Designing", Ontology of Designing, No 1, 2014, pp.42-64 (in Russian).

[11] Ch.S.Peirce, Collected Papers. Ed. by C.Hartshorne and P.Weiss. Vol.1-6; A.Burks. Vol.7-8. Cambridge MA: Harvard University Press, 1965-1967.

[12] A.S.Narinyani, "Non-Factors and Knowledge Engineering: from Naïve Formalization to Natural Pragmatics", Proccedings of the 4^{th} National Conference on Artificial Intelligence (CAI'94, Rybinsk, Russia. September 15-21, 1994),. vol.1. Tver: AAI, 1994, pp.9-18 (in Russian).

[13] J.P.Delgrande, J.Mylopoulos, "Knowledge Representation: Features of Knowledge", Fundamentals of Artificial Intelligence, W.Bibel, Ph. Jorrand (Eds.). Berlin: Springer Verlag, 1986, pp.3-36.

[14] G.V.Rybina, "Knowledge Acquisition in the Presence of Non-Factors", No 2, 2004, pp.82-94 (in Russian).

[15] Yu.R.Val'kman, "Non-Factors Modeling as a Basis for Intelligent Computer Technologies", No 2, 2004, pp.64-81 (in Russian).

[16] V.B.Tarassov, "Non-Factors: from Semiotic Analysis to Formalization Methods", No 2, 2004, pp.95-114 (in Russian).

[17] V.N.Vagin, "Knowledge in Intelligent Systems", Artificial Intelligence News, No 6, 2002, pp.8-18 (in Russian)

[18] J.Ferber, Les systemes multiagents. Vers une intelligence collective. Paris: InterEditions, 1995.

[19] V.B.Tarassov, From Multi-Agent Systems to Intelligent Organizations. Moscow: Editorial URSS, 2002 (in Russian).

[20] A.S.Rao, M.P.Georgeff, "Decision Procedures for BDI Logics", Journal of Logic and Computation, vol.8, No 3, 1998, pp.293-344.

[21] D.Moffat, N.H.Frijda, Where there's a Will there's an Agent, Intelligent Agents: Theories, Architectures, and Languages, M. Wooldridge and N.R.Jennings (Eds.), LNAI 890. Heidelberg: Springer-Verlag, 1995, pp.245-260.

[22] S.Russell, P.Norvig, Artificial Intelligence: a Modern Approach, Englewood Cliffs NJ: Prentice-Hall, 1995.

[23] M.G.Gaaze-Rapoport, D.A.Pospelov, From Amyoba to Robot: Behavior Models. Moscow: Nauka, 1987 (in Russian).

[24] M.H.Coen, SodaBot, a Software Agent Construction System, MIT AI Lab, USA, 1995.

[25] M.Wooldridge, N.Jennings, "Intelligent Agents:Theory and Practice", The Knowledge Engineering Review, vol.10, No 2, 1995, pp.115-152.

[26] L.A.Zadeh, "Fuzzy Logic, Neural Network and Soft Computing", Communications of the ACM, vol.37, No 3, 1994, pp.77-84.

[27] A.Bargiela, W.Pedrycz, Granular Computing: an Introduction. Dordrecht: Kluwer Academic Publishers, 2003.

[28] V.B.Tarassov, "Information Granulation by Cognitive Agents and Non-Standard Sets", Proceedings of the 6^{th} International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control (ICSCCW'2011, Antalya, Turkey, September 1-2, 2011). Kaufering: b-Quadrat Verlag, 2011, pp.59-74.

[29] Z.Pawlak, "Rough Sets", International Journal of Computer and Information Sciences, vol.11, 1982, pp.341-356.

[30] N.A.Vasiliev, "Logic and Meta-Logic", Logos, vol.1-2, 1912-1913, pp.53-81 (in Russian).

[31] L.Puga, N.C.A.da Costa, "On the Imaginary Logic of N.A.Vasiliev", Mathematical Logic Quaterly, vol.34, No 3, 1988, pp.2015-211.

[32] W.James, The Meaning of Truth. Cambridge MA: Harvard University Press, 1975.

[33] K.Ajdukiewicz, Pragmatic Logic. Warsaw: Dr. Reidel Publishing Company, 1974.

[34] A.A.Iwin, Logic of Evaluations amd Norms. Moscow: Prospect Publishers, 2016 (in Russian).

[35] O.P.Kuznetsov, "A Formal Approach to the Concept of "Knowledge" and the Problem of Modeling Various Knowledge Types", in V.D.Solovyov, T.V.Chernigovskaya (Eds.). Collection of Scientific Papers on Cognitive Studies, issue 2. Moscow: Institute of Psychology of RAS Editions, 2008, pp.265-275 (in Russian).

[36] J.Lukasiewicz, Selected Works, L.Borkowski (Ed.). Amsterdam: North-Holland, 1970.

[37] Ya.Shramko, H.Wansing, Truth and Falsehood. An Inquiry into Generalized Logical Values. Series "Trends in Logics", vol.36. Berlin: Springer, 2012.

[38] J.M.Dunn, "An Intuitive Semantics for First Degree Entailment and Coupled Trees", Philosophical Studies, vol.29, 1976, pp.149-168.

[39] V.B.Tarassov, "On Multi-Valued and Fuzzy Dia-Logics", Proceedings of the 5^{th} IEEE International Conference on Soft Computing, Computing with Words and Perceptions in Systems Analysis, Decision and Control (Famagusta, North Cyprus, 2-4 September 2009).

[40] V.B.Tarassov, M.N. Svyatkina, "Cognitive Measurements: the Future of Intelligent Systems", Software and Systems, No 4, 2013, pp.74-82.

[41] L.Wittgenstein, Prototractatus: An Early Version of Tractatus Logico-Philosophicus, G.von Wright (Ed.).London:Routledge and Kegan,1971.

[42] Ya.Shramko, "The Logical Way of Being True: Truth Values and the Ontological Foundation of Logic", Logic and Logical Philosophy, vol.23, 2014, pp.119-131.

[43] A.S.Karpenko. The Development of Many-Valued Logics. Moscow: LKI Publishers, 2014 (in Russian).

[44] M.Ginsberg, "Multi-Valued Logics: a Uniform Approach to Reasoning in AI". Computer Intelligence, vol.4, 1988, pp.256-316.

[45] A.A.Zenkin, Cognitive Computer Graphics. Moscow: Nauka, 1991 (in Russian)

[46] M.N.Svyatkina,V.B.Tarassov, A.I.Dolgiy, "Logical-Algebraic Methods in Constructing Cognitive Sensors for Railway Infrastructure Intelligent Monitoring System". Proceedings of the 1^{st} International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16, Sochi, May 16-21, 2016), vol.1, A.Abraham, S.Kovalev, V.Tarassov, V.Snazel (Eds.). Advances in Intelligent Systems and Computing, vol.450. Cham, Switzerland: Springer International Publishing, 2016, pp.191-206.

ГРАНУЛЯЦИЯ ИНФОРМАЦИИ, КОГНИТИВНАЯ ЛОГИКА И ЕСТЕСТВЕННАЯ ПРАГМАТИКА ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ

Тарасов В.Б.

Московский государственный технический университет им. Н.Э. Баумана

В работе предложена концепция когнитивной логики для интеллектуальных агентов, основанная на грануляции информации, обобщённых логических значениях и алгебраических структурах, логических мирах с естественной прагматикой и их визуализации с помощью цветных диаграмм Хассе.

On Terminal State Control Of Fuzzy Dynamical Systems

Paliukh B.V., Egereva I.A., Emelyanova I.I.
Tver State Technical University
Tver, Russia
Pboris@tvstu.ru
Irina.egereva@gmail.com
Emelyanova-123@yandex.ru

***Abstract*—The paper explores the question of the terminal state control of fuzzy dynamical systems, characterized by classical fuzzy relations. On the basis of modern methods of the dynamical systems general theory asymptotic properties of solutions for obtained functional equation are studied. The problem of existence and construction of suboptimal autonomous control law with feed-back has been studied.**

***Keywords*—fuzzy dynamical systems, terminal state control, Bellman equation, suboptimal autonomous control law.**

I. Introduce

Initially, the methods of fuzzy set theory have been directed to use logical methods of decision-making, based on the compositional inference rule (see., eg, [1]). Subsequently the methods of the dynamic programming theory and fuzzy sets theory was used to develop control problem-solving techniques of deterministic and stochastic systems with fuzzy objectives and restrictions (see., eg, [2]).

It enabled to form a general theory of mathematical programming and the theory of decision-making with fuzzy objectives and restrictions (see, eg, [3] - [5]). Further development of fuzzy dynamic programming can be found, for example, in [6] overview, where, in particular, the problem of deterministic and stochastic systems with the fuzzy end time and infinite horizon control are considered.

The main goal of this paper lies in the further development of the [2]-[6] results. All the constructions are actually based on the almost apparent modification of classical composition operation of fuzzy sets.

Let X, Y and Z be certain sets. Let's assume that at a $X \times Y$ set a fuzzy relation A with membership function μ_A, is defined and at a $Y \times Z$ set a fuzzy relation B with membership function μ_B is defined. Therefore theis defined. Therefore the $A \circ B$ composition of fuzzy set A and B is the fuzzy relation in $X \times Z$ space with the membership function

$$\mu_{A \circ B}(x, z) = \sup_{y \in Y} \min[\mu_A(x, y), \mu_B(y, z)] \qquad (1)$$

(see, eg. [1]).

Let's now assume that in X space the fuzzy set R with membership function μ_R s defined. Therefore the fuzzy relation μ_A induces the fuzzy set $R \circ A$ n the Y spase. In accordance with (1) the membership function $\mu_{R \circ A}$ of $R \circ A$ set is given by an equation

$$\mu_{R \circ A}(y) = \sup_{x \in X} \min[\mu_R(x), \mu_A(x, y)].$$

Let's assume that at X set the fuzzy relation S with membership function μ_S is defined. Further, let's assume that in X space G set with membership function μ_G s also defined. Therefore we can determine the $S \circ G$ composition of fuzzy sets S and G and following (1) the membership function $\mu_{S \circ G}$ of $S \circ G$ set will be defined by equation

$$\mu_{S \circ G}(x_1) = \sup_{x_2 \in X} \min[\mu_S(x_1, x_2), \mu_G(x_2)]. \qquad (2)$$

One can readily see, for each $x_1 \in X$ the membership degree of $\mu_{S \circ G}(x_1)$ of x_1 the fuzzy set G is defined by equation (2).

II. Terminal control

Let X and U be certain compact metric spaces. Let's consider Ş control system when X is state space and U is control space.

Let's assume that evolution of Ş system state is characterized by the fuzzy relation S representing fuzzy set S in $X \times U \times X$ pace with membership function μ_S. Provided that the initial state $x_0 \in X$ is defined. As a result of choosing of $u_0 \in U$ control the system goes into some new state x_1 which was earlier unknown. It is only known that with u_0 and x_0 fixed, x_0, u_0 and x_1 variables are related by the fuzzy relation S with membership function $\mu_S(x_0, u_0, x_1)$. In other words with x_0 and u_0 fixed at point of time n=0 the state x_1 can be defined only by value of membership function $\mu_S(x_0, u_0, x_1)$. However at point of time n=1 we can observe exact value of state x_1.

Example 1. Let X and U be intervals $[-1, 1]$. Let‘s assume that the evolution of Ş system is characterized by an approximate equation

$$x_{n+1} \simeq x_n + u_n, \quad n = 0, 1, 2, \ldots$$

Then the fuzzy relation S is the fuzzy set defined at the direct product of spaces $X \times U \times X$ with membership function

$\mu_S\colon X \times U \times X \to [0,1]$. The μ_S function can for example be of the form:

$$\mu_S(x_n,u_n,x_{n+1}) = \frac{1}{1+(x_{n+1}-x_n-u_n)^2},$$

(see [2]).

Let's consider that the control aim is characterized by fuzzy goal set G in X space with membership function μ_G. Let's also assume that both functions μ_S and μ_G are continuous in the range of their definition.

Now let's assume that time N of end of system work is defined. The control problem is to search the sequence

$$u_0, u_1, \ldots, u_{N-1} \tag{3}$$

of points of U set maximizing the membership degree of x_0 states to fuzzy set G with fuzzy relations with membership functions

$$\mu_S(x_0,u_0,x_1), \mu_S(x_1,u_1,x_2), \ldots, \mu_S(x_{N-1},u_{N-1},x_N).$$

Therefore the fuzzy set G is the control aim and the problem consists in searching the control sequences (3) providing the maximal membership degree of the state x_0 to the fuzzy set G with that the evolution of system state is described be the composition of fuzzy sets S and G. By equation

$$D_N = \underbrace{S \circ \ldots \circ S}_{N} \circ G$$

let's put for consideration the fuzzy set D_N being conditional for variables (3) in the X space with membership function μ_{D_N} satisfying the equation

$$\mu_{D_N}(x_0 \,|\, u_0,u_1,\ldots,u_{N-1}) =$$
$$= \max_{x_1,x_2,\ldots,x_N} \min[\mu_S(x_0,u_0,x_1), \mu_S(x_1,u_1,x_2), \ldots,$$
$$\mu_S(x_{N-1},u_{N-1},x_N), \mu_G(x_N)].$$

Therefore according to equation (2) $\mu_{D_N}(x_0 \,|\, u_0,u_1,\ldots,u_{N-1})$ he values of function μ_{D_N} have the form of the membership degree of the state x_0 to G set with the use of any fixed sequence of control of (3) kind.

Let's set

$$\mu_N(x_0) = \max_{u_0,u_1,\ldots,u_{N-1}} \mu_{D_N}(x_0 \,|\, u_0,u_1,\ldots,u_{N-1}). \tag{4}$$

Following [1] let's consider the initial task in the context of task family where x_0 and N are variable values. Therefore with $N=0$ the required membership degree x_0 to G et with the fuzzy relation S is prescribed by the equation

$$\mu_0(x_0) = \mu_G(x_0). \tag{5}$$

Function μ_0 s continuous by convention over all of the intervals at X set. Moreover because of continuity of functions it is easy to note that for each function f which is defined and continuous over all of the intervals at X and possesses values at the interval [0,1], the function

$$g(x_0,u_0,x_1) = \min[\mu_S(x_0,u_0,x_1), f(x_1)]$$

is continuous over all of the intervals. But X and U spaces are compact. Therefore, the function

$$h(x_0) = \sup_{u_0,x_1} \min[\mu_S(x_0,u_0,x_1), f(x_1)] =$$
$$\max_{u_0,x_1} \min[\mu_S(x_0,u_0,x_1), f(x_1)]$$

is continuous over all of the intervals at X set. Provided that

$$\max_{u_0,u_1,\ldots,u_N} \mu_{D_{N+1}}(x_0 \,|\, u_0,u_1,\ldots,u_N) =$$
$$= \max_{u_0,x_1} \min[\mu_S(x_0,u_0,x_1),$$
$$\max_{u_1,u_2,\ldots,u_N} \mu_{D_N}(x_1 \,|\, u_1,u_2,\ldots,u_N)]$$

(see, eg. [9]). Then by virtue of (4) for certain N the equation

$$\mu_{N+1}(x_0) = \max_{u_0,x_1} \min[\mu_S(x_0,u_0,x_1), \mu_N(x_1)], \tag{6}$$

is executed where $\mu_{N+1}(x_0)$ is the maximal membership degree of the statex_0 to the G set with the relation S and the condition where end of system work time is equal to $N+1$, and $\mu_N(x_1)$ is the maximal membership degree of the state x_1 to the G set with relation S and the condition where end of system work time is equal to N.

One can readily see that recurrence relationship (6) with the condition (5) is similar to Bellman's functional equation for classical problems of dynamic programming. This relationship interprets the control u_0 as function of time N and the state x_0, i.e.

$$u_0 = u_0^*(x_0,N), \quad N = 1,2,3,\ldots \tag{7}$$

Now let‘s note that for each N the function μ_{N+1} is defined. In addition if in-equations

$$\min_{x_0 \in X} \mu_G(x_0) > 0$$

and

$$\min_{x_0,u_0,x_1} \mu_S(x_0,u_0,x_1) > 0,$$

are executed, then for all $N = 0,1,2,\ldots$ the inequation

$$\mu_N(x_0) > 0.$$

is correct. It is obvious that in this case we can always imply a well-defined task.

Example 2. Let‘s assume that in the conditions of example 1 the G is a set of real numbers essentially larger than zero. Therefore the function of membership μ_G to the G-set can be defined by

$$\mu_G(x) = \begin{cases} 0, & x \in [-1,0], \\ (1+x^{-2})^{-1}, & x \in (0,1]. \end{cases}$$

(see, eg. [2]). Therefore, the recurrence relationship (6) is as follows

$$\mu_{N+1}(x_0) = \max_{u_0,x_1} \min \left[\frac{1}{1+(x_1-x_0-u_0)^2} \, , \, \mu_N(x_1) \right],$$

where $N = 0, 1, 2, \ldots$ and

$$\mu_0(x_0) = \begin{cases} 0, & x_0 \in [-1, 0], \\ (1 + {x_0}^{-2})^{-1}, & x \in (0, 1]. \end{cases}$$

III. Autonomous Law

In many practical cases it is appropriate to replace control law (7) by autonomous law

$$u = u^*(x) \tag{8}$$

(see, eg. [3]). In order to understand the availability of getting such a law let's study asymptotic properties of relationship (6).

Let's set ļ as an set closure operation. Let $C(X, [0, 1])$ be a space of continuous functions defined at X set and possessed value at interval [0,1]. For certain function $\varphi \in C(X, [0, 1])$ let's assume

$$A\varphi = \max_{u_0,x_1} \min[\mu_S(x, u_0, x_1), \varphi(x_1)],$$

where A is an operator mapping the space $C(X, [0, 1])$ into itself. Therefore the following sentences are correct.

Proposition 1: Proposition 1. Let's assume that the X space is finite. Therefore the set

$$\Omega(\mu_0) = \bigcap_{N \geq 0} \text{ļ} \left(\bigcup_{k \geq N} A^k \mu_0 \right)$$

isn't empty, it is compact in the topology of simple convergence and is invariant. In such a case the equation

$$\lim_{k \to +\infty} A^k \mu_0 = \Omega(\mu_0). \tag{9}$$

Sentence 1 proposition. One can readily see that the A operator is continuous over all of the intervals at the X. set by convention. Thus the operator set discrete dynamical system

$$A^N = \underbrace{A \cdots A}_{N}, \quad N = 0, 1, 2, \ldots,$$

where A^0 is an operator of identity transformation (see, eg. [10]).

Let x be an arbitrary point of the Xset. Therefore set of points

$$A^N \mu_0(x), \quad N = 0, 1, 2, \ldots,$$

is relatively compact. But the X set is finite. Thus $\Omega(\mu_0)$ set isn‘t empty, is compact in the topology of simple convergence and invariant (see, eg. [10]). Moreover it is easy to notice that in this case the equation (9) results immediately from the definition of $\Omega(\mu_0)$ set.

Proposition 2: Proposition 2. Let M be a set of functions

$$\mu_0, \mu_1, \ldots, \mu_N, \ldots \tag{10}$$

Therefore if the X space is finite then the set

$$\Omega(\text{ļ}M) = \bigcap_{N \geq 0} A^N \text{ļ}M \tag{11}$$

isn't empty, is compact in the topology of simple convergence and is invariant. In such a case the equation

$$\lim_{N \to +\infty} A^N M = \Omega(\text{ļ}M). \tag{12}$$

Sentence 2 proposition. First of all let‘s notice that by the X set finiteness the set ļM isn‘t empty and is compact in the topology of simple convergence. Thus by repeating the proof almost as it stands in lemma 4.2.2 in the book [10] (with substitution according to sentence 1 Banach space to metric space) we get nonemptiness, compactedness and invariance of $\Omega(\text{ļ}M)$set and equation (11) and (12).

In general case the X space can be approximated by a finite set to a high accuracy. Thus the conditions of sentences 1 and 2 are shown with prescribed accuracy . In addition the compact space approximation by its certain finite part is justified in many practical situations for modeling of fuzzy systems (see, eg. [1]). For this reason sentences 1 and 2 set asymptotic properties of relationship (6) applicable for practice.

IV. Asymptotic Autonomous Control Law

Asymptotic properties of relationship (6) set by sentences 1 and 2 prevent from thinking directly of the optimal autonomous control law existence without any additional requirements.

Actually we can speak about the existence of such a law only with sequence convergence (10). In this case according to A operator continuity in some cases the function μ defined at the X set by the equation

$$\mu(x) = \lim_{N \to +\infty} \mu_N(x), \tag{13}$$

is a continuous solution of the equation

$$\mu(x_0) = \max_{u_0,x_1} \min[\mu_S(x_0, u_0, x_1), \mu(x_1)]. \tag{14}$$

For the purpose that the equation (14) is followed by existence of the equation (14) continuous solution it is sufficient the convergence (13) is uniform at Xset and control laws (7) are continuous. Again the necessary and sufficient condition of uniform convergence in the equation (13) as is known lies in the fact that the set (10) is equicontinuous (see, eg. [9]). Then in the case under consideration the check of equicontinuity of the set (10) is rather difficult in virtue of representation of A operator.

Let's note that the situation is extremely rare where at the X set the equation is simply (even nonuniformly) satisfied (see, eg. [9]). The situation becomes complicated by the fact that even if the equation (14) has a unique solution it doesn't mean that the control law (8) corresponding to this solution will be unique. Thus we have to speak in the majority of practical situations only about the existence of suboptimal autonomous control law.

For the development of such law let's assume that the X set is finite. Therefore according to the sentence 2 the closure ļM of the set (10) isn't empty, is compact and invariant. But each compact, invariant set contains compact minimum set (see attachment 2). Thus under conditions of sentence 2 the set $\Omega(ļM)$ contains the compact minimum set . In this case by the finiteness of the X set one can find such finite system of compact, invariant set

$$\Omega(ļM) \supset M_1 \supset M_2 \supset \ldots \supset M_N, \tag{15}$$

that equation

$$\mathcal{M} = \Omega(M) \cap M_1 \cap \ldots \cap M_{N_M}, \tag{16}$$

is satisfied where N is a certain sufficiently great positive integer (see, eg. [9]).

If the system (15) is built then according to the equation (16) the set is also built. Let μ be arbitrary function of the set. Therefore by virtue of the fact that is a minimum set the function

$$A\mu = \max_{u_0, x_1} \min[\mu_S(x, u_0, x_1), \mu(x_1)] \tag{17}$$

belongs to and v.v. (see [10]).

One can readily see that maximum in the equation (17) is attained with the use of the certain control law

$$u = u_\mu(x). \tag{18}$$

Moreover by equation (13 and (14) it is easy to note that if the optimal law (8) exists, the law is the same as law (18).

Thus a certain kind of control law (18) corresponds to each function $\mu \in$. Any of these laws in general case is only suboptimal. However by sentences 1 and 2 the equation (17) not only sets the existence of such suboptimal laws but provides a procedure of its construction.

ATTACHMENT

In order to avoid any misreading let's give fundamental definitions of the discrete dynamical systems theory (see., eg, [9]).

Definition 1. Let R be certain metric space with d metric and let A be R continuous self-mapping. An iteration family

$$A^N, \quad N = 0, 1, 2, \ldots,$$

of transformation A is known as discrete dynamical system, if A^0 is an operator of identity transformation.

Definition 2. The $F \subset R$ set is known as invariant, if inclusions

$$AF \subset F$$

and

$$AF \supset F$$

are realized.

Definition 3. The $M \subset R$ set is known as minimum if it isn't empty, is closed and invariant and free of proper subset having these three properties.

ACKNOWLEDGMENT

The research was supported by the Russian Foundation for Basic Research (project nos. 17-07-01339).

REFERENCES

[1] L. Zadeh, *The concept of linguistic variable and its application to the adoption of the approximate solutions*. 0.5em minus 0.4emMoscow, Russia: Mir, 1973.

[2] R. Bellman and L. Zadeh, *Decision-making in vague terms*, Questions analysis and decision-making procedures. Moscow, Russia: Mir, 1976, p. 172–215.

[3] S. A. Orlovsky, *Decision-making problems with fuzzy initial information*. Moscow, Russia: Nauka, 1981.

[4] K. Negoytse, *The application of systems theory to control problems*. Moscow, Russia: Mir, 1981.

[5] L.A. Zadeh, J. Kacprzyk (Eds.) *Fuzzy logic for the management of uncertainty*. Wiley, New York, 1992.

[6] J. Kacprzyk, A.O. Esogbue, *Fuzzy dynamic programming: Main developments and applications*. Fuzzy Sets and Systems, 1996, vol 81, p. 31–45.

[7] D.A. Novikov and S. N. Petrakov, *Course theory of active systems*. Moscow, Russia: SINTEG, 1999.

[8] A.P. Ryzhov, *Elements of the theory of fuzzy sets and measure the fuzziness*. Moscow, Russia: Dialog – MSU, 1998.

[9] A.P. Afanas'ev, and Dzyuba, S.M., *Ustoichivost' po Puassonu v dinamicheskikh i nepreryvnykh periodicheskikh sistemakh (Poisson Stability in Dynamical and Continuous Periodic Systems)*. Moscow, Russia: LKI, 2007.

[10] J. Hale,*Theory of Functional Differential Equations, New York: Springer, 1977. Translated under the title Teoriya funktsional'no-differentsial'nykh uravnenii*. Moscow: Mir, 1984.

ОБ УПРАВЛЕНИИ КОНЕЧНЫМ СОСТОЯНИЕМ НЕЧЕТКИХ ДИНАМИЧЕСКИХ СИСТЕМ

Палюх Б.В., Егерева И.А., Емельянова И.И.
Тверской государственный технический университет

В работе рассмотрена задача управления конечным состоянием динамических систем, характеризуемых классическими нечеткими отношениями. Решение задачи сведено к решению функционального уравнения типа уравнения Беллмана. На основе современных методов общей теории динамических систем изучены асимптотические свойства решений полученного функционального уравнения. Изучена проблема существования и построения субоптимального автономного закона управления с обратной связью.

Development of a Fuzzy Knowledge Base as a Core of the Athene Platform

Aleksey Filippov
Department of Information Systems
Ulyanovsk State Technical University
Ulyanovsk, Russia
al.filippov@ulstu.ru

Vadim Moshkin
Department of Information Systems
Ulyanovsk State Technical University
Ulyanovsk, Russia
postforvadim@yandex.ru

Nadezda Yarushkina
Department of Information Systems
Ulyanovsk State Technical University
Ulyanovsk, Russia
jng@ulstu.ru

***Abstract*—The article describes the process of developing a fuzzy knowledge base (KB) as a core of the Athene platform. The content of fuzzy KB is formed as a result of the analysis of the contexts of the problem area (PrA). In this case, the context is a certain "point of view" on the PrA and its features. A graph database (DB) is used as a basis for storing the contents of the KB in the form of an applied ontology. An attempt is made to implement the mechanism of inference by the contents of a graph database. The mechanism is used to dynamically generate the screen forms of the user interface to simplify the work with the KB.**

***Keywords*—ontology, fuzzy knowledge base, context analysis, problem area, graph database, inference, dynamic ui**

I. Introduction

Post-industrial society operates with huge volumes of information both in everyday and professional activities. A large amount of information causes difficulties in making decisions within the framework of rigid time constraints [1], [2].

A variety of software automation of human activities are used to solve this problem. However, it is necessary to adapt them to the specifics of a particular problem area (PrA) and its contexts for the effective operation of these tools [3]–[6].

Thus, "trained" automation software solves the tasks more efficiently, but they require considerable resources (human and temporary) for training.

In this paper, an attempt is made to construct a fuzzy knowledge base (KB) as a core of the Athene platform [6]. The content of the fuzzy KB is an applied ontology. The basic requirements for fuzzy KB are (fig. 1):

- adaptation to the specifics of PrA based on contexts;
- reliability and speed of ontology storage;
- the presence of a mechanism of logical inference;
- the presence of a mechanism of logical inference;
- availability of mechanisms for importing data from external information resources.

As you can see from figure 1, the KB consists of the following subsystems:

1) Ontology store:
   - Neo4j;
   - content management module;
   - ontology import/export module (RDF, OWL).
2) Inference subsystem:
   - inference module.
3) A subsystem for interaction with users:
   - screen forms generation module.
4) A subsystem for importing data from external information resources:
   - a module for importing data from external wiki-resources;
   - a module for filling external wiki-resources.

This work was financially supported by the Russian Foundation for Basic Research (Grant No. 16-47-732054)

II. The organization of the ontology store of fuzzy knowledge base

Ontology is a model of the representation of the PrA in the form of a semantic graph [7], [8].

Graph-oriented database management system (Graph DBMS) Neo4j is the basis of the ontology store for fuzzy KB. Neo4j is currently one of the most popular graph databases and has the following advantages:

1) Having a free community version.
2) Native format for data storage.
3) One copy of Neo4j can work with graphs containing billions of nodes and relationships.
4) The presence of a graph-oriented query language Cypher.
5) Availability of transaction support.

Neo4j was chosen to store the description of the PrA in the form of an applied ontology, since the ontology is actually a graph. In this case, it is only necessary to limit the set of nodes and graph relations into which ontologies on RDF and OWL will be translated.

The context of an ontology is some state of ontology, obtained during versioning or building an ontology using different "points of view".

Formally, the ontology can be represented by the following equation:

$$O = \langle T, C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i} \rangle, i = \overline{1,t}, \qquad (1)$$

where

t is a number of the ontology contexts;

$T = \{T_1, T_2, \ldots, T_n\}$ is a set of ontology contexts;

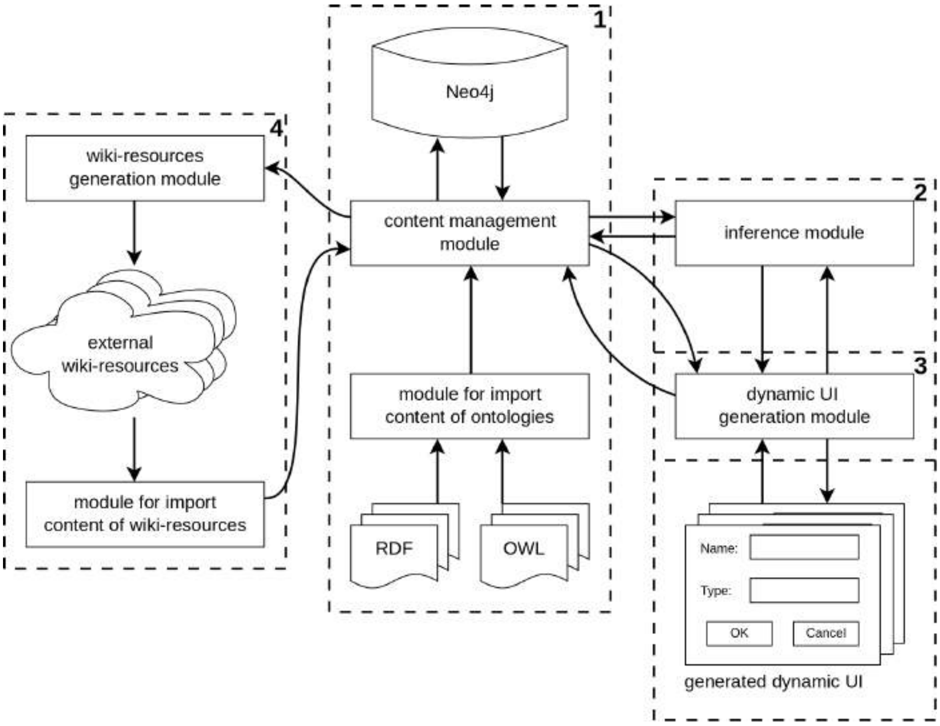


Figure 1. Architecture of a fuzzy knowledge base of the Athene platform.

$C^{T_i} = \{C_1^{T_i}, C_2^{T_i}, \ldots, C_n^{T_i}\}$ is a set of ontology classes within the i-th context;

$I^{T_i} = \{I_1^{T_i}, I_2^{T_i}, \ldots, I_n^{T_i}\}$ is a set of ontology objects within the i-th context;

$P^{T_i} = \{P_1^{T_i}, P_2^{T_i}, \ldots, P_n^{T_i}\}$ is a set of ontology classes properties within the i-th context;

$S^{T_i} = \{S_1^{T_i}, S_2^{T_i}, \ldots, S_n^{T_i}\}$ is a set of ontology objects states within the i-th context;

$F^{T_i} = \{F_1^{T_i}, F_2^{T_i}, \ldots, F_n^{T_i}\}$ is a set of the logical rules fixed in the ontology within the i-th context;

R^{T_i} is a set of ontology relations within the i-th context defined as:

$$R^{T_i} = \{R_C^{T_i}, R_I^{T_i}, R_P^{T_i}, R_S^{T_i}, R_F^{T_i}\},$$

where

$R_C^{T_i}$ is a set of relations defining hierarchy of ontology classes within the i-th context;

$R_I^{T_i}$ is a set of relations defining the 'class-object' ontology tie within the i-th context;

$R_P^{T_i}$ is a set of relations defining the 'class-class property' ontology tie within the i-th context;

$R_S^{T_i}$ is a set of relations defining the 'object-object state' ontology tie within the i-th context;

$R_F^{T_i}$ is a set of relations generated on the basis of logical ontology rules in the context of i-th context.

Principles similar to the paradigm of object-oriented programming are at the basis of the ontology of the fuzzy knowledge base:

- ontology classes are concepts of the PrA;
- classes can have properties, the child-class inherits properties of the parent class;
- objects of ontology describe instances of the concepts of the PrO;
- specific values for the properties of objects inherited from the parent class are determined by the states;
- logical rules are used to implement the functions of inference by the content of fuzzy KB.

III. The inference on the contents of fuzzy knowledge base

The inference is the process of reasoning from the premises to the conclusion. Reasoners are used to implement the function of inference. Reasoners form logical consequences on the basis of many statements, facts and axioms [9], [10]. The most popular at the moment reasoners are:

- Pellet;
- FaCT++;
- Hermit;
- Racer, etc.

These reasoners are actively used in the development of intelligent software. However, Neo4j does not assume the possibility of using similar default reasoners. Thus, there is a need to develop a mechanism for inference based on the content of a fuzzy KB.

Currently the Semantic Web Rule Language (SWRL) is used to record logical rules. These SWRL rules describe the conditions under which object a has "nephew-uncle" relation with object c. Formally the logical rule of the ontology of the fuzzy knowledge base is:

$$F^{T_i} = \langle A^{Tree}, A^{SWRL}, A^{Cypher} \rangle,$$

where

T_i is i-th context of the ontology of the fuzzy KB;

A^{Tree} is a tree-like representation of a logical rule F^{T_i};

A^{SWRL} is a SWRL representation of a logical rule F^{T_i};

A^{Cypher} is a Cypher representation of a logical rule F^{T_i}.

The tree-view A^{Tree} of a logical rule F^{T_i} is:

$$A^{Tree} = \langle Ant, Cons \rangle,$$

where

$Ant = Ant_1 \Theta Ant_2 \Theta \ldots Ant_n$ is the antecedent (condition) of the logical rule F^{T_i};

$\Theta \in \{AND, OR\}$ is a set of permissible logical operations between antecedent atoms;

$Cons$ is a consequent (consequence) of a logical rule F^{T_i}.

Figure 2 shows an example of a tree-like representation of a logical rule. This rule describes the nephew-uncle relationship.

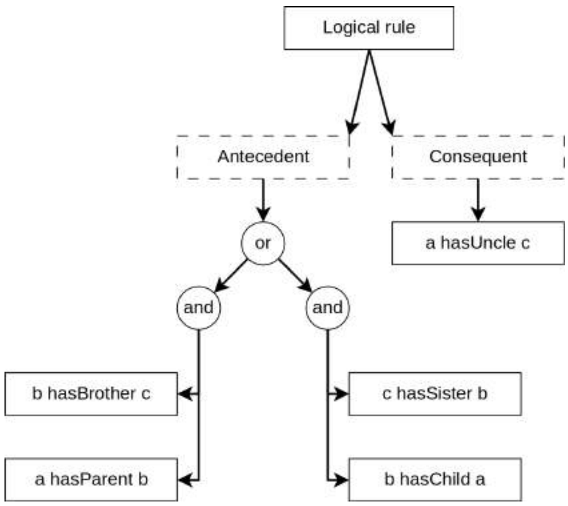


Figure 2. Example of a tree-like representation of a logical rule.

The tree-like logical rule (fig. 2) is translated into the following SWRL:

hasParent(?a, ?b) ∧ hasBrother(?b, ?c) → hasUncle(?a, ?c)
hasChild(?b, ?a) ∧ hasBrother(?b, ?c) → hasUncle(?a, ?c).

and the following Cypher view:

```
MATCH (a:Object)<-[:RANGE]
  -(s:Statementname:"hasParent")
  -[:DOMAIN]->(b:Object)
MATCH (b:Object)<-[:RANGE]
  -(s1:Statementname:"hasBrother")
  -[:DOMAIN]->(c:Object)
MERGE (a)<-[:RANGE]
  -(s2:Statementname:"hasUncle")
  -[:DOMAIN]->(c)
MATCH (b:Object)<-[:RANGE]
  -(s:Statementname:"hasChild")
  -[:DOMAIN]->(a:Object)
MATCH (c:Object)<-[:RANGE]
  -(s1:Statementname:"hasSister")
  -[:DOMAIN]->(b:Object)
MERGE (a)<-[:RANGE]
  -(s2:Statementname:"hasUncle")
  -[:DOMAIN]->(c).
```

Thus, the rules are translated into their tree-view when imported into the KB of logical rules in the SWRL language.

The presence of a tree-like representation of a logical rule allows to form both a SWRL-representation of a logical rule and a Cypher-representation based on it.

Relations of a special type are formed by using Cypher to represent the logical rule between entities of the ontology of the fuzzy KB. These relations correspond to the antecedent atoms of the logical rule. Formed relationships provide the inference from the contents of the fuzzy KB.

IV. Building a Graphical User Interface (GUI) based on the contents of a fuzzy knowledge base

The dynamic graphical user interface (GUI) mechanism is used to simplify the work with KB of untrained users and control of user input [11].

You need to map the fuzzy KB ontology entities to the GUI elements to build a GUI based on the contents of the fuzzy KB. Formally, the GUI model can be represented as follows:

$$UI = \langle L, C, I, P, S \rangle, \quad (2)$$

where

$L = \{L_1, L_2, \ldots, L_n\}$ is a set of graphical GUI components (for example, ListBox, TextBox, ComboBox, etc.);

$C = \{C_1, C_2, \ldots, C_n\}$ is a set of ontology classes;

$I = \{I_1, I_2, \ldots, I_n\}$ is a set of ontology objects;

$P = \{P_1, P_2, \ldots, P_n\}$ is a set of properties of ontology classes;

$S = \{S_1, S_2, \ldots, S_n\}$ is a set of states of ontology objects of fuzzy KB.

The following function is used to build a GUI based on fuzzy KB:

$$F(O) = \{C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i}\} \rightarrow \{L, C, I, P, S\},$$

where

$\{C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i}\}$ is a set of ontology entities of fuzzy KB represented by expression 1 within the i-th context;

$\{L, C, I, P, S\}$ is a set of GUI entities of fuzzy KB represented by the expression 2.

Thus, the contents of the fuzzy KB are mapped to many GUI components. This makes it easier to work with KB for a user who does not have skills in ontological analysis and knowledge engineering. It also allows you to monitor the logical integrity of the user input, which leads to a reduction in the number of potential input errors.

V. Interaction of fuzzy knowledge base with external wiki-resources

At present, wiki-technologies are used to organize corporate KB. It is necessary to solve the task of importing the content of such wiki-resources into fuzzy KB [12], [13]. Table I contains the result of mapping the fuzzy KB ontology entities to the wiki resource entities.

Thus, it becomes possible to import the content of external wiki resources for initial filling of KB contents. There is also the possibility of the reverse process – generation of wiki-resources based on the contents of fuzzy KB.

Table I
Matching ontology elements of fuzzy KB and wiki-resources

The ontology element of fuzzy KB	The element of wiki-resources
Class	Category
Subclass	Subcategory
Object	Page
Class properties	The infobox elements (properties)
Object states	The infobox elements (values)
Relations	Hyperlinks

VI. Conclusion

Thus, the use of fuzzy KB stored in the Graph DBMS in the decision support process presupposes the existence of a certain set of mechanisms:

- organization of inference on the content of fuzzy KB by translating SWRL-rules into Cypher-structures;
- building a graphical user interface based on the contents of fuzzy KB;
- automated import of knowledge from internal and external wiki-resources.

These mechanisms allow to automate the learning process of KB and simplify the work of specialists with KB.

The application of a contextual approach to the storage of knowledge raises the effectiveness of the use of subject ontologies, allowing to adapt the KB to the characteristics of the PrA and to the requirements of specialists. This approach provides them with a tool that is convenient in a software dynamically changeable depending on the contents of the KB.

References

[1] Bova V.V., Kureychik V.V., Nuzhnov Ye.V. Problemy predstavleniya znaniy v integrirovannykh sistemakh podderzhki upravlencheskikh resheniy [Problems of representation of knowledge in integrated systems of support of administrative decisions]. Izvestiya YUFU [News of SFU], 2010, vol. 108, no. 7, pp. 107-113.

[2] Chernyakhovskaya L.R., Fedorova N.I., Nizamutdinova R.I. Intellektual'naya podderzhka prinyatiya resheniy v operativnom upravlenii delovymi protsessami predpriyatiya [Intellectual decision support in the operational management of business processes of the enterprise]. Vestnik UGATU. Upravleniye, vychislitel'naya tekhnika i informatika [Bulletin of the USATU. Management, Computer Science and Informatics], 2011, vol. 42, no. 2, pp. 172-176.

[3] Zagorul'ko YU.A. Postroyeniye portalov nauchnykh znaniy na osnove ontologii [Building portals of scientific knowledge on the basis of ontology]. Vychislitel'nyye tekhnologii [Computing technologies], 2007, vol. 12, no S2, pp. 169-177.

[4] Karabach A.Ye. Sistemy integratsii informatsii na osnove semanticheskikh tekhnologiy [Information integration systems based on semantic technologies]. Nauka, tekhnika i obrazovaniye [Science, technology and education], 2014, vol. 2, no. 2, pp. 58-62.

[5] Tuzovskiy A.F. Razrabotka sistemy upravleniya znaniyami na osnove yedinoy ontologicheskoy bazy znaniy [Development of knowledge management systems based on a single ontological knowledge base]. Izvestiya Tomskogo politekhnicheskogo universiteta [Proceedings of Tomsk Polytechnic University], 2007, vol. 310, no. 2, pp. 182-185.

[6] Filippov A.A., Moshkin V.S., Shalayev D.O., Yarushkina N.G. Yedinaya ontologicheskaya platforma intellektual'nogo analiza dannykh [Unified Ontological Data Mining Platform]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2016, pp. 77-83.

[7] Moshkin V.S., Yarushkina N.G. Metodiki postroyeniya nechetkikh ontologiy slozhnykh predmetnykh oblastey [Methods for constructing fuzzy ontologies of complex subject domains]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2015, pp. 401-406.

[8] Tarasov V.B., Kalutskaya A.P., Svyatkina M.N. Granulyarnyye, nechetkiye i lingvisticheskiye ontologii dlya obespecheniya vzaimoponimaniya mezhdu kognitivnymi agentami [Granular, fuzzy and linguistic ontologies to provide mutual understanding between cognitive agents]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2012, pp. 267-278.

[9] Makhortov S.D. Ob algebraicheskoy modeli raspredelennoy produktivnoy sistemy [On the Algebraic Model of a Distributed Production System]. KII-2016. 2016, vol. 1, pp. 64-72.

[10] Moshkin V.S., Yarushkina N.G. Logicheskiy vyvod na osnove nechetkikh ontologiy [Logical output based on fuzzy ontologies]. Integrirovannyye modeli i myagkiye vychisleniya v iskusstvennom intellekte [Integrated models and soft computations in artificial intelligence], 2015. vol. 1, pp. 259-267.

[11] Gribova V.V., Kleshchev A.S. Upravlenie proektirovaniem i realizatsiei pol'zovatel'skogo interfeisa na osnove ontologii [Managing the design and implementation of the user interface based on ontologies]. Problemy upravleniya [Management issues], 2006, vol 2, pp. 58-62.

[12] Shestakov V.K. Razrabotka i soprovozhdeniye informatsionnykh sistem, baziruyushchikhsya na ontologii i Wiki-tekhnologii [Development and maintenance of information systems based on ontology and Wiki-technology]. Elektronnyye biblioteki: perspektivnyye metody i tekhnologii, elektronnyye kollektsii [Electronic Libraries: Advanced Methods and Technologies, Digital Collections], 2011, pp. 299-306.

[13] Suchanek F.M., Kasneci G., Weikum G. YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. In Proceedings of the 16th International Conference on World Wide Web, 2007, pp. 697-706.

РАЗРАБОТКА ЯДРА ПЛАТФОРМЫ Athene – НЕЧЕТКОЙ БАЗЫ ЗНАНИЙ

А. Филиппов, В. Мошкин, Н. Ярушкина
Кафедра «Информационные системы»
Ульяновский государственный технический университет,
Россия, Ульяновск

В работе описывается процесс разработки ядра платформы Athene – нечеткой базы знаний (БЗ), содержимое которой формируется в результате анализа контекстов проблемной области (ПрО). В данном случае под контекстом понимается некоторая «точка зрения» на ПрО и ее особенности. В качестве основы для хранения содержимого БЗ в виде прикладной онтологии используется графовая база данных (БД). Представлена попытка реализовать механизм логического вывода по содержимому графовой БД. Для упрощения работы с БЗ используется механизм динамической генерации экранных форм интерфейса пользователя.

Features of Development of Internet Resource for Supporting Developers of Intelligent Decision Support Systems

Zagorulko Yu.A.
A.P. Ershov Institute of Informatics Systems
Novosibirsk State University
Novosibirsk, Russia
zagor@iis.nsk.su

Zagorulko G.B.
A.P. Ershov Institute of Informatics Systems
Novosibirsk State University
Novosibirsk, Russia
gal@iis.nsk.su

***Abstract*—The paper describes the approach to development of Internet resource providing support of building of Intelligent Decision Support Systems in weakly formalized areas. In this resource the ontology is used for the formalization and systematization of knowledge, data, facilities for processing and analysis of information integrated in the resource, as well as the organization of convenient content-based access to them. Therefore, the creation of the resource is mainly reduced to the development and population of its ontology. The paper discusses the methodology for developing such an ontology, which includes methods for constructing a target ontology based on a representative set of basic ontologies and ontology design patterns. It is shown that the use of ontology design patterns allows not only to save resources spent on the ontology development, but also to provide a consistent representation of all entities of the ontology, which in turn allows to increase the clarity of the ontology and minimize the number of errors of ontological modeling.**

***Keywords*—information-analytical Internet resource, decision support system, weakly formalized area, ontology, ontology design pattern**

I. Introduction

Support for decision-making in weakly formalized areas, such as medicine, power engineering or education, is currently an urgent problem. The development of decision support systems (DSS) for this class of areas is not an easy task because modern tools for the development of DSS either are not applicable to weakly formalized areas, or are unavailable due to their high cost.

The concept of complex support of the process of developing intelligent DSS in weakly formalized areas [1], which assumes providing support for DSS developers at all stages of its creation, including support at three levels: conceptual, component and information, was suggested in the A.P. Ershov Institute of Informatics Systems of the Siberian Branch of the Russian Academy of Sciences.

At the conceptual level, all groups of specialists involved in the development of intellectual DSS (knowledge engineers, programmers, experts) are provided with a unified system of concepts. In the proposed concept, the ontology of the knowledge area "Support for decision-making in weakly formalized areas" is used for these purposes.

At the component level, DSS developers (programmers) are given direct access to implementations of decision support methods (DS-methods) in the form of services assembled in a special repository.

The information level is provided by the information-analytical Internet resource for decision support (IAIR DS), which provides the DSS developers with all the necessary information about specific DS-methods, the classes of tasks solved by these methods, the possibilities and limitations of each of them, and so on.

The conceptual basis of the IAIR DS is the ontology that serves to formalize and systematize various types of knowledge, data, information resources, information processing and analysis tools integrated into information space of the resource, and to organize convenient, content-based access to them. In view of the fact that ontology plays such an important role in this resource, its creation is reduced mainly to the development and population of its ontology. For this reason, most part of the paper is devoted to the discussion of methods for developing the ontology of the IAIR DS.

II. The Information-analytical Internet Resource for Decision Support

IAIR DS is an information system accessible via the Internet which integrates and systematizes the knowledge and information resources of the knowledge area "Support for decision-making in weakly formalized areas" and the means for processing and analyzing information developed within this area, as well as providing a content-based effective access to him.

As said above, the conceptual basis of the IAIR DS is the ontology containing the formal specifications of the concepts of the knowledge area of the IAIR DS, the types of information resources and methods of supporting decision-making. According to these specifications, structures for representing information on the real objects of the modeled area of knowledge, information resources to be integrated in this resource and of DS methods are being constructed. On the basis of ontology, convenient navigation through the content of the resource and the content-based search of the necessary information are also organized.

The user can navigate along the full content of the IAIR DS, using as a conductor a tree of ontology concepts which built on the basis of a ”general-private” relationship. In this case, it is possible to select an object of any concept (class) and view it in detail, as to transit to other objects related to this relation specified in the ontology. If the object description contains links to information resources or web services, the user can view their descriptions and go to the selected resource or execute the web service required by him.

The user is also has the opportunity to search for objects of a certain class. To do this, he can set restrictions on the values of the attributes of the desired instances of the selected class and on the values of the attributes of the instances associated with them using an ergonomic interface generated on the basis of an ontology.

It should be noted that the IAIR DS provides access not only to descriptions of decision support methods, but also to their implementations stored in the repository of the DS methods. This allows the user to directly test different methods, as well as gain access to their implementations, which can be used by him to create DSS and other applications.

Thus, this resource can be useful to the following types of users:

- developers of intelligent DSS (programmers, knowledge engineers and experts),
- decision-makers,
- researchers developing new methods of decision making,
- persons studying the knowledge area “Theory of decision-making”.

III. Structure of Ontology of IAIR DS

The ontology used in the IAIR DS, along with a description of various aspects of the modeled area of knowledge, contains a description of the structure and typology of information resources and methods for processing and analyzing information associated with this area. In this connection, the ontology of the IAIR DS consists of three interconnected ontologies responsible for the representation of the three components of knowledge above mentioned, namely: ontology of the knowledge area of IAIR DS, ontology of the tasks and methods of DS, ontology of scientific Internet resources.

Ontology the knowledge area of IAIR DS sets a system of concepts and relationships designed for a detailed description of the knowledge area of IAIR DS and scientific and research activities performed in this area.

The ontology of tasks and methods of DS describes the tasks to be solved in the knowledge area of IAIR DS, the methods of solving these tasks and their implementation. In particular, this ontology presents the connections of the decision-making stages with the tasks to be solved at these stages and the methods used for their solution, the connections of the DS methods with the software products (services) implementing them, the connections of such products with the persons and the organization which developed them, with publications describing them and Internet resources providing access to them.

The ontology of scientific Internet resources serves to describe information resources presented in the Internet and relevant to knowledge area of the resource.

IV. Approach to Development of Ontology of IAIR DS

The development of such a complex ontology as ontology of IAIR DS from scratch is not an easy task, therefore for its development methods and tools provided by the technology for creating thematic intellectual scientific Internet resources (ISIR) [2], proposed by the Institute of Informatics Systems, were used.

A. *Method for Development of Ontology*

The essence of the suggested method is to build an ontology based on a small but representative set of base ontologies including only the most general entities that do not depend on specific areas of knowledge. This set includes (see Fig. 1): the ontology of scientific knowledge, the ontology of scientific activity, the base ontology of tasks and methods, the base ontology of scientific information resources.

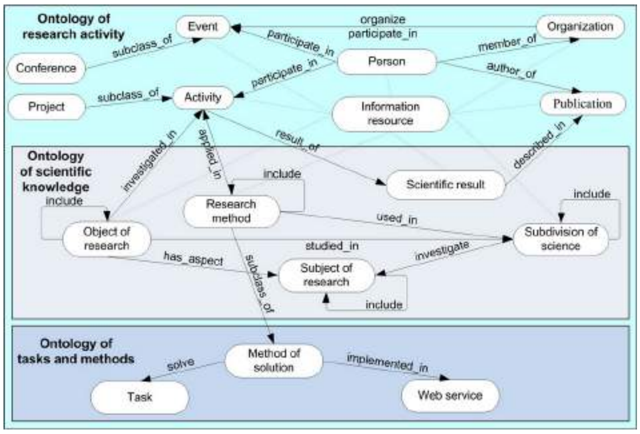


Figure 1. Base ontologies.

The ontology of scientific knowledge contains classes that define structures for describing concepts used in any scientific area of knowledge. Such concepts are the Division of Science, the Object of Research, the Subject of Research, the Method of Research, the Scientific Result, etc.

The ontology of scientific activity includes classes of concepts related to the organization of scientific and research activities, such as Person, Organization, Event, Scientific Activity, Project, Publication, etc.

The base ontology of scientific information resources includes the Information resource class as the main class. The set of attributes and connections of this class is based on the Dublin core standard [3]. To represent information about the resource sources and its co-hosts, as well as related events, organizations, persons, publications and other entities, special relations are introduced that connect the Information resource class with the classes of other ontologies of the resource.

The base ontology of tasks and methods includes classes such as Task, Solution Method, and Web Service. Using the concepts and relations of a given ontology, we can describe tasks that can be solved in the knowledge area of the resource being created, methods for solving such tasks and web services implementing these methods.

In accordance with the proposed methods, the ontology of the knowledge area is constructed on the basis of the first two base ontologies, the ontology of the tasks and methods of the DS – on the basis of the third base ontology, the ontology of scientific Internet resources – on the basis of the fourth base ontology.

B. Ontology Design Patterns Used in Design of IAIR DS Ontology

As said above, the IARP DS is developed using the technology for creating thematic intellectual scientific Internet resources [2]. The base ontologies provided by this technology are developed in OWL language [4]. The use of this language caused problems associated with the fact that the proposed concept of knowledge and data representation in the ISIR is not always convenient to fit into the capabilities of the OWL language. To solve these problems, as well as other problems arising in the development of ontologies, we use ontology design patterns.

Ontology design patterns (ODPs) [5], [6] have been used for more than ten years to streamline the development of ontology and reduce its laboriousness. ODPs are documented descriptions of proven solutions of problems of ontological modeling [7]. ODPs allow us to describe both typical and specific problems arising in the development of ontologies, as well as recommendations and agreements proposed by the developers for their solution. The main catalog of such patterns is presented on the portal of the Association for Ontology Design & Patterns (ODPA) created within the framework of the NeOn project [9].

Depending on the type of problems for solution of which the patterns are designed, there are several types of patterns: structural patterns, correspondence patterns, content patterns, logical inference patterns, presentation patterns and lexico-syntactic patterns [8], [9]. In this paper, we will focus only on those that were used in the development of ontologies for the IAIR DS, namely: structural logical patterns, content patterns and presentation patterns.

The need to use structural logical patterns arose due to the lack of expressive means in OWL for the presentation of complex entities and constructions that are relevant to the construction of the IAIR DS ontology. The technology of creation of thematic INIR offers structural patterns for representing such complex constructions necessary for the construction of ontologies, as domains (ranges) of admissible values, attributed and n-ary relations, relations between the object of one class and the class of other objects.

When creating the ontology of the knowledge area of the IAIR DS, it is important to unambiguously and consistently represent scientific concepts and their properties used in basic ontologies. To solve this problem, content patterns defining ways of representing typical ontology fragments are required. Ontologies of a whole class of subject domains can be built on the basis of such patterns.

When developing the ontology of the DS knowledge area, content patterns developed earlier for the implementation of projects for the creation of scientific Internet resources for a number of scientific subject areas are used. On the basis of such patterns, basic ontologies of scientific activity and scientific knowledge are built. These patterns, serving to represent objects and methods of research, performed activities, scientific results, etc. (see Fig. 2 and 3), are common for many scientific subject areas. Such patterns, in turn, can be part of more complex content patterns.

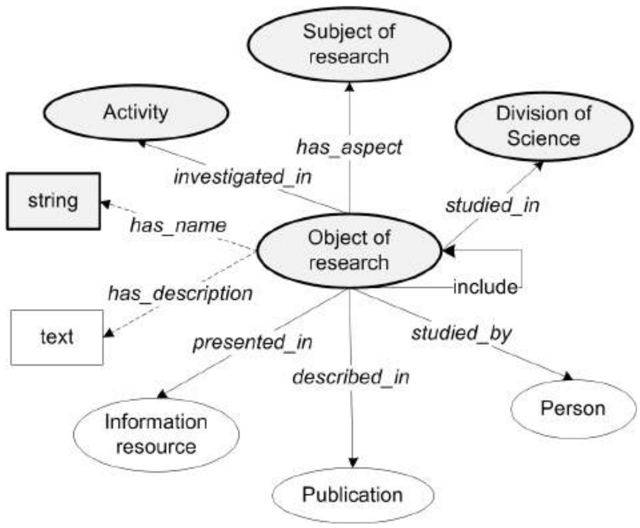


Figure 2. Object of research pattern.

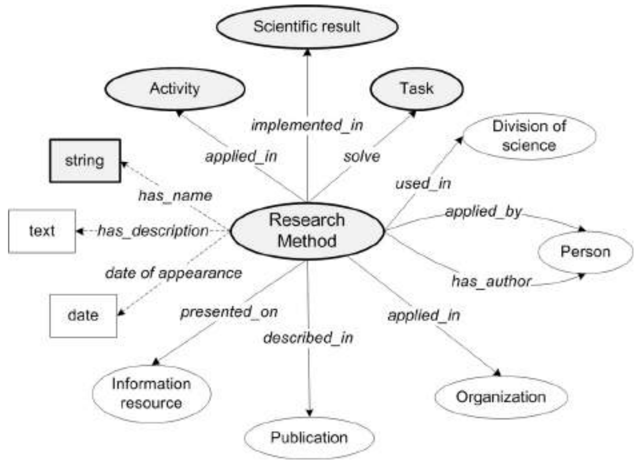


Figure 3. Method of research pattern.

It should be noted that the content patterns include not only the concepts presented by them, but also concepts from "adjacent" patterns, as well as patterns of other base ontologies, which allows you to specify a related description of the modeled area.

In fact, the content patterns are fragments of base ontologies that become constituent parts of the ontology of the modeled area of knowledge after supplementing with new concepts and specializing the concepts contained in them.

It is very important to provide a user-friendly and understandable ontology view. For these purposes, presentation patterns that define conventions for naming, annotating, and visualizing ontology elements for the end user are used.

The technology of thematic ISIR creation provides the possibility to customize the display of information objects (ontology class instances) included in the ISIR content when they are displayed on the monitor screen. For these purposes, it offers and implements the presentation pattern mechanism [2], which allows you to customize the visualization of objects of selected classes when they are displayed to users or editing.

V. Conclusion

The paper considers the approach to the development of an information-analytical Internet resource that supports the process of creating intelligent DSS in the weakly formalized areas. Its peculiarity is that the creation of this resource basically boils down to the development and populating of its ontology, which serves not only the purposes of formalizing and systematizing knowledge, data, processing and analysis tools that are integrated in the resource, but also used to organize user-friendly content access to them.

During the development of the IAIR DS, methods and components of the technology for creating thematic intellectual scientific Internet resources based on the ontological approach [2] and Semantic Web tools [10] were used.

It should be noted that the use of patterns of ontological design in the development of IAIR DS allowed to save resources in the development of ontologies and provide a consistent representation of all ontology entities, which in turn allowed minimizing the number of ontological modeling errors, increasing the clarity of ontology and provide the possibility of collective development of ontologies.

Acknowledgment

This work was supported by the Russian Foundation for Basic Research (grant No. 16-07-00569).

References

[1] Zagorulko, G.B., Zagorulko Yu.A. Podkhod k organizatsii kompleksnoi podderzhki protsessa razrabotki intellektual'nykh SPPR v slaboformalizovannykh predmetnykh oblastyakh [The approach to the organization of integrated support for the process of developing intelligent DSS in weakly formalized subject areas]. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, 2016, pp.61-64 (in Russian)

[2] Zagorulko Yu.A., Zagorulko G.B., Borovikova O.I. Tehnologija sozdanija tematicheskih intellektual'nyh nauchnyh internet-resursov, bazirujushhajasja na ontologii [Technology for building subject-based intelligent scientific internet resources based on ontology]. *Programmnaja inzhenerija [Software Engineering]*, 2016. no. 2, pp. 51-60 (in Russian).

[3] Hillmann D. Using Dublin Core. Available at: http://dublincore.org/documents/usageguide/ (accessed 2017, Nov)

[4] Antoniou G., Harmelen F. Web Ontology Language: OWL. *Handbook on Ontologies*. Berlin: Springer Verlag, 2004. pp. 67-92.

[5] Gangemi A. Ontology Design Patterns for Semantic Web Content. Gil Y., Motta E., Benjamins V.R., Musen M.A. (eds). *The Semantic Web - ISWC 2005. Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 2005. vol. 3729. pp. 262-276.

[6] Gangemi A., Presutti V. Ontology Design Patterns. *Handbook on Ontologies*. Springer, 2009. pp. 221-243.

[7] Hitzler P. Ontology Engineering with Ontology Design Patterns: Foundations and Applications. *Studies on the Semantic Web*. Hitzler P., Gangemi A., Janowicz K., Krisnadhi A., Presutti V. (eds.) IOS Press/AKA, 2016.

[8] NeOn Project. Available at: http://www.neon-project.org/ (accessed 2017, Nov)

[9] Hammar K. *Towards an Ontology Design Pattern Quality Model. Linköping Studies in Science and Technology*. Linköping University, 2013. vol. 1606. 159 p.

[10] Hitzler P. , Krötzsch V. , Rudolph S. *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, 2009. 455 p.

ОСОБЕННОСТИ РАЗРАБОТКИ ИНТЕРНЕТ-РЕСУРСА ДЛЯ ПОДДЕРЖКИ РАЗРАБОТЧИКОВ ИНТЕЛЛЕКТУАЛЬНЫХ СППР

Загорулько Ю.А., Загорулько Г.Б.
Институт систем информатики имени А.П. Ершова СО РАН
Новосибирский государственный университет
г. Новосибирск, Россия

Рассматривается подход к разработке интернет-ресурса, обеспечивающего поддержку процесса создания интеллектуальных систем поддержки принятия решений в слабоформализованных областях. Данный ресурс представляет собой доступную через Интернет информационную систему, интегрирующую и систематизирующую знания и информационные ресурсы области знаний «Поддержка принятия решений в слабоформализованных областях» и разработанные в рамках этой области средства обработки и анализа информации, а также обеспечивающую содержательный эффективный доступ к ним. Для формализации и систематизации знаний, данных, средств обработки и анализа информации, интегрируемых в данном ресурсе, а также организации удобного содержательного доступа к ним используется онтология. В связи с этим создание данного ресурса в основном сводится к разработке и наполнению его онтологии.

В статье обсуждается методология разработки такой онтологии, которая включает методы построения целевой онтологии на основе представительного набора базовых онтологий (путем их специализации и дополнения новыми понятиями) и паттернов онтологического проектирования. Показано, что использование паттернов онтологического проектирования позволяет не только сэкономить ресурсы, затрачиваемые на разработку онтологии, но и обеспечить согласованное представление всех ее сущностей, что, в свою очередь, позволяет повысить «понимаемость» онтологии разработчиками и минимизировать число ошибок онтологического моделирования.

One Managerial Approach to Knowledge Representation: Developing a Knowledge Map

Elvira Grinberg
Graduate School of Management
St. Petersburg University
St. Petersburg, Russia
st057454@student.spbu.ru

***Abstract*—Knowledge representation potential is understudied in such area as enhancing knowledge management processes. This work offers an overview of the existing approaches for mapping business knowledge. Furthermore it discusses the problems of knowledge maps created by managers and indicates their advantages and disadvantages. It was revealed that managers have shortage of competencies in information structuring and underestimate the role of ontology. That is a significant obstacle in the implementation of knowledge models. The paper supposes a systematic approach to the construction of knowledge map to overcome that issue a in a knowledge-intensive organization. It is taking into account specifics of Russian companies on one hand and low knowledge engineering expertise of managers on the other hand.**

I. Introduction

There is a limited amount of Knowledge Representation literature related to business. It deeply explores several distinct areas of management such as business process automation, digitalization, and decision support systems development. That stream of research is tied to the information theory and IT architecture of the company. However, KR potential is understudied in such area as enhancing knowledge management processes. That stream of research invades cognitive psychology and human resource management. The present paper aims to bridge these fields of studies via literature synthesis and case study analysis. The focus of present work is knowledge map development as it is a starting point of knowledge management system implementation. One of the obstacles of business-knowledge map development is large variety of definitions and approaches to that subject. That is why it is necessary to establish common language not only inside the managerial body of literature but also across the domains. The family of ISO standards can be one of the possible common grounds for both management field and knowledge engineering field as the standards are aimed at broad number of business issues and are applicable to any kind of firm. According to the ISO 9001:2015 standard [1], knowledge is one of the main resources of the organization. Knowledge mapping allows to approach knowledge management systematically. Regular mapping, categorizing, and bench-marking of organizational knowledge can not only make it available for all employees, but also get greater value from the efforts to prioritize and focus in organizational learning.

At the moment, researchers focus on developing variety of software tools implementing the mapping procedure even for graphs with thousands of vertices [2], and displaying the business processes using a common notation [3]. However, it is often overlooked that a crucial role in the processing of available data into knowledge is played by the skills of the manager to summarize, interpret, and organize information. Lack of such skill effects negatively the quality of management decisions. Present paper suggests an approach which partially covers that gap.

II. Existing Approaches to Knowledge Mapping

Variety of knowledge-map related terms and approaches was studied in the previous work (Gavrilova, Alsufyef, Grinberg 2017). This section briefly presents evolution of the term and discusses the purposes of a knowledge map.

A. The evolution of the "knowledge map" term

Some top level terms in knowledge management have been already established and fixed in an above-mentioned standard. Knowledge of the company is defined in ISO standard as the combination of staff expertise, which is recorded and not recorded at the corporate tangible and intangible media.[1] Knowledge management is defined as a unified approach or a group of methods of knowledge creation, storage, protection and dissemination. However, the definition of term "knowledge map" is either vague and includes any type of knowledge representation (from Gantt charts to metaphors) or it is very narrow and supposes answer to one specific question only (“Where can I find a certain knowledge?”)[4].

The term "knowledge map" appears in business literature in the end of XX century but authors do not clearly define it. Knowledge map denotes different types of charts, graphs, matrices and other graphic objects. In some studies it is defined as an analogue of the plan of the area that allows manager to figure out where and what knowledge can be found, while other authors denote by this term the totality of diagrams that present knowledge.

Among the trends of this area it is highlighted the desire of researchers to expand the study of applications of knowledge mapping in various business practices. In the processing of information (for example, in the context of Big Data) the use of mapping technique as clear and succinct presentation of data allows one to go to deeper levels of interpretation. Till now the scientific literature has not formed a uniform classification of knowledge representation techniques. As a result, practitioners work with the knowledge representation, not knowing which method to use in order to solve certain business problems.

In 2010s managerial literature converges to a common understanding of the "knowledge map" term. It has happened ten years after classical definitions were given by seminal authors. Today's common understanding is that the creation of knowledge map involves locating important knowledge of organization and representation that shows where to find that knowledge. Knowledge map points to the people, the documents, and databases. Knowledge Map is a navigation assistant for tacit and explicit knowledge that illustrates the knowledge flows in the organization [5]. Knowledge map describes the sources, flows, and limits of knowledge organization. Knowledge maps help to understand the relationship between knowledge stores and dynamics [6].

B. Purpose of a knowledge map

The company can group its learning experience in critical mass around certain strategic areas of knowledge [7]. Knowledge map can address a wide range of questions. Some most common a re listed below.

1) Identifying the knowledge necessary for the functioning of the business processes: The main goal of any knowledge map is identification of knowledge assets and their location. Availability of knowledge assets becomes one of the main issues for practitioners and regulatory organizations [8]

2) Spotting areas of potential competencies' development: It becomes possible to compare existing competencies of employees with those which are necessary to implement the strategy. That analysis yields a clear understanding of staff learning and development strategy.

3) Identifying areas of organizational knowledge that need additional development: Knowledge gaps remain invisible until knowledge map is developed. Creation a list of missing knowledge is a first step in bridging those gaps. After that an action plan should be developed. It includes several options besides employee development mentioned above. The gaps can be closed via knowledge acquisition from external sources: hunting necessary experts, inviting consulting company, employing external trainers, buying patents, databases and software, etc.

4) Regular evaluation of the current level of knowledge: Knowledge audit should be conducted at least once a year; it should be done more frequently in the knowledge intensive firms. Knowledge update and development should be paired with organizational forgetting[9]. Out-of-date knowledge causes significant problems in a knowledge intensive firm which works in a turbulent, frequently changing environment.

5) Knowledge commercialization: Knowledge map shows areas of strong expertise of the company. Knowledge intensive company sells products and knowledge. That is why it is important to inventory all knowledge and reveal those results of intellectual activity which can be commercialized in form of patents or services for external clients.

III. Problem areas of knowledge maps developed by managers

Most scientific articles suggest that knowledge map is developed by an experienced professional. However, it is often not the case. We consider the situation when a knowledge map is developed by a company managers who usually lack knowledge audit experience and had taken none knowledge engineering courses. This section discusses of most common misconceptions related to the subject.

A. "We do not need an ontology"

Many managers misinterpret the term "ontology". Moreover, they highly underestimate the importance of business glossary as a basis for knowledge management processes. For instance, knowledge map which is a result of knowledge audit [4] can be developed in a company without any ontology. Despite the fact that half of the methods of a knowledge map building underline the importance of preliminary ontology development [10], managers tend to get around those acute angles. The reason for that is an absence of the relevant expertise. Knowledge engineering course is usually not included in the MBA or equivalent programs. Ontology or conceptualization of the specification [11] is a hierarchical domain model with solid mathematical and programming study. Without the ontology it is rather difficult to build any representation, especially if the number of elements on the lower level is broad as it is usually a case in big organizations.

B. "I know everything"

This mistake is a straightforward result of the previous one. Since manager can collect but cannot aggregate knowledge from employees and develop an ontology, he or she develops knowledge representation based on personal perception. Such map answers another question: "What does manager know about employees' knowledge". There are several positive aspects of such approach. Firstly, it allows a quick building of a knowledge representation without time-consuming knowledge elicitation from employees. Secondly, it stimulates knowledge structuring process at least of one manager. Still, it is not the representation of organizational knowledge, it reflects an individual point of view only.

C. "Any hierarchy can fit"

Another way out from the absence-of-ontology trap is taking one of the existing hierarchies; for instance, it can be competence list, business processes list or production cycle as a basis for the knowledge representation. This approach creates a solid basis for a well-structured knowledge representation. If a manager lacks synthesizing skills that can be a recommended

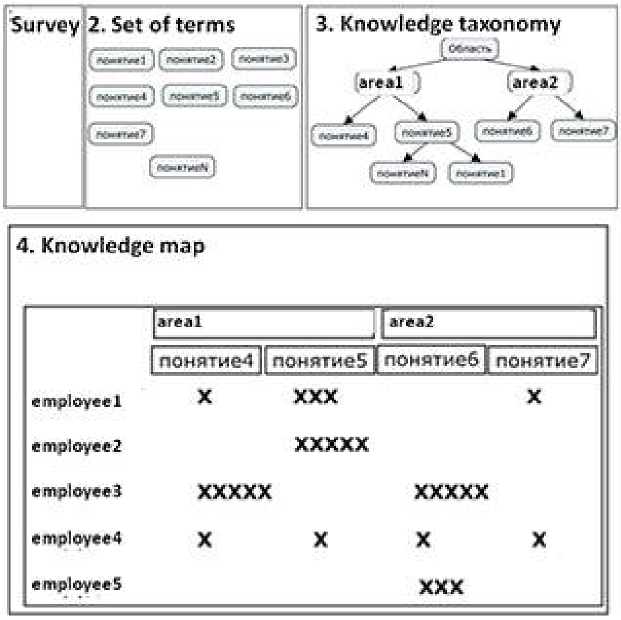


Figure 1. Proposed steps of knowledge map development

way of action. However, the result of such KR-development is questionable since every ontology has a very specific focus question and should be built around it. Thus, if one takes taxonomy developed for other purposes he or she most likely will get an irrelevant result. For instance, a manager develops a map of competencies instead of a knowledge map. It is related to knowledge but it is not the knowledge itself. Such map is misleading.

IV. The Proposed Approach to the Specified Problem

Present section suggests a sequence of steps towards proper knowledge representation developed. They are as follows: conducting a survey, which becomes a source of key concepts; developing taxonomy of those concepts; developing a visual representation of the taxonomy in a form of a knowledge map (see Figure 2 for details).

A. Conducting a survey

The main part of the mapping knowledge is conducting a survey among employees. The survey can be developed on the basis of the approved scientific papers [4][6][10] taking into account the specifics of the company. It considers existing business glossary of the firm and uses a database of a company employees. The list of questions should be fixed in the form of a questionnaire template. The result of mapping is creating a knowledge map in accordance with the preliminary established way of visualization. Knowledge map answers the question "Who owns what knowledge and where it is?". Thus, the survey ensures availability of knowledge to the extent necessary. The analysis of the survey results among professionals of one department allows us to develop a taxonomy of knowledge a department. The basis of taxonomy are the answers to the questions of the survey. The purpose of these questions is to identify key knowledge areas that employees are willing to share and/or consider significant. Specified areas of expertise form the basis for building a hierarchical structure of knowledge functions, which not only lists the knowledge possessed by a Company, but also describes the relationship between these elements. This provides a basis to identify gaps in knowledge.

B. Developing a taxonomy

The taxonomy helps the user of knowledge by making it easier to find and use someone else's experience. Common taxonomies distinguish between explicit and implicit, General and specific, depending on context, personal and situational knowledge. Knowledge can also be characterized by type: descriptive ("know") and procedural ("know how"), causal ("know why"), knowledge about the terms ("know when") and attitudinal ("you know who" or "know who"). These categories allow managing knowledge at the process level. After type of desired knowledge representation is agreed upon, a manager can follow several steps to build a taxonomy as described in [12]. First, it is necessary to allocate of the basic concepts of the subject area. The set of basic concepts is based on employees ' responses to the above-described survey. Then it is identified the number of levels of the taxonomy and the preliminary distribution of concepts by levels. Further it is necessary to build relationships between concepts. Finally, the manager consults with various specialists to avoid contradictions and inaccuracies. It is done via the procedures of removal of synonymy and selection of commonly accepted meta-concepts. The last stage is iterative, after it one may need to return to the previous stages and reconsider relationships between concepts, or even redistribute them across levels.

C. Developing a knowledge map representation

Knowledge map representation is carried out based on the taxonomy of knowledge. Knowledge map can be realized as a table in MS Excell with some visual elements [13]. It is supposed to be available to each employee. Graphical elements denote two dimensions: the level of proficiency in a particular area and willingness to share. The level of proficiency in a particular area is determined subjectively. Further validation of this level can be done by a supervisor and/or on the basis of annual evaluation of personnel. It shall be entered in the appropriate cell at the intersection of the row "employee name" and column "knowledge". Name of knowledge domain is identical with that presented in the developed taxonomy. The levels of proficiency are displayed in the table as follows: 1 - beginner, 2 - basic, 3-qualified, 4 - professional, 5 - expert. For clarity, the level of proficiency can be displayed in the form of horizontal histograms using conditional formatting in MS Excell. Willingness to share defines the degree of readiness of an employee to share a particular knowledge. It is entered in the appropriate cell at the intersection of the row "employee name" and "Knowledge" column (same as above). It is denoted by colors as follows: a) green means "ready to share knowledge", b) yellow stands for "may not always share knowledge", and c) red denotes absence of desire to share knowledge.

V. Conclusion

Knowledge map can be applied at two levels: operational and strategic. In the first case, the activities of knowledge management are aimed at understanding and disseminating their existing knowledge. The premise is that the Company has a broad knowledge and necessary to identify them. Thus, the core operating activities is the establishment of a mechanism for detection/mapping and dissemination of knowledge through the intranet. Following presentations of knowledge map can be developed at the operational level:1) Map knowledge necessary for business processes, 2) Map knowledge of the regions, 3) Map geographical location of the knowledge holders, 4) Visual search of knowledge and others. That can be done by mapping the developed knowledge map with other ontology or image. The choice will depend on the certain need of stakeholders. The strategic application of knowledge map is not only the cataloguing of existing knowledge, but also in comparison them with the knowledge necessary for the formulation and implementation of the strategy of the organization. In addition, the card can be used to compare the knowledge of the Company with knowledge of competitors (Zack 2009). The main issue identified in the present research is lack of specialized engineering knowledge in a managerial skill-set. The proposed solution to this problem is to involve an expert on knowledge engineering to build the taxonomy. Thus, knowledge representation in a form of a knowledge map is a convergence point for the strategic knowledge management and knowledge engineering.

References

[1] Wilson J. P. et al. Developing a knowledge management policy for ISO 9001: 2015. Journal of Knowledge Management, 2016. No 4(20). pp. 829-844.

[2] Kolomeychenko M.I., Chepovskiy A.M. Vizualizatsiya i analiz grafov bol'shikh razmerov [Huge graph visualization and analysis]. Business Informatics, 2014. No 2(30), pp. 7–16 (in Russian). (37). P. 62 to 71.

[3] Fiodorov I.G. Overcoming expressiveness deficit of business process modeling languages. Business Informatics, 2016. No 3(37), pp. 62–71.

[4] Liebowitz J. et al. The knowledge audit. Knowledge and process management, 2000. No 1(7). p. 3.

[5] Davenport T. H., Prusak L. Working knowledge: How organizations manage what they know. Harvard Business Press, 1998. p. 74.

[6] Grey, D. Knowledge mapping: a practical overview. SWS Journal, 1999. Retrieved from http://kmguru.tblog.com/post/98920, Accessed January 18, 2017.

[7] Zack M., McKeen J., Singh S. Knowledge management and organizational performance: an exploratory analysis. Journal of knowledge management. 2009. No6(13). pp. 392-409.

[8] Wilson, J. P., Wilson, J. P., Campbell, L., Campbell, L. Developing a knowledge management policy for ISO 9001: 2015. Journal of Knowledge Management, 2016. No 20(4), 829-844.

[9] Argote, L., Ingram, P. Knowledge transfer: A basis for competitive advantage in firms. Organizational behavior and human decision processes, 2000. No 82(1), 150-169.

[10] Balaid A. Knowledge maps: A systematic literature review and directions for future research. International Journal of Information Management. 2016. No3(36). pp. 451-475.

[11] Gruber T.R. A translation approach to portable ontologies. Knowledge Acquisition, 1993. Vol. 5, No2, pp. 199–220.

[12] Gavrilova T., Farzan R., Brusilovsky P. One practical algorithm of creating teaching ontologies. 12th International Network-Based Education Conference NBE, 2005. pp. 29-37.

[13] Eppler M. J. Making knowledge visible through knowledge maps: concepts, elements, cases. Handbook on Knowledge Management 1. Springer Berlin Heidelberg, 2004. pp. 189-205.

ОДИН УПРАВЛЕНЧЕСКИЙ ПОДХОД К ПРЕДСТАВЛЕНИЮ ЗНАНИЙ: РАЗРАБОТКА КАРТЫ ЗНАНИЙ

Эльвира Гринберг
Высшая Школа Менеджмента, СПбГУ,
Санкт-Петербург, Россия

Потенциал представления знаний недостаточно изучен в такой области, как совершенствование процессов управления знаниями. Эта работа предлагает обзор существующих подходов к отображению бизнес-знаний. Кроме того, обсуждаются проблемы карт знаний, созданных менеджерами, и указываются их преимущества и недостатки. Выявлено, что менеджеры испытывают дефицит компетенций в структурировании информации и недооценивают роль онтологии. Это является серьезным препятствием на пути внедрения моделей знаний. В статье предполагается системный подход к построению карты знаний для преодоления этой проблемы в наукоемкой организации. Она учитывает специфику российских компаний, с одной стороны, и низкий уровень компетенций менеджеров в области инженерии знаний, с другой стороны.

Ontological modeling using the system «Binary Model of Knowledge»

Gerald S. Plesniewicz
National Research University MPEI
Moscow, Russia
salve777@mail.ru

Dmitry E. Masherov
National Research University MPEI
Moscow, Russia
masherovdy@mail.ru

***Abstract*—In the paper, a brief description of the ontological modeling system «Binary Model of Knowledge» (BMK) is given. Using the system, experts have an opportunity to create heavy-weight ontologies for complex systems. A case study of applying BMK to development problems of computer systems for railway safety is presented.**

***Keywords*—ontological modeling; knowledge bases; highweight ontologies; railway safety systems**

I. Introduction

An ontology aims at representing knowledge (at conceptual level) of the problem domain and the functionality of a given modelled system.

«Binary Model of Knowledge» (BMK) is an ontological modeling system which is under development at National Research University MPEI (Moscow) and at Institute of Information and Computing Technologies of MES RK (Almaty).

BMK is supported by concept-type languages which provide a friendly interface for experts who are busy in developing ontologies.

In Section 1 we describe briefly the main languages of the system «Binary Model of Knowledge». In Section 2 we present the case study of ontological modeling related to the problem of developing a computer railway safety system.

II. Brief description of the ontological modeling system «Binary Model of Knowledge»

Ontological modeling is an activity in creating ontologies for subject domain and functionalities of applications. For ontological modeling, the methods and methodology of ontological engineering are used [2, 3, 5].

BMK is the ontological modeling system supported by conceptual-type languages which are well readable and friendly to experts-ontologists.

Semantics of BMK languages is based on the formal model of concept [3]. Formal concepts are constructed from names. A *formal concept* with the name C has the following components: (1) set U^C – the *universe* of the concept; (2) set Γ of so-called *points-of-reference*; (3) set $E^C\gamma$ for each $\gamma \in \Gamma$ – the set of *instances* of the concept at point-of-reference γ; (4) reflexive, symmetric and transitive relation $\tilde{}^C\gamma \subseteq E^C\gamma \times E^C\gamma$ for each $\gamma \in \Gamma$ – *coreferentiality* relation; (5) pair $(E^C\gamma, \tilde{}^C\gamma) = Ext^C\gamma$ for each $\gamma \in \Gamma$ – the *extension* of the concept at point-of-reference γ; (6) set family $\{(E^C\gamma, \tilde{}^C\gamma) \mid \gamma \in \Gamma\}$ – the *full extension* of the concept.

T. Gruber define an ontology as «explicit specification of conceptualization» [2]. A conceptualiz/ation includes a choice of suitable concepts and relations between them. (Let us note that the relations can be considered also as concepts.) We define a *formal conceptualization* as a (finite) set of formal concepts with the same set Γ of points-of-reference.

For formal specification of formal concepts suitable formal languages are used. There are different types of sentences in such languages, depended on what components of formal concept they specify. *Structural sentences* specify concept universes; *logical sentences* specify concept extensions; *transitional sentences* specify modifications of concept extensions.

BMK includes the language **LS** for specifying ontologies and its sublanguages **LS-S**, **LS-L**, **LS-T** and **LS-DT** which consist in structural, logical, transitional sentences, and data type specifications (correspondingly). In **LS**, two kinds of concepts are used: *classes* and *binary links* (*relations*).

A. Structural sentences

Elementary **LS-S** sentences are expressions of the forms:

$$C[E], C[A:E], C[A:T], (C\ L\ D), (C\ L\ D)[E], \\ (C\ L\ D)[A:E], (C\ L\ D)[A:T]. \quad (1)$$

Here C, D, E are names of classes, L is a name of a binary link, A is an attribute name, and T is a data type specification.

C and $(C\ L\ D)$ in elementary structural sentences (1) are their *heads*, and $[E]$, $[A{:}E]$, $C[A{:}T]$ are their *tails.* Several elementary statements with identical heads can be merged into one statement joining their tails. For example, the elementary statements $C[E]$, $C[A{:}D(*)]$, $C[B{:}$(String,Integer)] are merged into one statement $C[E{,}A{:}D(*)$, $B{:}$(String,Integer)].

A *structural scheme* is a (finite) set of structural sentences.

Example 1.

```
SCHEME EducationInfo
1. Student[Name:String,
        Belongs_to:StudGroup].
2. StudGroup[Title:String,Tutor:Student].
3. (Student PassedExam Course).
4. PassedExam[ExamList].
```

```
5. Teacher[Name:String,
        Position:{prof,assist},
        Works_at:Department].
6. (Course Conducted_by Teacher).
7. Course[Name:String].
8. Department[Name:String,Staff:
             (Teacher OR Ingineer)(*)]
9. ExamList[Name-of-course:String,
          Name-of-student:Dtring,
          Date-of-exam:Date].
END
```

The statements from (1) define (correspondingly) the following universes of C and L:

U^C = Surr U Iname U {[E:x] | x ∈ Surr},
U^C = Surr U Iname U {[A:x] | x ∈ Surr},
U^C = Surr U Iname U {[A:x] | x ∈ T},
U^L = Surr U Iname U{[C:x, D:y] | x, y ∈ Surr},
U^L = Surr U Iname U {[C:x,D:y,E:z] | x, y, z ∈ Surr},
U^L = Surr U Iname U {[C:x,D:y,A:z] | x, y, z ∈ Surr},
U^L = Surr U Iname U {[C:x,D:y,A:z] | x, y ∈ Surr, z ∈ T}.

Here Surr is the set $\{\#1, \#2, \#3, \ldots\}$ of so-called *surrogates* (object identifiers) and Iname is the set of object names.

For example, the following tuple belongs to $U^{Teacher}$, where #27 is the surrogate of some department:
e = [Name:V.Falk',Position:prof, Work_at:#27].

In a structural scheme, every concept from a structural sentence with a tail defines derived concepts. For example, for the concept Teacher we have the derived concepts

```
  Teacher(Name = 'V.Falk'),
  Teacher(Name = 'V.Falk; Positon =
prof),
  Teacher(Name=prof;Work_at = #27),
  Teacher(Work_at.Name = AppliedMath)
```

and so on. The first term denotes the teacher V. Falk. The second term denotes the teacher V. Falk which is a professor. The third term denoes set of professors working an the department with the surrogate #27. The third term denotes the set of all teachers working at the Applied Mathematics department.

B. Representation of facts

A *fact* is a statement about belonging a given object e to instances of a given concept C at a given point-of-reference γ, i.e. $e \in E^C_\gamma$ or $e \in U^C \setminus E^C_\gamma$. The first fact is *positive* and the second fact is *negative.* In the language **LS**, the facts are written as follows:

- $+$(e IN C / POR γ) and –(e IN C / POR γ) when C is a class. Here POR is the abbreviation for «point-of-reference»;
- $+$(e_1 L e_2 C / POR γ) and –(e_1 L e_2 C / POR γ) when L.

Here POR is the abbreviation for «point-of-reference», and the signs «+» and «–» denote that fact are positive and negative.

In the BMK, a tabular representation of facts data representing first sort facts is used.

Example 2. Consider the sentences 3 from Example 1 for the binary relation PassedExam. Table I is an example of a tabular representation of a set of facts for this relation.

Table I
TABULAR REPRESENTATION OF THE BINARY RELATION «PASSEDEXAM»

PassedExam				
Surr	Sign	Coref	Student	Course
.....	....	....	....	....
#33	+	[exam4]	#10	#24
#34	-	[]	#10	#27
.....	....	....	....	....

The rows from Tables I represent the facts: «The student with the surrogate #10 passed the exam on the course with the surrogate #24, and does not pass the exam with the surrogate #27». Also here we have the coreferentialities:

#33 ~exam4 ~[Student:#10, Course:#24],

The language **LS** includes conjunctive queries.

Example 3.

Query 2. What students from the student's group A13-09 passed examination in algebra at 15-01-2017 and received the assessment 75?

```
?X.Name --(X Passed_exam:Y Course);
    X.Belongs_to.Title = A13-09;
    Y.ExamList.Name-of-course = algebra;
    Y.ExamList.Data = 15-01-2017;
    Y.ExamInfo.Assesment = 4.
```

C. Logical sentences

Logical sentences of the language **LS-L** are constructed of three type terms: C-*terms*, L-*terms* and P-*terms*. C-*terms* denote subsets of the set Surr, L-*terms* denote subsets of the set Surr × Surr, and P-*terms* denote unary predicates defined on Surr.

Consider several examples of C-*terms*, L-*terms*, P-*terms* and **LS-L** sentences.

Examples 4.

1) Student, Student(Name = 'A. Kotov), Student(Belongs_to.Title = A13-11), Belong_to.Name = A13-11) OT PassedExam SOME Course,
2) PassedExam SOME Course THAT Conducted_by Teacher(Name='V.Falk'),
3) Course(Name=CraphTheory)Conducted_by.

The first P-*term* presents the predicate which is true for every student that did not pass at least one exam. The third P-term presents the predicate which is true for all teachers who conducted the course «graph theory».

We use the following metavariables in the description of syntax and semantics of terms and sentences of the language **LS-L**: c for individual constant; C, D, E for classes; L, M, N for binary links; P, Q, R for predicates; S, S_1, S_2 for sentences; V for parameters.

For any expression *exp*, we denote by $\|exp\|$ its value under a given interpretation. For concept and binary link names we suppose $\|C\| \subseteq$ `Surr` and $\|L\| \subseteq$ `(Surr, Surr)`.

Syntax of C-terms:

C, D::= `NOT` C | $C:V$ | (C `AND` D) | (C `OR` D) |
C `THAT` P | C `SOME` L | C `ONLY` L | $L.1$ | $L.2$.

Here $L.1 = D_1$ and $L.2 = D_2$ where $(D_1\ L\ D_2)$ specifies L.

Semantics of C-terms:

$\|$ `NOT` $C\| =$ `Surr` $\setminus \|C\|$,
$\|C:V\| = \|C\|$,
$\|C_1$ `AND` $C_2\| = \|C_1\| \cap \|C_2\|$,
$\|C_1$ `OR` $C_2\| = \|C_1\| \cup \|C_2\|$,
$\|C$ `THAT` $P\| = \{x \in \|C\| \mid \|P\|(X)\}$,
$\|C$ `SOME` $L\| = \{x \in$ `SurrR` $\mid \exists y \in$ `Surr`.
$Y \in \|C\| \wedge (x,y) \in \|L\|\}$,
$\|C$ `ONLY` $L\| = \{x \in$ `Surr` $\mid \forall y \in$ `Surr`.$(x,y) \in \|L\| \rightarrow y \in \|C\|\}$,
$\|L.1\| = \|D_1\|$,
$\| L.2 \| = \| D_2 \|$.

Syntax of L-terms:

L::= `NOT` L | (L_1 `AND` L_2) | (L_1 `OR` L_2) | `INV` (L).

Semantics of L-terms:

$\|$ `NOT` $L\| =$ `(Surr, Surr)` $\setminus \|C\|$,
$\|L_1$ `AND` $L_2\| = \|L_1\| \cap \|L_2\|$,
$\|L_1$ `OR` $L_2\| = \|L_1\| \cup \|L_2\|$,
$\|$ `INV` $(L)\| = \{(y,x) \mid (x,y) \in \|L\|\}$.

Syntax of P-terms:

P ::= `NOT` P | (P_1 `AND` P_2) | (P_1 `OR` P_2) |
L `SOME` C | L `EACH` C | ($L\ V$) | ($L\ c$).

Semantics of P-terms:

$\|$ `NOT` $P\|(x) \Leftrightarrow \neg\|P\|(x))$,
$\|P_1$ `AND` $P_2\|(x) \Leftrightarrow \|P_1\|(x) \wedge \|P_2\|(x)$,
$\|P_1$ `OR` $P_2\|(x) \Leftrightarrow \|P_1\|(x) \vee \|P_2\|(x)$,
$\|L$ `SOME` $C\|(x) \Leftrightarrow \exists y \in \|C\|.(x,y) \in \|L\|$,
$\|L$ `EACH` $C\|(x) \Leftrightarrow \forall y \in$ `Surr`.$(x,y) \in \|L\| \rightarrow y \in \|C\|$,
$\|(L\ V)\|(x) \Leftrightarrow (x,V) \in \|L\|$,
$\|(L\ c)\|(x) \Leftrightarrow (x,\|c\|) \in \|L\|$.

Syntax of sentences:

S ::= `NOT` S | S_1 `AND` S_2 | S_1 `OR` S_2 | S_1 `IMP` S_2 |
`EACH` $C\ P$ | `FOR-EACH` $C\ P$ | `FOR-SOME` $C\ P$ |
C_1 `ISA` C_2 | $C_1 = C_2$ | $L_1 = L_2$ | `EXIST` C |
`EXIST` L | `NULL` C | `NULL` L. | `PRECOND` $(\star\ S\ *)$ |
`POSTCOND` $(\star\ S\ *)$.

Semantics of sentences:

$\|$ `NOT` $S\| \Leftrightarrow \neg\|S\|$,
$\|S_1$ `AND` $S_2\| \Leftrightarrow \|S_1\| \wedge \|S_2\|$,
$\|S_1$ `OR` $S_2\| \Leftrightarrow \|S_1 \vee \|S_2\|$,
$\|$ `EACH` $C\ P\| \Leftrightarrow \|$ `FOR-EACH` $C\ P\| \Leftrightarrow \forall y \in \|C\|.\|P\|$,
$\|$ `SOME` $C\ P\| \Leftrightarrow \|$ `FOR-SOME` $C\ P\| \Leftrightarrow \exists y \in \|C\|.\|P\|$,
$\|C_1$ `ISA` $C_2\| \Leftrightarrow \|C_1\| \subseteq \|C_2\|$,
$\|C_1 = C_2\| \Leftrightarrow \|C_1\| = \|C_2\|$,
$\|L_1 = L_2\| \Leftrightarrow \|L_1\| = \|L_2\|$,
$\|$ `EXIST` $C\| \Leftrightarrow \|C\| \neq \emptyset$,
$\|$ `EXIST` $L\| \Leftrightarrow \|L\| \neq \emptyset$,
$\|$ `NULL` $C\| \Leftrightarrow \|C\| = \emptyset$,
$\|$ `NULL` $L\| \Leftrightarrow \|L\| = \emptyset$,
$\|$ `PRECOND` $(\star\ S\ *)\| = \|S\|$,
$\|$ `PRECOND` $(\star\ S\ *)\| = \|S\|$.

D. Specification of changes

In **LS** we can create schemas for determining dynamics of fact bases. The dynamic aspects include:

- the operations that are possible;
- the relationships between their inputs and outputs;
- the changes of fact bases that happen.

Examples 5. An examination list can be considered as an object changing its content during examination. We describe it by means the following schemes.

```
SCHEME ExamList
1. Records SUBS (Name, {20 .. 100})(*).
2. assess: Name --> {20..100}.
3. FOR-EACH X IN RecordsedNames EXIST
   Y IN {20..100} THAT assess(X) = Y.
END
```

Here `SUBS` is the abbreviation for «subset» and « –> » informs that `assess` is a function, possibly partial. The sentence 3 says that the names of students are recorded together with assessments.

```
SCHEME AddRecord
1. ExamList.
2. name?:Name.
3. assessment?: {20..100}.
4. PRECOND(*name? NOT IN Records*).
5. Records : =  Records ADD
                         (name?,assessment?)
END
```

The names with question marks are used as input variables, So, when the specific values `a` and `b` are assigned to the variables `name?` and `assessment?` then the pair (a,b) is added to `Records`. The sentence 4 act as precondition for executing 5.

```
SCHEME FindAssesment
1. name?: Name.
2. assessment!: {20..100}.
3. PRECOND (*name? IN Records*).
4. assesment! = assess(name?).
END
```

The exclamation mark informs that `assessment!` is an output variable. Thus, the scheme `FindAssesment` acts as a query.

III. Railway safety control system

A distributed railway control system (RCS) is a critical safety system. RCS consists of:

- switch boxes (SB), each one locally controlling a point, i.e. the boundary between two segments of a single track or a railway crossing,
- train control computers (TCC) residing in the train engines and collecting the local state information from switch boxes along the track in order to derive the decision whether the train may enter the next track segment [4].

We introduce a formal approach for domain specification and decision inference using BMK schemas with some constucts of Z-notation [9,10]. The presented approach allows agents to make decisions to signal trains and update the control system.

Consider the system configuration depicted in Fig. 1. The tasks of train control and interlocking are distributed on computers residing in each train $t1$, $t2$ and switch boxes $sb1$, $sb2$, each one controlling a single point, the boundary between two segments of a single track or a railway crossing.

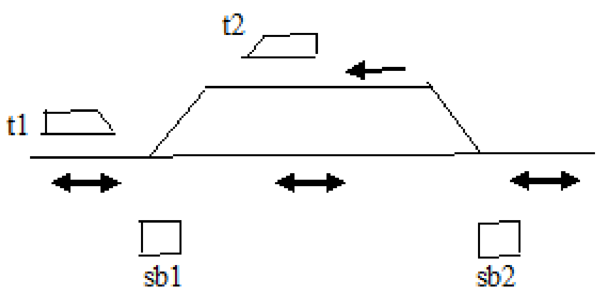


Figure 1. Example of a network with trains

The system TCC acts as follows:

- Each switch box stores the local safety-related information. In particular, this information includes the state of the traffic lights guarding the railway crossing (i.e., the track segments that are currently connected by the local point, or whether a train is approaching the switch box).
- In order to pass a railway crossing or to enter a new track, a train's TCC communicates with the relevant switch boxes to make a request for blocking a crossing, switching a point, or just reserving the relevant track segments at the SB for the train to pass.
- Depending on their local state, the switch boxes may or may not comply with the request received from a TCC. In any case, each SB returns its local state information to the requesting TCC. After having collected the response from each relevant SB, the TCC evaluates the SB states to decide whether it is safe to approach the crossing or to enter the next track segment.

We implemented a model to describe RCS. Specification schemas are used to define the domain concepts and operator schemas to enforce the train system safety. The definitions are split into the static and dynamic parts. The static part defines classes which are not changed during the system operation and the dynamic defines classes with changing states. The static part corresponds to the `TrainDefinitions` schema and the dynamic to the `TrainSystem` schema.

```
SCHEME TrainDefinitions
    RouteSegments SUBS
        (Segment,{'<--','-->','<->'})(*).
    Arcs SUBS(RouteSegments,RoutSegments).
    Network == SUBS (RoutSegments,Arcs).
    Train[Network,SwitchBlok(*),Lock(*),
    /   Direction: ,{'<--','-->'},
        Reservation(*)].
        Reservation: SwitchBlock -->
                                    SwitchBlock(*)].
END

SCHEME TrainSystem
    TrainParameters[
    speed: Train --> Speed,
    envelope: Train -->
                    (Points, Points)(*),
    segment: Train --> Segment,
    direction: Train --> {'-->', '<--'}].
    TYPE TrainSignals == {ContinueGreenZone,
    StopTrainRedZone,IncreaseSpeed,
    ContinueCrossing, StopForCrossing} .
    TYPE BlockStatus = {Free,Occupied,Clear,
    RedZoneForEveryTrain}.
    ControlRoom [
    sectorStatus: BlockStatus,
    crossingAreaStatus: BlockStatus,
    ///sectorOccupiedBy: Train(*),
    segmentsOccupied: Segment(*),
    segmentOccupiedBy: Train(*)s,
    directionOfRedZone: Direction(*)s,
    timeInSector:(Sectors,Train) <-> Time,
    signal: Train <-> TrainSignals
    ].
END.
```

The class `TrainParameters` stores current state of `TrainSystem` which includes

- speed (type Speed is defined as an alias of integers);
- train «envelopes». A train envelope is a neighborhood of the train depending on the train speed;
- segments which are currently occupied by trains train directions.

The class `ControlRoom` stores railway sector information:

- `sectorStatus` and `crossingAreaStatus` are statuses of a sector and a crossing which are inferred with train system state;
- a set of trains occupying the sector `sectorOccupiedBy`;
- a set of segments occupied by any train from `segmentsOccupied`;
- helper variable for storing the train on a segment from `segmentOccupiedBy`;
- helper variable for storing the direction of train on an unsafe segment from `directionOfRedZone`;
- time of trains in sector `timeInSector`.

The following operator schemas are similar to Z-notation schemas: the input of schemas is a set of variables which can be used by the schemas to update the state and to output other variables [10]. Operator schemas can use classes, attributes and types from other schemas using *schema importing*. To import a schema, its name and an import annotation must be explicitly written at the start of the current schema definition. An import annotation specifies whether the current schema modifies the state of an imported schema. Import annotations use symbols «=» and « <> » where « = SN » means that that current schema does not modify the schema SN and « <> SN » means that the schema SN is modified.

Like Z-notation schemas input variables in BKM schemas use «?» suffix and output variables use «!». Variable names can match any attribute name from the imported schemas. If names match and «=» import annotation was used, then both the variable and the attribute change when they are assigned to a value. If « <> » import notation was used the attribute is left unmodified.

`LinearSafety` and `CrossingSafety` schemas set the train signals to enforce safety on linear segments and at railway crossings.

```
SCHEME LinearSafety
    =TrainSystem
    <> ControlRoom,
    train?: Train,
    sector?: Sectors,
    directionOfRedZone!: Directions,
    sectorStatus!: BlockStatus
    PRECOND
    segmentSectors(segment(train?))
    sector?
    (*FOR-EACH
    other:Trains EMPTY(envelope(other)
    AND envelope(train?)sectorStatus':=
                                       Free
    signal'(train?) := ContinueGreenZone
    *)
    (*EXIST other:Trains THAT
      NOT_EMPTY(envelope(other) AND
      envelope(train?))
      timeInSector(sector?, train?) >
      timeInSector(sector?, other?)
      AND speed(train?) < speed(other)
      sectorStatus' := Occupied
      segmentOccupiedBy := other
      signal'(other) := StopTrainRedZone
      signal'(train?) := IncreaseSpeed
      directionOfRedZone! :=
      direction(other)
    *)
    (*EXIST other:Trains THAT
      NOT_EMPTY(envelope(other) AND
      envelope(train?))
      timeInSector(sector?,train?) <
      timeInSector(sector?, other?)
      AND speed(train?) -> speed(other)
      sectorStatus' := Occupied
      segmentOccupiedBy' := other
      signal'(other) := IncreaseSpeed
      signal'(train?) := StopTrainRedZone
      directionOfRedZone!:=
      / direction(other)
    *)
END
```

`LinearSafety` schema receives input variables `train?` and `sector?`. The precondition checks that the `train?` is in the `sector?`. If the precondition fails, then the schema returns a corresponding message, otherwise the schema infers train signals. If there are no other trains near the `train?`, then it is reported that the sector is free. If there is other train behind the `train?`, then the `train?` is signaled to accelerate and the other is signaled to stop. If there is other train in front of the `train?`, then the signals are produced for the corresponding trains.

```
SCHEME CrossingSafety
    =TrainSystem,
    <> ControlRoom
    crossing?: Crossing(*)
    train?: Trains
    crossingAreaStatus!:BlockStatus
    PRECOND
    segment(train?) IN
            crossingSegments(crossing?)
    (*EXIST other:Trains THAT
      NOT_EMPTY(envelope(other) AND
      envelope(train?)) AND EXIST
      envelopeSegment:
      pointSegments(envelope(train?))
      envelopeSegment==crossingSegments
```

```
    (crossing?)
    crossingAreaStatus! :=
                  RedZoneForEveryTrain
    signal'(train?) := StopForCrossing
  *)
  (*
    EXIST other:Trains THAT
    EMPTY(envelope(other) AND
    envelope(train?)) AND
    FOR-ALL envelopeSegment:
    pointSegments(envelope(train?))
    envelopeSegment ==
    crossingSegments(crossing?)
    crossingAreaStatus! := Clear
    signal'(train?) := ContinueCrossing
  *)
    *No train near the crossing*
END
```

`CrossingSafety` schema receives input variables `train?` and the state of the `crossing?`. The precondition checks that the `train?` is near the `crossing?`. If the precondition fails, then the schema returns a corresponding message, otherwise the schema infers train signals. If there is other train near the train and crossing, then the `train?` is signaled to continue, otherwise it is signaled to stop. Besides signaling the train system is updated.

IV. Conclusion

We gave a brief description of the ontological modeling system «Binary Model of Knowledge» inteded for designing and interpreting heavyweight ontologies. Such ontologies are needed when describing problem domains for advanced applications. We present a case study of modeling the railway safety control system. We can analyze the resulting ontology using the reasoning block of the system «Binary Model of Knowledge». The case study demonstrates adequacy of the system «Binary Model of Knowledge» for designing and interpreting heavyweight ontologies for complex systems.

Acknowledgment

This work was supported by Russian Foundation for Basic Research (project 14-07-0387) and Ministry of Education and Science of Kazakhstan (project 0115 RK 00532).

References

[1] T. R. Gruber. “Toward principles for the design of ontologies used for knowledge sharing.”. International Journal of Human-Computer Studies, Vol. 43, Issues 4-5, pp. 907-928, November 1995.
[2] V. Devedzic. “Undestanding ontological engineering.” Communications of the ACM, Vo.l. 45, Issue 4, pp. 136-149, 2002.
[3] T. Halpin. “Ontological modeling: Part 1”. Available in: http://www.orm.net/pdf/OntologicalModeling1.pdf
[4] A.E. Haxthausen, J. Peleska. "Formal developing and verification of a distributed railway control system.” IEEE Transaction on Software Engineering, 26 (8) , pp. 687-701.
[5] R. Mizoguchi. “Tutorial to ontological engineering”. Available in: https://ru.scribd.com/document/56477570/Mizoguchi-2004-Tutorial-on-Ontological-Engineering
[6] G.S. Plesniewicz. “Formal ontologies”. Open semantic technolologies for intelligent systems (OSTIS 2012). – Minsk, 2012, pp. 163-168. (In Russian.)
[7] G.S. Plesniewicz., D.E. Masherov, Nguen Thi Min Vu, A.B. Karabekov. “Binary Knowledge Model: specifying, instantiating and interpreting advanced ontologies.” Proceedings of the 9th International Conference on Application of Information and Communication Technologies. – Rostov-on-Don, Russia, 2015, pp.314-319. (In Russian.)
[8] Plesniewicz., D.E. Masherov, Nguen Thi Min, A.B. Karabekov. Methods and Languages for Ontological Modeling. – Almaty: IICT publ., 2015, p. 179. (In Russian.)
[9] J.M. Spivey. The Z notation: a reference manual. – Oxford (UK): Prentice Hall Int., 1989, p. 158.
[10] J. Woodcock, J/ Davis. “Uzing Z.” Available in: http://www.cs.cmu.edu/ 15819/zedbook.pdf

ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ С СИСТЕМОЙ «БИНАРНАЯ МОДЕЛЬ ЗНАНИЙ»

Г.С. Плесневич, Д.Е.Машеров

В статье дано краткое описание системы онтологического моделироваия «Бинарная Модель Знаний» (БМЗ). Используя эту систему, эксперты могут создавать «весомые» (heavyweight) онтологии для сложных приложений. Рассмотрен пример применения БМЗ к задачам разработки компьютерных систем безопасности железнодорожного движения.

Model Transformations for Intelligent Systems Engineering

Nikita O. Dorodnykh
and Alexander Yu. Yurin
Matrosov Institute for System Dynamics and Control Theory,
Siberian Branch of the Russian Academy of Sciences (IDSTU SB RAS),
Irkutsk, Russia
tualatin32@mail.ru
iskander@icc.ru

Nikolai Yu. Pavlov, Sergey A. Korshunov
and Dmitriy Yu. Sopp
CentraSib LLC.,
Irkutsk, Russia
Telephone: (3952) 603–911
info@centrasib.ru

***Abstract*—The paper discusses the application of model transformations in the process of intelligent systems engineering. The model-driven approach is used as a basic methodology and the own its implementation is proposed. The conceptual models presented in XML-like formats are used as the initial data. Particular attention is paid to the first stage in the chain of model transformations: the stage of the formation of a computation-independent model. A new domain-specific language – Transformation Model Representation Language (TMRL) and special tools are used for implementing model transformations. The description of the main TMRL constructions and an example of its application are presented.**

***Keywords*—model transformation, code generation, intelligent system, knowledge base, domain-specific language.**

I. Introduction

The model transformation is one of the main principles of the Model Driven Engineering (MDE) [1]. MDE approach is based on the use of software information models of varying degrees of abstraction as the main artifacts in the development of software systems. In this case, the process itself is a sequence (chain) of transformation of these models. In this paper we proposed to consider the application of model transformations in the process of intelligent systems engineering, in particular, for development of knowledge bases (KB) with the use of transformation of conceptual models presented in XML-like formats. The main results are analysis of related works in the field of model transformation and the new domain-specific declarative language for the description of transformations – Transformation Model Representation Language (TMRL), designed for transforming conceptual models to the KB.

II. Model Transformation

Recently, in the field of software engineering there has been a tendency to use approaches that consider models not only as artifacts of documentation (technical specifications), but also as central artifacts in software development providing automatic synthesis of program codes. This allows any to significantly reduce development time, reduce the risk of programming errors, to involve end-users in the process of software engineering. These approaches are related to the Model Driven engineering (MDE) or Model Driven Development (MDD) areas [2]. Today, the main MDE implementations (initiatives) are the following: OMG Model Driven Architecture (MDA), Eclipse Modeling Framework (EMF), Model Integrated Computing (MIC), Microsoft Software Factories, JetBrains MPS.

The central concepts of MDE/MDD are:

- A model is an abstract description of the some characteristics of the system (process) in a formal language. As a rule, the models are visualized with the aid of a certain graphic notation and serialized in XML.
- A metamodel is a model of a language used to create models (model of models).
- A meta-metamodel is a language that describes metamodels. The most common languages for metamodeling are: MOF (Meta-Object Facility), Ecore, KM3 (Kernel Meta Meta Model), etc.
- A four-layer metamodeling architecture is the concept that defines the different layers of abstraction (M0-M3), where the objects of reality are represented at the lowest level (M0), then the level of models (M1), the level of metamodels (M2) and the level of the meta-metamodel (M3).
- A model transformation is automatic generation of a target model from a source model with the accordance of a set of transformation rules [3]. In this case, each transformation rule describes the correspondence between the elements of the source and target metamodels.

A more detailed description of these concepts can be found in [2], [4].

There are many works devoted to the model transformations. At the same time, model transformations can be considered from different viewpoints.

In particular, there are two types of transformation [5]:

- Model-to-Model (M2M);
- Model-to-Text (M2T) and Text-to-Model (T2M). In this case, the output text can be in the form of a source code, documentation, specifications, and etc.

Two types of model transformations in accordance with the modelling languages used to describe the source and target models [5]:

- an endogenous transformation is a transformation between models that are using one modeling language;
- an exogenous transformation is a transformation between models that are using different modeling languages.

The model transformations can be classified according to the abstraction level on which the source and target model are resided [5]:

- a vertical transformation is a transformation of models of different abstraction levels;
- a horizontal transformation is transformation of models of the same abstraction level.

The model transformations can also be classified according to the transformation direction:

- an unidirectional transformation is a transformation where only the target model can be obtained from the source model;
- a bidirectional transformation is a transformation where the target model can be obtained from the source model and vice versa.

Currently, there are some ways to implement the model transformation:

- using graph grammars (graph rewriting) (e.g., VIsual Automated model TRAnsformations (VIATRA2) [6], Graph REwriting And Transformation (GReAT) [7], etc.);
- using visual design of transformation rules and category theory (e.g., Henshin [8]);
- using transformation standards (e.g., Query/View/Transformation [9]);
- using hybrid (declarative-imperative) approach for specifying and constructing transformation rules (e.g., ATLAS Transformation Language [10]);
- using declarative and procedural programming languages [11];
- using languages for transforming XML documents (e.g., eXtensible Stylesheet Language Transformations [12], etc.).

The transformations constructed using these ways should satisfy the following main requirements [13], [14], [15]:

- completeness: it should allows one to represent any necessary transformation in accordance with the defined models;
- formality: it should allows automatic execution;
- flexibility: it should not depends on a specific subject domain.

In this paper, it is proposed to consider the application of model transformations in the process of intelligent systems engineering and KBs based on the implementation of MDE / MDD.

III. Model Transformations for Intelligent Systems Engineering

The MDE/MDD implementation used is based on the principles of MDA [16] and assumes a clear separation of three levels (viewpoints) of the software representation:

- A computation-independent level, which is a description of the basic concepts and relationships of the subject domain, we will express it in the form of ontologies. Various conceptual models can be used to automate its formation. It is proposed to limit the number of conceptual models by XML-like formats, in particular, for representation: UML models – XMI (XML Metadata Interchange), concept maps – CXL (Concept Mapping Extensible Language), event trees – ETXL (Event Tree Mapping Extensible Language).
- A platform-independent level, which provides a representation of the domain model in the context of the formalism used to represent knowledge. In particular, logical rules or frames. It is advisable to use problem-oriented notations or UML profiles, in particular RVML (Rule Visual Modeling Language) [17].
- A platform-dependent level representing a formalized description of KBs, taking into account a certain software platform. In the context of intelligent systems engineering, such a platform is the programming languages for KBs, for example, CLIPS.

Thus, the process of intelligent systems engineering is a sequential transition between the considered levels and can be described by the following sequence of steps:

- Formation of a conceptual model by means of third-party programs. At this step some problem-oriented notations are used: UML, concept maps or memory cards, event trees. At the end of the step, the conceptual model is presented in XML-like formats: XMI (StarUML, IBM Rational Rose), CXL (IHMC CmapTools), ETXL (ET-Editor).
- Analysis of XML document of conceptual model with identification of the concepts and relations. A computation-independent model (CIM) is formed on the basis of the selected concepts, represented in the form of ontology. The automatically generated ontology is edited to clarify it.
- Formation of a platform-independent model (PIM) based on ontology. This model depends on the formalism of knowledge representation, but does not take into account the features of languages and the tools for implementing these formalisms.
- Formation of a platform-specific model (PSM), taking into account the features of languages and means for implementing formalisms.
- Generation of program codes or specifications of KBs and intelligent systems based on generated models.

It is suggested to use model transformations to implement the transitions between the steps and to consider in detail the transformations of the first step, which are related to the M2M type.

In the context of the MDA approach, it is recommended to use the OMG standard for implementation of M2M transformations, which is called QVT including Operational, Relational and Core model transformation languages.

However, programmers who use the QVT should:

- know the specific syntax of the model transformation language;
- be able to describe transformation rules with the aid of the model transformation language;
- know the additional languages, such as an Object Constraint Language (OCL) that can be used to construct the transformation rules;
- be able to describe metamodels for the source and target languages (to support the transformation process).

It should be noted that all model transformation languages is supported by a specific software tools, that, in turn don't provide an opportunity to visualize the development process, i.e. the transformation rules are defined in special text editors focused on programmers. The combination of these factors makes difficult to use these languages and tools in a practical way by end-users (e.g., subject domain experts, knowledge engineers, analysts, etc.), in particular, when developing KBs and ESs on the basis of conceptual models. So, in practice developers prefer "ad-hoc" solutions for particular tasks with using declarative and procedural programming languages.

We propose a new domain-specific language – TMRL to address these shortcomings.

IV. Transformation Model Representation Language

TMRL is focused on the representation and storage of the so-called a transformation model, which is a scenario (program) for transforming the source conceptual model to the target KB. Thus, TMRL transformation model is the core of the software component for transformation, providing analysis (parsing) of conceptual models and synthesis (generation) of the KB code [18].

The structure of the transformation model can be defined as follows:

$$M_T = \left\langle \; MM_{IN}, MM_{OUT}, T \; \right\rangle, \tag{1}$$

where MM_{IN} is a metamodel of the source (input) conceptual model; MM_{OUT} is a metamodel of the target (output) model of knowledge representation (KB); T is an operator for model transformation (a set of rules).

The TMRL grammar belongs to the class of context-free grammars (LL(1) CF-grammars) [19]. The TMRL constructs allow one to describe the elements of the transformation model in a declarative form, in particular, the rules for the correspondence of metamodel elements, as well as the mechanism of interaction with previously developed (external) software components for the transformations. TMRL specifications meet the requirements of accuracy, clarity and completeness [20], i.e. the TMRL specifications contain all the necessary information (for the considered transformations) to solve the task, all objects of the model are well formalized, while the specifications are compact enough and at the same time understandable (readable).

The main difference between TMRL and existing model transformation languages is its ease of use, achieved through a limited set of elements. TMRL is not an extension of other languages and does not use the constructions of other languages, as other model transformation languages very often do, in particular, ATL uses the OCL. In addition, TMRL has human-readable syntax for the purpose of making the necessary modifications to the model of the transformation manually, if necessary. An additional feature of TMRL is the ability to describe interaction with previously developed software components of the transformation in support of the import of different formats of conceptual models.

A. Transformation model structure in TMRL

The transformation model in TMRL consists of three main blocks. Consider this structure in an example that describes the transformation of a UML class diagram into an ontology model (CIM). In this case, the language elements are highlighted in bold.

Block 1. Description of the elements and relationships of the source metamodel:

```
Source Meta-Model UML-diagram-class {
   Elements [
      Model,
      Class attributes (xmi.id, name),
      ...
   ]
   Relationships [
      Model is associated with Namespace.
          ownedElement,
      Namespace.ownedElement is associated
          with Class,
      DataType(xmi.id) is Attribute(type),
      ...
   ]
}
```

This block contains a description of the UML class diagram: "UML-diagram-class", including elements of the model ("Elements"). In this example, these are the "Model" and "Class" elements, the "Class" element has "xmi.id" and "name" properties. In addition to describing the elements, the source metamodel contains a description of the relationships between the elements of the metamodel ("Relationships"), including by identifiers, for example, linking an attribute to a data type ("DataType(xmi.id) is Attribute(type)").

Block 2. Description of elements and relationships of the target metamodel:

```
Target Meta-Model Ontology {
   Elements [
      Ontology attributes (about),
      Class attributes (id, label),
      ...
   ]
```

```
    Relationships [
        Ontology is associated with Class,
        ...
    ]
}
```

The block contains a description of the ontology model: "Ontology". The structure of the block is similar to the structure of the block of the source metamodel.

Block 3. Description of rules for transforming models:

```
Transformation UML-diagram-class to Ontology {
    Rule Model to Ontology priority 1 [
        Ontology(about) is Model or ModelElement
            .name
    ]
    Rule (Class, ModelElement.name) to Class
        priority 2 [
        Class(label) is Class(name) or
            ModelElement.name
        Class(id) is Class(name)
    ]
    ...
}
```

This block describes the rules for converting the elements of the source ("UML-diagram-class") metamodel to the target ("Ontology") metamodel.

The Knowledge Base Development System (KBDS) is used to support TMRL [21]. KBDS provides interactive visual construction of transformation models and their automatic generation on TMRL.

V. Conclusion

The paper describes the application of model transformations in the process of intelligent systems engineering and, in particular, KBs. The various conceptual models presented in XML-like formats are used as initial data. Particular attention is paid to the stage of formation of CIM. The detailed description of the new domain-specific language (TMRL) is provided. The KBDS is used as tool for supporting TMRL.

References

[1] L. G. Cretu and D. Florin, *Model-Driven Engineering of Information Systems: Principles, Techniques, and Practice*. Apple Academic Press, 2014.

[2] A. R. Da Silva, *Model-driven engineering: A survey supported by the unified conceptual model*. Computer Languages, Systems & Structures. 2015. Vol. 43. P. 139-155.

[3] A. Kleppe, J. Warmer and W. Bast, *MDA Explained: The Model-Driven Architecture: Practice and Promise*, 1rd ed. New York: Addison-Wesley Professional, 2003.

[4] J. Sprinkle, B. Rumpe, H. Vangheluwe and G. Karsai, *3 Metamodelling. State of the Art and Research Challenges*. Model-Based Engineering of Embedded Real-Time Systems. 2010. P. 57-76.

[5] T. Mens and P. V. Gorp, *A Taxonomy of Model Transformations*. Electronic Notes in Theoretical Computer Science. 2006. Vol. 152. P. 125-142.

[6] D. Varro and A. Balogh, *The model transformation language of the VIATRA2 framework*. Science of Computer Programming. 2007. Vol. 63, No. 3. P. 214-234.

[7] D. Balasubramanian, A. Narayanan, C. Buskirk and G. Karsai, *The graph rewriting and transformation language: GReAT*. Electronic Communications of the EASST. 2006. Vol. 1. P. 1-8.

[8] T. Arendt, E. Biermann, S. Jurack, C. Krause and G. Taentzer, *Henshin: advanced concepts and tools for in-place EMF model transformations*. Processing of the 4th International Conference on Model-Driven Engineering and Software Development (MODELSWARD 2010) / Lecture Notes in Computer Science, Springer Berlin Heidelberg. 2010. Vol. 6394. P. 121-135.

[9] *Query/View/Transformation (QVT)*. 2016. Available at: http://www.omg.org/spec/QVT/ (accessed 07.12.2017).

[10] F. Jouault, F. Allilaire, J. Bézivin and I. Kurtev, *ATL: A model transformation tool*. Science of Computer Programming. 2008. Vol. 72, No. 1. P. 31-39.

[11] A. F. Berman, O. A. Nikolaychuk and A. Y. Yurin, *Intelligent planner for control of failures analysis of unique mechanical systems*. Expert Systems with Applications. 2010. Vol. 37, No. 10. P. 7101-7107.

[12] *XSL Transformations (XSLT) Version 2.0*. 2007. Available at: http://www.w3.org/TR/xslt20/ (accessed 07.12.2017).

[13] T. Gardner, C. Griffin, C. Koehler and R. Hauser, *A review of OMG MOF 2.0 Query/Views/Transformations Submissions and Recommendations toward the final standard*. Object Management Group, OMG Document ad/03-08-02, 2003.

[14] S. Sendall and W. Kozaczynski, *Model Transformation – The Heart and Soul of Model-Driven Software Development*. IEEE Software. 2003. Vol. 20, No. 5. P. 42-45.

[15] K. Czarnecki and S. Helsen, *Feature-based survey of model transformation approaches*. IBM Systems Journal. 2006. Vol. 45, No. 3. P. 621-645.

[16] D. Frankel, *Model Driven Architecture: Applying MDA to Enterprise Computing*. New York: Wiley, 2003.

[17] A. Y. Yurin, *Notation for design of knowledge bases of rule-based expert systems*. Object systems. 2016. No. 12. P. 48-54. (In Russ.)

[18] I. V. Bychkov, N. O. Dorodnykh and A. Y. Yurin, *Approach to the development of software components for generation of knowledge bases based on conceptual models*. Computational Technologies. 2016. Vol. 21, No. 4. P. 16-36. (In Russ.)

[19] A. V. Aho, M. S. Lam and J. D. Ullman, *Compilers: Principles, Techniques, and Tools*, 2nd ed. Addison Wesley, 2006.

[20] V. N. Agafonov, *Specification of programs: conceptual means and their organization*. Novosibirsk: Nauka, 1987. (In Russ.)

[21] N. O. Dorodnykh, *Web-based software for automating development of knowledge bases on the basis of transformation of conceptual models*. Open Semantic Technologies for Intelligent Systems. 2017. P. 145–150.

ПРИМЕНЕНИЕ МОДЕЛЬНЫХ ТРАНСФОРМАЦИЙ ДЛЯ СОЗДАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Дородных Н.О.*, Коршунов С.А.**, Павлов Н.Ю.**, Сопп Д.Ю.**, Юрин А.Ю.*

*Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (ИДСТУ СО РАН)

**ООО "ЦентраСиб"

В работе рассмотрено применение модельных трансформаций в процессе создания интеллектуальных систем на основе предлагаемой реализации модельно-управляемого подхода. В качестве исходных данных предлагается использовать концептуальные модели, представленные в XML-подобных форматах. Особое внимание уделено первому этапу в цепочке модельных трансформаций: этапу формирования вычислительно-независимой модели. В качестве инструментария реализации модельных трансформаций предложено использовать новый предметно-ориентированных язык – Transformation Model Representation Language (TMRL) и поддерживающее его программное средство. Приведено описание основных конструкций TMRL и пример его применения.

From training intelligent systems to training their development tools

Golenkov V. V., Gulyakina N. A.,
Grakova N. V., Davydenko I. T., Nikulenka V. Y.
The department of Intelligent Information Technology
Belarusian state University of Informatics and Radioelectronics
Minsk, Belarus
{golen, guliakina, grakova}@bsuir.by
ir.davydenko@gmail.com, v.nikulenka@jetsys.io

Eremeev A. P.
The department of applied mathematics
National research University "MEI"
Moscow, Russia
eremeev@appmat.ru

Tarasov V. B.
The department of computer systems of production automation
Moscow state technical University N. E. Bauman
Moscow, Russia
vbulbov@yahoo.com

***Abstract*—The article considers the approach to the transition implementation from the training of intellectual systems based on knowledge to the training of tools for their development. At the same time, the architecture of such an intelligent system is considered as the basis for ensuring its flexibility and learnability. In addition, the article examines the areas of learning and self-learning of intelligent systems, as well as their ability to acquire knowledge and skills from various sources. The justification of OSTIS technology using for the development of knowledge-based intelligent systems is provided.**

***Keywords*—learnability of intellegent systems; skill; intellegent system; problem; external environment; self-training; reflexivity; unlimited learnability; flexibility; machine learning; training of intellegent systems; artificial neural network; genetic algorithm; technology of development of intellegent systems; ostis-system; self-training of intellegent systems; metaknowledge**

I. Introduction

Here are the main points of this article:

1) ***Learnability of intelligent system***, that is, its ability to acquire new ones and to improve already acquired ***knowledge*** and ***skills*** is the main characteristic of intelligent systems. *Learnability of intelligent system* creates the necessary conditions to ensure the rapid pace of ***intelligent systems*** evolution, to intensively expand the set of ***problems*** which it solves, to continuously improve the quality of the *problems* solution, to quickly adapt *intelligent systems* to changes in ***the external environment*** and operating conditions.
2) *The level of learnability of intelligent system* is determined by the level of development of the ***self-learning*** means and, first of all, by the level of its ***reflexivity***, that is, the ability for self-analysis.
3) The highest form of *learnability of intelligent systems* is their ***unlimited (universal) learnability***, assuming the ability to acquire and improve <u>any</u> (!) types of knowledge and skills.
4) *Learnability of intelligent system* is determined by the degree of its ***flexibility***, that is, the complexity of making various changes to the *intelligent system*, as well as the variety of types of possible changes.
5) Modern work in the field of ***machine learning*** does not consider the whole complex of problems associated with the ***training of intelligent systems***, but focuses mainly on improving various *skills*. It also applies to the training of ***artificial neural networks*** and to the training of ***genetic algorithms***.
6) The most important criterion for the quality of the proposed ***technologies for development of intelligent systems*** is *the level of learnability* of developed *intelligent systems* which it provides.
7) *Intelligent systems* developed using ***OSTIS Technology*** [1], which we call ***ostis-systems***, have a high level of *flexibility* and *unlimited learnability*.
8) ***Flexibility of ostis-systems*** is determined by:
    - the sense nature of the ***internal knowledge representation***;
    - the universality of the language of the internal ***knowledge representation***;
    - the developed level of ***the associative organization of memory*** of *ostis-systems*;
    - ***agent-oriented management of knowledge processing***, managed by ***knowledge base***.
9) ***Learnability of ostis-systems*** is determined by the ability of *ostis-system* to detect ***contradictions*** (errors), ***information holes*** and ***information waste*** that appear in the current state of *knowledge base* both as a result of acquiring *knowledge* and *skills* from outside and in the

process of solving various *problems*. In the second case, each detected *contradiction*, *information hole* or *information waste* is clearly associated with the ***information process*** that generated it to clarify the reason for its occurrence.

10) ***Unlimited learnability of ostis-systems*** is determined by:
    - universality of ***internal sense language of knowledge representation***;
    - unlimited opportunities for transition in ***internal representation of knowledge bases*** from *knowledge* to ***metaknowledge*** and, as a result, unlimited possibilities for *structuring knowledge bases*;
    - universal character of *agent-oriented model of knowledge processing*.
11) *OSTIS Technology* also has high level of *flexibility* and *unlimited learnability*, because it is implemented in the form of ***intelligent metasystem***, which is also *ostis-system.*

II. The features of intelligent systems

Transition of ***traditional computer systems*** to *intelligent systems*, that is ***to the systems based on knowledge*** is the most important tendency of transition to the next generation of the computer systems having ***semantic compatibility***, *learnability* and ability to formation of temporary collectives of computer systems in need of the collective solution of difficult tasks. At the same time it is essential to emphasize that any modern computer system can be realized in the form of functionally equivalent *systems based on knowledge*. Such transformation of modern computer systems will allow to increase their quality and competitiveness significantly. As a result of transformation of modern computer systems in functionally equivalent *systems based on knowledge* we will receive ***semantic compatible systems based on knowledge*** with *the high level of learnability*, differing among themselves in structural complexity of *knowledge bases* and functional complexity of means of processing of knowledge.

It is necessary to distinguish *traditional computer systems* using some methods of artificial intelligence (for example, artificial neural networks, genetic algorithms, nonclassical logics) from *intelligent systems*, which consists of (1) the base of all necessary *knowledge* and *skills* (2) ***interface with external environment*** and (3) ***integrated problem solver*** which consists of *skill* base acquired by intelligent system (methods for solving various classes of problems) and ***interpreter of knowledge and skills*** acquired by *intelligent system*. General architecture of *intelligent system* is shown in (Figure 1), which is explained below in formal language ***SCn-code*** [1] (Listing 1).

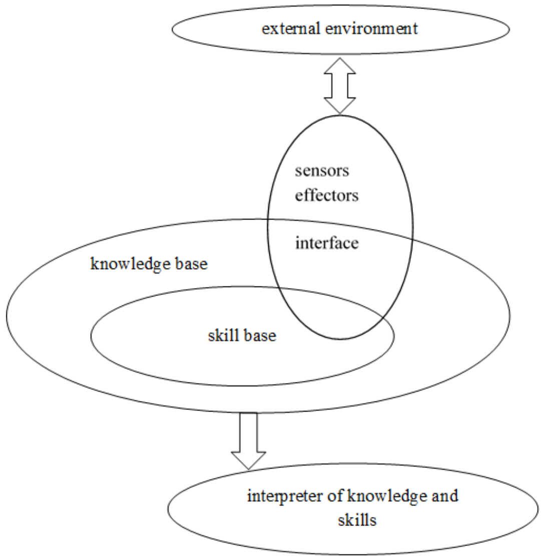


Figure 1. The architecture of intelligent system

intellectual system
= *trained system based on knowledge*

knowledge base
= *knowledge base and skills of intelligent system*
= *systematized mix of all knowledge and skills stored in memory of intelligent system*
<= *second domain**
   *knowledge base**
= *Class of information structures, each of which is knowledge base of some intelligent system*

knowledge base*
= *Binary oriented relation connecting each intelligent system with its knowledge base*

skill base
= *skill base of intelligent system*
<= *generalized inclusion**
   knowledge base
= *a systematized complex of all methods of solving various problems and classes of problems known to intelligent system*
= *part of knowledge base, which is systematized mix of all skills that intelligent system has acquired to the present moment, each of which is knowledge about how to solve specific problem or some class of problems*

memory of intelligent system
= *internal environment of intelligent system in which its knowledge base is stored and processed and which is contrasted to external environment of this intelligent system*

external environment of intelligent system
= *mix of all kinds of entities which are not signs or sign constructions stored in memory of intelligent system and*

which are in relation to the indicated internal signs, can only implement the role of their denotations (that is, the role of described entities)

interface of intelligent system
= *interface of intelligent system with its external environment*

sensors of intelligent system
= *sensory means of intelligent system that ensure reflection of events and situations occurring in external environment of intelligent system (including reflection of intelligent system behavior in its external environment), events and situations occurring in memory of this intelligent system*

effectors of intelligent system
= *effector means of intelligent system, providing transformation of planned actions specifications of intelligent system in external environment*

interpreter of knowledge and skills
= *interpreter of knowledge base and skill base of intelligent system*
= *base machine of intelligent system*
= *processor of intelligent system*
= *engine of intelligent system*
= *set of means to solve problems in intelligent systems*
= *set of means by means of which the intelligent system carries out the solution of tasks*

integrated problem solver
= *integrated problem solver of intelligent system*
<= *generalized decomposition**
{
- *skill base*
- *interpreter of knowledge and skills*
}

Listing 1. Explanation of general architecture of intelligent system in formal language SCn-code

The architecture of *intelligent system* is the basis for ensuring its *flexibility* and *learnability* due to its clear stratification to independent components, expansion and improvement of each of them sufficiently is not dependent on changes in other components.

Traditional computer systems are mainly focused on solving problems, the initial data of which are localized (fully specified). To solve the problems there is no need for additional information. The specialized solvers can exist for these problems which can use convenient specialized methods of representing and storing the processing data. In order to integrate these specialized solvers into the developing computer system it is sufficient to organize the exchange of a finite number of initial (transmitted) and returned parameters with a priori known semantic interpretation or the exchange of finite information constructions presented in a specified format.

But for *problems* solved by *intelligent systems*, in general, it is not known what data and knowledge should be used to solve these problems, what methods should be used to solve them. When you solve these problems, it is necessary to localize a fragment of the *knowledge base* which comprises the data and the *knowledge* which are sufficient for the solution of the *problem* and excluding those data and *knowledge* that are obviously not needed for this, as well as the distinguishing of the available solutions from the variety of methods that are sufficient for solving this *problem*.

The competitiveness of *intelligent system* ensures:
- permanent *training of intelligent system* during the operation;
- permanent ***self-learning of intelligent system*** during the operation;
- *learnability of intelligent system*;
- ***flexibility of intelligent system***.

The circumstances that constrain the development of *intelligent systems* market:
- inter-subject character of the development of *intelligent systems*;
- high science-intensive development of *intelligent systems*;
- the need for close cooperation with experts of applied fields;
- lack of ***general theory of intelligent systems***;
- lack of ***common technology for the development of intelligent systems*** which is available to a wide range of engineers;
- high and not always adequate social interest in artificial intelligence;
- substantially more complicated training of engineers of intelligent systems in comparison with the training of traditional computer systems engineers.

The main features of the next generation intelligent systems:
- the variety of problems types solved by the developed intelligent systems:
  - solving problems that require the joint coordinated use of problem-solving models variety, types of knowledge and knowledge representation models;
  - solving problems which a priori aren't known how to solve and the necessary knowledge and also models (principles) for solving problems aren't known;
- low deadlines and relatively low labor content for the development of the initial versions of *intelligent systems*, which are *trained* and do *self-learning* during the operation;
- high rates, low labor content and unlimited possibilities for the evolution (improvement) of *intelligent systems* during the operation with the maintaining the compatibility of these systems;
- high rates of evolution, low labor content and unlimited possibilities for the evolution of the *technology for the development of intelligent systems* with the automation of making appropriate changes to the operating *intelligent systems* while ensuring the compatibility of these systems.

III. Training of intelligent systems

Under ***the training of intelligent systems*** we don't mean only the training, for example, of *artificial neural networks* that are part of *intelligent systems*. We interpret *the intelligent system training* as a process of new knowledge and skills acquiring and improving for this system, implemented by some team of ***teacher-developers*** of the specified system.

Training of intelligent system has the following stages:

- preliminary debugging (testing and tuning) of *intelligent system* to solve some problems that is implemented before the introduction of *intelligent system* in operation. It is clear that the stage of creating (synthesizing, assembling) the initial (starting) version of this system should precede the debugging of *intelligent system*;
- *training of intelligent system* is implemented directly during its operation on the basis of a permanent analysis of the trained *intelligent system* activity.

Training of intelligent system is implemented:

- for improvement of its *knowledge base*;
- for improvement of its *integrated problem solver*;
- for improvement of its verbal and, primarily, ***user interface*** with various external subjects;
- for improvement of its sensory-effector subsystems, providing interaction with the *external environment*.

IV. Self-learning of intelligent system

Self-learning of intelligent system is a transition from a passive form of *training intelligent system*, when the trained system is treated as a passive "container" filled by *teachers-developers* with new *knowledge* and *skills*, to an active form of training, when the trained system becomes the participant of the training process. *Self-learning* of *intelligent system* is the automation of various processes aimed at training of *intelligent system* and implemented by *intelligent system.* These automated processes include:

1) permanent ***analysis of knowledge base quality*** and *integrated problem solver* of trained *intelligent system*, which has led to, for example, the identification and specification of various kinds of errors (contradictions), *information waste*, *information holes* (lack of *knowledge* and *skills*), assessment of the reliability of new acquired (including input) *knowledge* and *skills*, as well as the specification of the *knowledge* and the *skills* (who is the author, the moment of appearance in the system, the type and other);
2) Some automated types of improving the current state of *knowledge base* and *integrated problem solver* - automatic correction of some errors, the *information waste* removal, the specification of *information holes*, the systematization of acquired *knowledge* and *skills*, the extraction of implicit *knowledge* from the given (inductive inference, *self-learning* by using of precedents);
3) Coordination of the activities of *intelligent system developers (teachers)*. It means that *intelligent system* developers exchange information among themselves only through the *knowledge base* of the developed *intelligent system*. The developers of *intelligent system* become independent *agents* (subjects) of the *intelligent system* development (training), managed by the *knowledge base* of this system. The developed (trained) *intelligent system* also becomes one from these *agents*.

The development of *intelligent system* that has developed *self-learning skills* is fundamentally different from development of *intelligent system* that does not have a high level of *self-learning*. This is due to the fact that the *self-learning intelligent system* becomes one of the subjects of its own training, that is, one of its teachers. And it becomes an essential factor in increasing the effectiveness of training, because no one knows better than the trained *intelligent system* about its *knowledge base*, the *integrated problem solver*.

So ***the method of training the intelligent system*** is largely determined by the tools of *self-learning* that the trained *intelligent system* has.

Areas for *self-learning* of *intelligent system* are:

- acquisition of new *knowledge* from different sources;
- extracting implicit *knowledge* from acquired *knowledge*:
  - •• detection of regularities;
  - •• structuring *knowledge base*;
- maintaining integrity of *knowledge base* (consistency, completeness, flexibility);
- increasing efficiency of solving problems based on analysis of own activities.

V. Learnability of intelligent system

Learnability of intelligent system is a level of ability to solve *self-learning* problems, for example:

- integrate new *knowledge* and *skills* into *knowledge base*;
- identify and resolve *contradictions (*errors) in *knowledge base*;
- structure *knowledge* and *skills* for localization of solution areas of problem sections;
- identify *information holes* in *knowledge base*;
- identify and remove *information waste*;
- analyze the quality of its own activity in solving various problems and learn from its own mistakes.

It is clear that learnability of *intelligent system* is based on ***reflection***, that is, the ability to analyze and evaluate its own quality and quality of its activities, as well as the ability to learn new *knowledge* and *skills* and improve the acquired *knowledge* and *skills* quickly. For example, the learnability of artificial neural networks is determined by the availability of a method for automatically adjusting of *the neural network* which is based on the results of its testing, aimed at constructing of optimal connections' structure, setting up parameters of connections (the stochastic gradient method, the method of back propagation of the error).

Thus, *learnability* of *intelligent system* is ensured by:

1) Systematization of internal representation of *knowledge* and *skills* – all accumulated *knowledge* and *skills* should be brought into coherent system;
2) Fairly simple model for integration (immersion) of new *knowledge* and *skills* into *knowledge base*;

3) Unlimited possibilities to represent in *knowledge base* all necessary information for self-analysis, containing a sign of own ***Myself***, complete own documentation, description of its connections with other *entities* including description of its own point of view from the point of view of other subjects;
4) Ability to *reflection* and a fairly simple model for analyzing quality of the current state of *knowledge base* (quality of *knowledge base* structure, ***completeness of knowledge base***, presence and localization of detected contradictions and errors);
5) Level of development of means of detecting and eliminating unnominal (including erroneous) situations in the process of functioning of *intelligent system*;
6) Level of development of means to improve quality of the current state of *knowledge base* (improving system of accumulated *knowledge* and *skills*).

VI. Unlimited learnability of intelligent systems

The highest level of learnability of *intelligent systems* is ***unlimited learnability***, striving for permanent and unlimited reduction in those problems which *intelligent system* has not managed with or managed not well enough. Unlimited (universal) learnability of *intelligent system* is its ability to acquire (grasp) quickly and without any restrictions and apply new necessary *knowledge* and *skills* effectively, and also ability to modernize and improve the acquired *knowledge* and *skills* without any restrictions.

When unlimited learnability of *intelligent system* implies absence of any restrictions on the types of acquired *knowledge* and *skills*, it requires ensuring compatibility within *knowledge base* of all kinds of *knowledge* and *skills*, as well as all possible models for their presentation, and leads to unlimited expansion of set of problems solved by *intelligent system*.

unlimited learnability
= *universal learnability*
= *ability to acquire, extract and systematize a variety of knowledge and skills, as well as to integrate and joint use them in solving complex problems*

Listing 2. Formal interpretation of the concept of unlimited learnability

Unlimited learnability of intelligent system is provided by:

- variety of used *knowledge* and *skills*' types;
- absence of restrictions on the acquisition and deep integration of fundamentally new (previously unknown) types of *knowledge* and *skills* into *knowledge base*;
- compatibility of used *knowledge* and *skills*.

VII. Flexibility of intelligent systems

Flexibility of intelligent system is the basis of its *learnability*. *Flexibility of intelligent system* is determined by the labor content (simplicity) of making various changes in *intelligent system*, implemented at various levels of *intelligent system* during its *training*.

flexibility
= *level (degree) of flexibility*
= *modifiability*
= *ease of modification*
= *reconfiguration*
= *softness*

Listing 3. Formal interpretation of the concept of flexibility

The principal difference between the characteristic of *flexibility of intelligent systems* and the characteristic of their *learnability* lies in the following: *flexibility* determines the possibility and complexity of various transformations in the *intelligent system*. In contrast, *learnability* determines a fundamental possibility and complexity of holistic transformations of a higher level that translate the *knowledge* and *skills* of *intelligent system* to a qualitatively new content level. So, for example, for the training of *intelligent system*, it is important to bring to its *knowledge base* a new regularity, not previously known to it. Whereas flexibility of *intelligent system* should ensure a fairly simple integration of the regularity's formulation into the composition of the *knowledge base*. Thus flexibility is a kind of "syntactical" aspect of learnability, which is the basis that provides this learnability. Therefore, there may be ***flexible systems***, but ***they don't have the learnability***.

Flexibility of *intelligent system* is determined not only by the labor content of making any changes in the *intelligent system*, but also by the labor content of ensuring *the integrity* of *intelligent system* when the changes are made. For example, new added information to the *knowledge base* of *intelligent system* may conflict with the *knowledge*, which has already located in the *knowledge base*. It means that it is necessary for the maintaining of the integrity (in this case, consistency) of the *knowledge base*:

1) to try finding in the current state of the *knowledge base* all statements that, for various reasons, contradict new information (for example, there may be statements about non-existence, statements about uniqueness);
2) if no such contradictions are found, the new information can be integrated into the current state of the *knowledge base*;
3) if such contradictions are found, then it is necessary to determine for each of them which of the contradictory statements has greater reliability, then either to eliminate (destroy) the contradiction or to fix the fact of its presence with an indication of that the *intelligent system* itself agrees with and that is considered by some other explicitly indicated external subject which is the author of the information entered into the *knowledge base*.

It is much more difficult to ensure the integrity of *intelligent system*, not with the addition of something new, but with adjustment (updating, changing) any fragments of *intelligent system*. Firstly, each change can also come into conflict with the current state of *intelligent system* and the contradictions should be solved. Secondly, for the ensuring of the *intelligent system* integrity some of the changes (for example, the replacement of concepts, the change in the structure of the *knowledge base*) require the implementation of not only these changes, but also a large number of other changes that are the consequences of the former. It is clear that not only the automation of making various changes to *intelligent system* should be provided for ensuring the flexibility of *intelligent system* updating, but also the possibility of rolling back to the previous states of the system (in case of erroneous changes), which requires storing the history of its evolution in *knowledge base* of *intelligent system* (its training).

Flexibility of *intelligent systems* is provided by the ***unification*** of the basic principles of information coding in the *intelligent systems*, memory, the basic principles of *intelligent systems* memory organization, as well as the basic principles of processing information in the memory of *intelligent systems*. As the training of *intelligent systems* comes down to the expansion and improvement of the system of acquired *knowledge* and *skills*, this unification significantly reduces the variety of forms of implementation of processes aimed at *intelligent systems*' training. It is of particular importance for the realization of the unlimited learnability of *intelligent systems*.

VIII. Flexibility of ostis-systems

Flexibility of ostis-systems (systems developed using *OSTIS Technology*) is provided by:

- the basic principles of coding information in the memory of *ostis-systems* are proposed by *OSTIS Technology*, the principles underlie the ***SC-code*** (Semantic Computer Code);
- the basic principles of organizing *ostis-systems* memory which are proposed by *OSTIS Technology* - the principles of *sc-memory* organization, which provides storage and processing of ***SC-code texts***;
- the basic principles of processing information in *OSTIS Technology* in memory of *ostis-systems* (in *sc-memory*).

A. SC-code Principles

Among the basic principles of coding information in memory of *ostis-systems* that provide flexibility of *ostis-systems* are, first of all, those principles that provide ***sense information representation*** with the following characteristics:

1) Among ***the signs*** from ***ostis-system knowledge base*** there should not be pairs of synonymous signs that is signs with the same ***denotation***, ***signs denoting*** the same ***entity***.
2) There should be no duplication of information not only in the form of multiple occurrence of signs denoting the same *entities*, but also in the form of multiple occurrence of ***semantically equivalent texts*** (*knowledge*). In this case, ***the logical equivalence of texts*** is allowed.
3) Among the *signs* from *knowledge base* there should not be *any homonymous signs* that is signs, which can *denote* different *entities* in different contexts.
4) All the *signs* from *knowledge base* should not have an internal structure, which analysis is necessary for understanding ***the sign constructions*** using these *signs*. The internal structure of *signs* is necessary only in those *sign constructions* in which the multiple occurrence of signs having the same *denotation* is allowed, and the analysis of the coincidence of the internal structures of such signs is means of establishing their synonymy or supposed synonymy specified at the stage of subsequent semantic analysis of the *sign construction*. Consequently, all the signs included into *sense information representation* can be *considered* abstract, because they abstract from their internal structure, they have only one characteristic - to designate one-to-one correspondent *denotations*, are invariants of the corresponding maximal ***classes of synonymic signs***.
5) The composition of *the sign construction*, which is *sense information representation*, should not include anything but the abstract *signs* considered above.
6) Within the constructing of *sense information representation*, such constructions as words, terms (phrases) should not be only used, but such language methods as declination, conjugation, punctuation marks (separators, delimiters).
7) From the syntactic point of view, all *the signs* included into *sense information representation* should be clearly divided into two types:
   - ***signs of connections*** between described *entities* (connection is considered one of the types of described *entities*);
   - *signs of entities* that are not connections.

The important advantage of *sense information representation* is that it explicitly and clearly defines the connections between the described *entities* (including the *connections* between the *connections*) in the form of *connections between the signs(!)* of these *entities*, the semantic type of each *connection* is clearly indicated. Stress that all the *connections* of each described entity in *sense information representation* are represented only by a set of *connections' signs* incident to the sign of the indicated *entity,* i.e. a set of *connections*, where the *sign* of the described entity is one of the components. Thus, any *sign construction* can be represented (in a *semantically equivalent* form) as a set of *signs* of the described *entities* and a set of *connection's signs* connecting these described *entities* with other *entities*. At the same time, there are no restrictions on the described *entities* and on the connections between them. The described *entities* can be:

- ***material*** (physical) and ***abstract*** (virtual) - ***numbers***, ***sets***, signs of any *entities*;
- ***really existing*** and ***fictitious***;
- ***fixed (constant)*** and ***arbitrary (variables);***
- *connections* between *entities*.

It is not difficult to see that the *sign construction*, which is the *sense information representation*, can't be ***linear*** in general, because each described *entity*, which is a *denotation* of the corresponding *sign,* can be connected by an unlimited number of connections with other *entities* described in the same *sign construction*. Thus, *sense information representation* is a ***graph structure***, which has a different theoretical-graph configuration. Such *graph structures*, which have the semantic characteristics described above, are also called ***semantic networks*** [2], [3], [4]. The interpretation of *semantic networks* described above makes the possibility of fully using the theory of graphs for studying the syntactic characteristics of ***semantic networks*** and for constructing algorithms for processing *semantic networks*.

It is necessary to clearly distinguish the *semantic network*, which is internal abstract information representation in *the memory* of *intelligent system*, from its coding within the chosen version of the technical implementation of the specified memory, as well as from various versions of its visualizing for users. The semantic network and its graphic or text image are not the same. Similarly, it is necessary to clearly distinguish the *graph structure* as an abstract object from various variants of its representation.

Within *OSTIS Technology*, we speak about creating formal means of describing the meaning of various types of *knowledge* and formal means of describing *knowledge* processing at sense level. This involves the development of the corresponding standard that distinguishes the basic universal language of *semantic networks* from the whole variety of *semantic networks abstract languages*, which is called ***SC-code*** [5]. *SC-code* is a language of unified semantic networks. The texts of this language are called ***sc-texts***. The *signs* in the composition of *sc-texts* are called ***sc-elements***. The transition from the general concept of the semantic network introduced above to unified *semantic networks* (*sc-texts*) is considered as setting a number of restrictions on *semantic networks* of a general form, but such restrictions that do not reduce the semantic power of the *semantic networks language claiming universality* [6].

The consequences of the considered basic principles of building ***knowledge bases* of *ostis-systems*** are:

- unlimited possibility of transition within each *knowledge base* from *knowledge* to ***meta-knowledge***, from *meta-knowledge* to ***meta-meta-knowledge*** and so on, and, consequently, an unlimited possibility of ***structuring knowledge bases*** on a variety of features. This structuring is, first of all, necessary for localizing ***the area of the each problem solution***, that is, for establishing connections between different problems and areas of their solution (those fragments of *knowledge bases* that contain sufficient information for solving the corresponding ***problems***);
- the ability to represent (encode) any (!) ***structure*** in the *sc-memory* in the form of isomorphic *sc-text* whether:
  - ***image structure*** (the structure of the connections between different image fragments);
  - ***the syntactic structure*** of a naturally language text (the structure of connections between different fragments of text at the level of letters, words, terms);
  - *the structure* that reflects the configuration of various connections between any (!) described *entities*.
- complete independence of *sc-texts* from ***terms*** (names) attributed to various described *entities*. These terms are also considered as described *entities*, and the connections connecting these terms to the *entities* which they are assigned to are given by the relation "**to be an *external sign***";

The *SC-code* is not *language* in habitual understanding of this word, and the unified way graph (nonlinear generally) codings of information, providing submission of information of any level (primary, secondary, metalevel) and the description of *communications* between these levels. So, for example, ***primary sc-text*** can be the description of structure of some image, and various fragments of this text can be images of various observed objects which will be presented in the ***secondary sc-text*** by the signs which are obviously connected with the fragments of *primary sc-text* corresponding to them.

The same way *primary sc-text* can be the description of *structure* of the accepted text message, that is primary description of structure of some ***line of symbols***. It is obvious that some fragments of this *primary sc-text* are descriptions of structures of external signs of the corresponding described *entities*, that is structures of the words or phrases entering the message. It is obvious also that the specified described *entities* in the *secondary sc-text* representing sense of the accepted message will be designated by the ***internal signs*** which are obviously connected with the fragments of the specified *primary sc-text* corresponding to them.

B. The principles of memory organization ostis-system

Memory of ostis-system (sc-memory) is a non-linear (graph) associative restructuring (graphodynamic) memory, in which the processing of information is reduced not only to a change in the state of memory elements (any of which is a sign that is a part of the processed sign structure) but also to a change in the configuration of connections, the incidence between them.

From a formal point of view, the memory of *ostis-system* is a dynamic *sc-text*, in which the following events can occur:

- remove ***sc-node*** with the removal of all the *incident* ***sc-links*** (signs of various connections);
- remove the *sc-link* with the removal of all the *sc-links* which it is a component;
- replacing the type of *sc-text* element (for example, *sc-node* can be converted to ***sc-connector***);
- adding a new *sc-text* element with mandatory indication of the connection of the new *sc-element* generated in memory with any *sc-element* already presented (stored) in memory.

In *sc-memory* it is possible to store and process any ***external information constructions*** (terms, hieroglyphs, texts, images, video information, audio information) presented in electronic form as corresponding ***files***. At the same time each *file* in *sc-memory* is represented by the corresponding *sc-node* denoting this file (actually, the *external information structure* encoded by this file), and the specified file is considered the content of the denoted *sc-node*. Through such files, in particular, information is exchanged between *ostis-system* and external subjects (users, other *ostis-systems*).

For ensuring *flexibility* of *ostis-systems* the developed form of associative access to the any kinds of *knowledge* and *skills* stored in memory of *ostis-systems* thanks to the developed means of the specification of required *knowledge* is of particular importance.

So, for example, the procedure of associative search of *signs* on the basis of aprioristic information on communications between *signs* significantly becomes simpler. Such search is carried out by means of wave navigation on space of the connected *signs*.

Significantly also the procedure of associative search of the fragments of the stored *knowledge base* satisfying to the set *inquiries* (requirements) becomes simpler. Significantly the variety of types of such *inquiries* extends:

- *inquiriy* of fragments of the *knowledge base*, isomorphic to the set sample;
- *inquiriy* of fragments of the *knowledge base*, homomorphic to the set sample;
- *inquiriy* of the fragments of the *knowledge base* satisfying to the set not atomic logical formula that is the fragments satisfying to the set logical properties;
- request of the specification (the semantic vicinity) the set essence;
- *inquiriy* of the *entities* similar to the set essence (including to the set *knowledge base* fragment);
- *inquiriy* of communications (including similarities and differences) between the set *entities*.

C. The principles underlying the integrated solver of ostis-system problems

The integrated solver of ostis-system problems is a hierarchical system of ***agents*** that implement the *knowledge base* processing represented in *SC-code* and stored in *sc-memory*, and which interact with each other only through the specified *sc-memory*. Thus, the whole process of processing the *knowledge base* in *ostis-system* is controlled by this *knowledge base*.

These *agents*, which are called sc-*agents*, are divided into:

- ***non-atomic sc-agents***, which are collectives of sc-*agents* of lower level;
- ***Atomic sc-agents***, which are not collectives of sc-*agents*.

In its turn *atomic sc-agents* to their implementation level are divided into:

- ***atomic programm's sc-agents*** implemented in the ***basic programming language*** of *sc-agents*;
- **atomic sc-agents of the base programming language interpreter of sc-agents** (the interpreters of *knowledge base* and integrated solver of *ostis-system* problems).

In this case, each *sc-agent* has its own class of situations or events in *sc-memory* that initiate the activity of this *sc-agent*, generating the corresponding ***information process in sc-memory***, the main characteristics and the current state are described in *sc-memory* and used in this process.

Example1. Here is an example of parallel asynchronous multi-agent solution of a problem in *sc-memory*. Consider a problem of calculating the indicated number.

Calculation is the construction of representation of a number in a given system of calculus (for example, decimal).

Example of an agent-oriented model for solving problems in *ostis-system*.

Example problem: calculate the value of a number. This *problem* is a condition for the activation of a number of *agents*. It is done with the help of a meta-agent, which activates *agents*, focused on solving particular problems:

- ***sc-agent addition agent***
  - •• checks whether the specified number is the sum of the other numbers. Moreover, if these other numbers are calculated,then addition is performed, and if any of the terms is not known, a subtask for its calculation is generated, and the information process goes into the waiting state for solving this subtask;
- ***sc-agent subtraction agent***;
- ***sc-agent multiplication agent***;
- ***sc-agent division agent***;
- ***sc-agent involution agent***;
- ***sc-agent evolution agent***;
- ***sc-agent logarithm agent***;

Unfortunately, not every problem for calculating a number can be solved with the help of the specified set of ***arithmetic sc-agents*** and even with the help of a wider set of *sc-agents* that provide the calculation of all possible numerical functions. Examples of such problems are the problems of solving quadratic equations (Figure 2), the problem of solving systems of linear equations, etc. To solve such problems additional *sc-agents* are introduced, in particular, those that reformulate the original problem so that it can be solved, for example, only with the help of *arithmetic sc-agents*. Thus, for example, *sc-agent for solving quadratic equation* replaces the formulation of the problem of solving a quadratic equation (Figure 2) in the form shown in Figure 3.

Moreover, this kind of problems reformulation can be described in the form of ***products*** and included in *knowledge base* (actually, in skill base), and then implement *sc-agent* that provides interpretation of any products stored in ***ostis-system skill base***.

The unambiguous representation of information *in memory of ostis-system*, that is, the elimination in *SC-code* of duplication (semantic equivalence) of both *signs* and whole *sign constructions* significantly simplifies the processing of *knowledge base*.

Listing 4 provides clarifying of some concepts that underlie the organization of integrated problem solver of ostis-system.

operation
= *purposeful holistic process implemented by a certain subject within some dynamic system and effecting a change in the state of this system*
= *action*

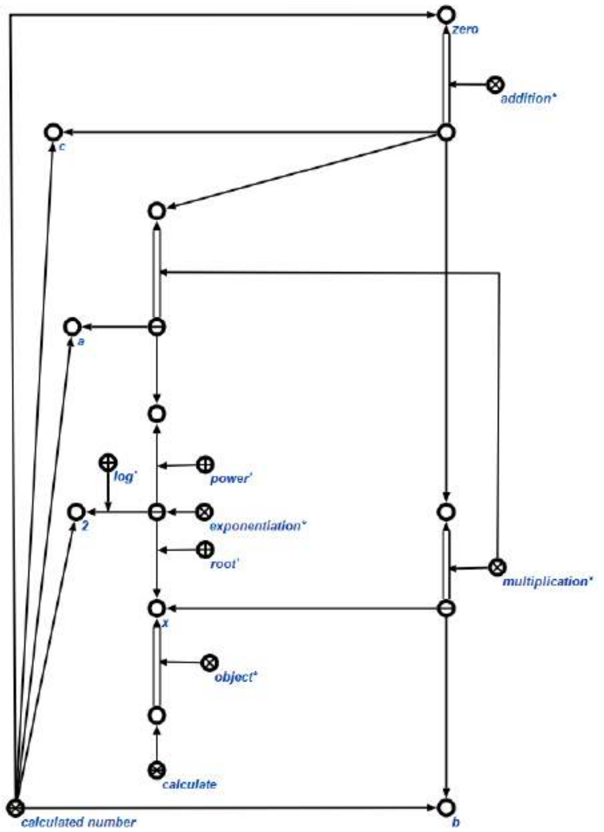


Figure 2. Formulation of the problem of solving a quadratic equation in SC-code

intelligent system operation
= *operation implemented by intelligent system*
⊃ *operation*
<= *partitioning **
{
- *internal operation of intelligent system*
  = *information process in memory of intelligent system*
  = *operation implemented in memory of intelligent system and aimed at changing the state of the stored knowledge base*
- *external operation of intelligent system*
  = *operation of intelligent system implemented in its external environment*
  ⊃ *atomic external operation of intelligent system*
    = *external operation of intelligent system implemented by one of its effectors*

}

problem
= *pragmatic specification of some (given) operation*
= *formulation of problem*
⊃ *problem of intelligent system*
  = *specification of intelligent system operation*

Listing 4. Clarifying concepts underlying the organization of integrated problem solver of ostis-system

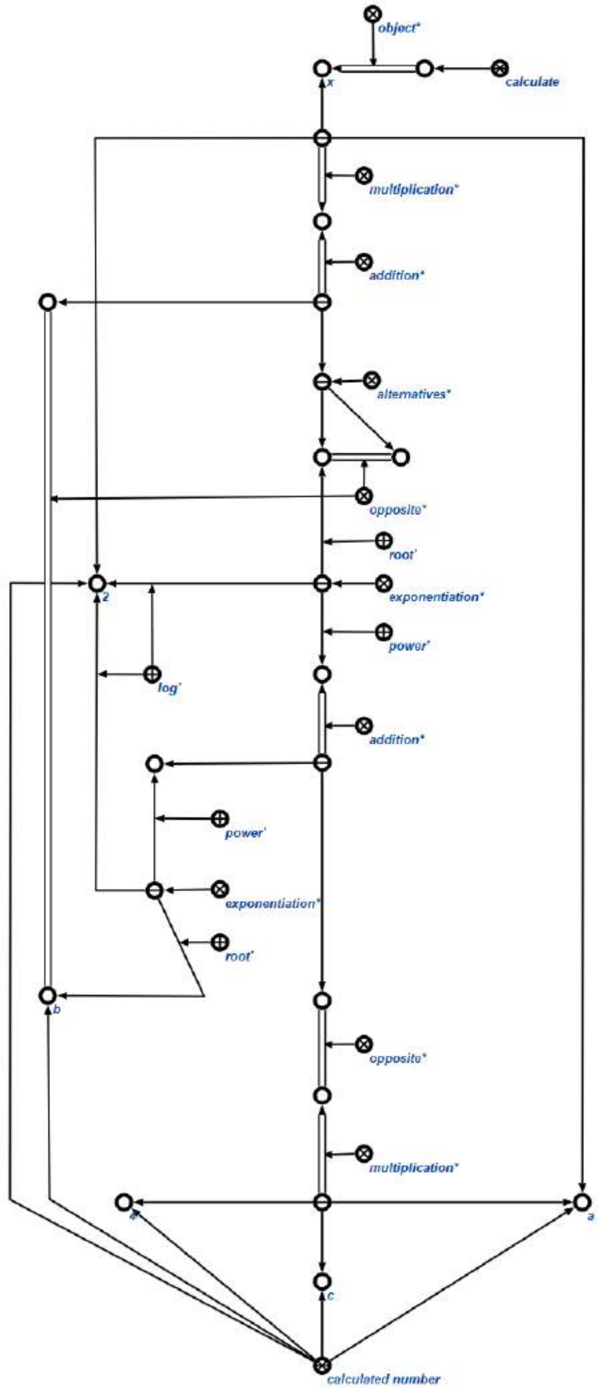


Figure 3. Result of reformulating the problem of solving a quadratic equation

Formulation of *problem*, as specification of the corresponding *operation,* can include:

- indication of ***classes of operations*** (preferably minimal) which the specified operation belongs to;
- indication of what should be done
  - •• either by indication of the ***target situation***;
  - •• or by indication of the *class of operations* which this *operation* belongs to with additional indication of objects which this *operation* is implemented with (action arguments);
- condition for initiating the specified *operation*. This condition can be either the completion of some other specified *operation,* or at least one of the specified *operations*, or the completion of all the specified *operations*, or the occurrence of the specified *situation* in *knowledge base*, or at least one of the specified *situations*, or the occurrence of the specified ***event*** in *knowledge base*, or at least one of the specified *events*, or all of these *events*;
- indication of a plan for the implemention of the specified *operation,* that is, hierarchical decomposition of this operation into operations of lower level;
- indication of an area for the implemention of the specified operation, that is, the structure included in *knowledge base*, which contains all (!) data that may be required to implement the specified operation;
- indication of temporal characteristics of the specified operation (start and / or end time, duration);
- indication of a customer for this operation;
- indication of implementers, that is, means of implemention of the specified operation (for internal specified operation of *ostis-system*, such mean is set of its several sc-*agents*);
- indication of status of the operation, that is, the indication that the specified implementer or collective of implementers may or may not, want or not want, should or should not implement the operation;
- indication of priority of the specified operation;
- indication of the state of the specified operation, which can be planned, implemented, directly implemented at the moment, interrupted.

skill
= skill of intelligent system
⊃ Skill of ostis-system = pragmatic specification of some (given) class of operations
= method (model) for solving problems belonging to given class of problems
= method to implement operations belonging to given class of operations
= knowledge of how to solve problems of given class
= knowledge of how to implement operations of given class

Listing 5. Formal interpretation of the concept of skill

In contrast to other types of *knowledge*, *skills* (Listing 3) have not only *denotation*al, but operational (functional) semantics, which describes what operations should be implemented when this skill is used.

Skill of ostis-system is called *knowledge*, which is a part of *knowledge base* of this system and which is sufficient to solve given class of problems (that is, to implement given class of operatons) within given *knowledge base* area using given family of *ostis-system agents*.

Each *Skill of ostis-system* is represented in *knowledge base* in the form of a specification of the *class of operations* implemented by using this skill.

Skill of ostis-system as a specification of ***some class of ostis-system operation*** includes:

- indication of the corresponding class of *ostis-system operation*;
- generalized formulation of *problems* of the corresponding class, that is, generalized specification of operations of the specified class (indicating what should be given and what is required);
- indication of area for solving all problems of given class, that is, the fragment of *knowledge base* that contains all the necessary information for solving all problems of given class;
- indication of declarative or procedural program, which interpretation guarantees implementation of all or almost all operations belonging to the given specified *class of operations*;
- means for implementation operations of given class, which are set of sc-*agents* that ensure implemention of all or almost all operations belonging to the specified *class of operations* and defining the operational semantics of the interpreted program mentioned above and, accordingly, the operational semantics of whole skill.

Stress that in most cases, area for solving problems is a certain ***subject domain*** together with its ***integrated ontology***, and interpreted program is ***logical ontology*** of the specified *subject domain*.

Hierarchy of *ostis-system skills*:

- skill of base (zero) level – the skill realized by one of the *agents* of base machine of *ostis-system*, which is the interpreter of *knowledge* and *skills* of *ostis-system* stored in its *sc-memory*;
- skill of level 1 – the skill, implemented by collective of all *agents* of base machine of *ostis-system*;
- skill of level 2 – the skill, implemented by one of *atomic program* sc-*agents*;
- skill of level 3 – the skill, implemented by collective of *atomic program* sc-*agents*;
- skill of level 4 – the skill implemented by collective consisting of *atomic program* sc-*agents* and collective consisting only of *atomic program* sc-*agents*.

Thus, activity of *ostis-system* is clearly *stratified* into several sufficiently independent levels, each of which ensures implementation of clearly defined set of operations:

- level of operations implemented by collectives of sc-*agents*;

- level of operations implemented by *atomic program* sc-*agents*;
- level of operations implemented by interpreter of *ostis-system knowledge* and *skills* when interpreting programs of *atomic program* sc-*agents*.

D. Flexibility of knowledge bases sc-models

The unambiguity of information representation in memory of *ostis-system*, that is, the elimination in *SC-code* of duplication (semantic equivalence) of both signs and whole *sign constructions* essentially simplifies the processing of the *knowledge base*.

The labor content of editing *sign constructions* stored in *sc-memory* at a semantic level essentially reduces compared, for example, with the editing of database or texts of natural languages. In particular, it is much easier to remove a sign from *sign construction* in *SC-code* without damage to its integrity.

Rules for removing *sc-elements*:

1) If *sc-element* is deleted, then all (!) *sc-links* (both atomic and non-atomic) incident to it are deleted recursively;
2) If ***sc-arc of membership*** leaving the sign of the ***formed set*** is deleted, then the indication of the formation of this set is deleted and an indication of the power of this set is added (if it was absent).

Flexibility of knowledge bases sc-models is also ensured by a clear separation of external and *internal signs* and complete independence of *knowledge bases* sc-model from the variety of forms and languages of this *knowledge base* external representation.

Here are some examples of editing *knowledge base* sc-model.

Example 2. Replacement of the term

A name (term) of each described entity is present in *ostis-system knowledge base* in one instance and is represented in *knowledge base* by *sc-node* denoting this name, which is connected by the relation external *sign with *sc-element* denoting the same entity as the specified name (Figure 4), and this *sc-element*, in turn, is connected by unlimited number of connections with *sc-elements*, which denote other *entities* described in *knowledge base*.

Therefore, replacement of a name (term) of an entity is implemented only in one place of *knowledge base*.

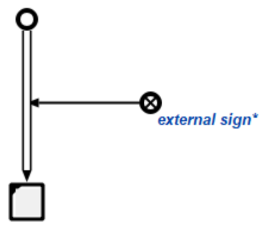


Figure 4. The relationship of a term with a sc-element that identify the same entity

Example 3. Replacement of concepts

Instead of the concept of *partitioning* * of set, proceed to use the following concepts:

- *union** of sets (covering a set);
- *family of pairwise disjoint sets*.

For such replacement, you need a concept that becomes unused to define through the used concepts (Figure 5).

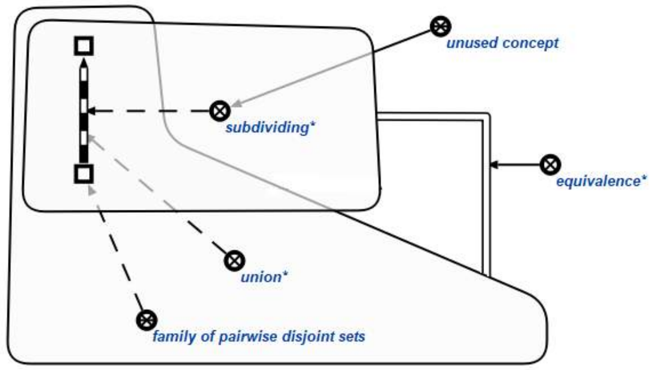


Figure 5. Definition for concept of partitioning*

The purpose of this kind of replacement of concepts is to minimize the number of defined concepts (some defined concepts are made unused). To eliminate the concept means to remove all arcs of membership from it, leaving the concept as unused.

1) removal of all arcs of memdership emerging from it;
2) construction of logically equivalent specifications for each instance of the concept presented in *knowledge base*;
3) inclusion of this concept in the number of unused concepts.

E. Flexibility of ostis-system integrated solver is provided by:

- autonomy of *sc-agents*;
- interaction of *sc-agents* only through *sc-memory*;
- using the emergence in *sc-memory* of formulations of initiated problems as the main events initiating activity of the corresponding *sc-agents*;
- developing of powerful language of the formulations of various kinds of problems on the basis of *SC-code*;
- clear separation of variety of various skills of the solution of the tasks stored by *ostis-system*, from the interpreter of these skills and ensuring independence of the specified skills of variety of options of realization of the specified interpreter;
- the fact that the skills acquired by *ostis-system* are a part of her *knowledge base* and are stored in her memory, therefore, them easily to fill up and correct.

Correction of *knowledge base* is based on the fact that there is no duplication of information both at *sc-element* level and at sc-construction level in *knowledge base*, it means that all such corrections are of local character (within the semantic area of fragment correction is not required global revision). But it is necessary to pay for such locality, it should be constantly maintained. This requires:

- constantly improve structure of *knowledge base* (system of subject domains and ontologies);
- use special sc-agents in *knowledge base* of *ostis-system* to perform permanent analysis of the quality of *knowledge base* for the presence and elimination of synonymous sc-elements and semantically equivalent sc-constructions (not to be confused with logical equivalence);
- each *sc-agent* within each *atomic sc-process* complies with certain rules of *sc-memory* usage that are common for all *sc-agents* of *ostis-system*
  - removes *information waste* (information necessary only for performance of the specified atomic process);
  - ensures that synonyms and semantic equivalent *sc-texts* weren't generated;
  - ensures that that the *atomic processes* which are carried out by him didn't interfere with each other.

F. Flexibility of ostis-systems

in general lies in the local character of various kinds of *intelligent system* transformations. It means that consequences of transformations should be considered within the known and clearly defined parts of system (fragments of *knowledge base* of integrated solver). When strict requirements presented by *OSTIS Technology* to the integrity of *intelligent systems* are observed, you can implement any(!) changes in *ostis-system knowledge base*, in sc-model of its integrated solver in base interpreter of sc-model of *intelligent system*. Compliance with these requirements ensure locality of these changes, that is, complete independence of some changes from others.

Localization of problem-solving processes in *ostis-system* is implemented due to the fact that *knowledge base* of *ostis-system* has hierarchical structure that provides clear allocation of solving area for each problem of the fragment of *knowledge base* which is sufficient for solving one form of problem or another. The structure of *knowledge base* is based on hierarchical system of *subject domains* and corresponding ontologies.

IX. Learnability of ostis-systems

Learnability of ostis-systems is expressed in the following:

- detection and localization of contradictions in *knowledge base*;
- elimination of some contradictions in *knowledge base*;
- detection and localization (specification) of *information holes* in *knowledge base*;
- detection and elimination of *information waste* in *knowledge base*;
- analysis of the solver performance in solving various kinds of problems;
- detection of contradictions (errors) that appeared in *knowledge base* as a result of the activity of sc-*agents* specifying specific problems and specific atomic information processes (as sources of errors);
- detection of contradictions, which are consequences of activity of *knowledge base* developers , with indicating execution time and authors of these actions and with organization of interaction with the developers to eliminate these contradictions;
- constant updating of enumeration of problems which could not be solved or succeeded, but not as desired (insufficiently accurate, insufficiently fast, insufficiently transparent, insufficiently simple, insufficiently clear);
- permanent improvement of *knowledge base* structure, the purpose of which is to increase accuracy of information resources localization sufficient for solving various problems solved by *intelligent system*, including fundamentally new problems for it;
- permanent search and improvement of ways and effective models for solving fundamentally new problems for *intelligent system*, which are necessary to be solved during functioning of *intelligent system*.

A. Ability to acquire new knowledge generated by SC-agents of ostis-system in internal sense language

To acquire new *knowledge* generated by some sc-agent of *ostis-system* and, accordingly, located in its *sc-memory*, is to immerse (integrate) the specified *knowledge* into *knowledge base* of *ostis-system*. The most important advantage of *SC-code* is that this procedure of immersing (integrating) new *sc-texts* into *knowledge base* is reduced to constructoin of correspondence of synonymy between *sc-elements* (both *sc-nodes* and *sc-connectors*) of new immersible *sc-text* and *sc-elements* that are part of the current state of *knowledge base*, followed by gradual gluing (identification) of synonymous *sc-elements*. The construction of this correspondence of synonymy is implemented, firstly, by the coincidence of main names * (names, *external signs*) corresponding to some *sc-elements*, and, secondly, on the basis of statements describing existence and uniqueness of *entities* possessing given characteristics. Stress that not all *sc-elements* are named, that is, having corresponding external names (*external signs*). Only the used concepts should be named, each of which should be researched at least in one of the selected *subject domains*. In addition, several names can correspond to some *sc-elements*, one of which is allocated as the main name * of the corresponding *sc-element*, which the lack of synonymy and homonymy towards the *main names** of other *sc-elements* is guaranteed for.

Example 4. An example of immersing a new *sc-text* in *knowledge base*. Figure 6 shows *sc-text* that is in the current state of *knowledge base*. Figure 7 shows *sc-text* that is immersed in *knowledge base* (introduced into *knowledge base*). Figure 8 shows the result of gluing *sc-elements* by the coincidence of their *main names**, after that several new *entities* appeared in the current state of *knowledge base* – a new triangle and two bisectors and their intersection point, a new circle corresponding to the new triandle.

Figure 9 shows the final result of immersion of considered new text in *knowledge base*, derived on the basis of using the following statements about uniqueness stored in *knowledge base* of ostis system:

1) Each *triangle* has a single *bisector* having a common element, which is the *vertex* of this triangle.
2) Every two *triangles* having the same *vertices* are the same triangle.
3) Each *triangle* has a unique inscribed circle.
4) All bisectors of one *triangle* intersect at one *point*.
5) The point of intersection of its bisectrix and the center of the inscribed circle coincide for every triangle.
6) Each two-fold constant positive *arcs of membership* going out of the same set that is not a multiset is the same arc.
7) Membership *bisectrix** does not have multiple pairs.

Comparing Figure 8 and Figure 9, it is easy to see that there was not really much new information contained in the new immersed *sc-text* – there is only another bisector of the already known triangle and that the center of the inscribed circle coincides with the point of intersection of its bisectors.

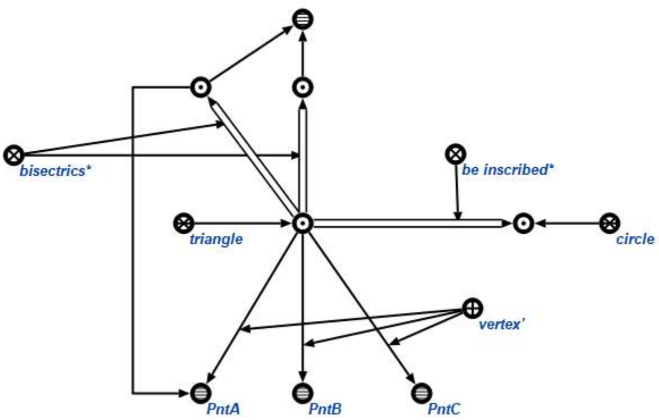


Figure 6. *Sc-text* that is in the current state of *knowledge base*

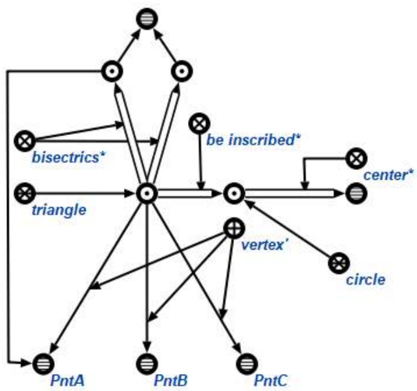


Figure 7. *Sc-text*, immersed in *knowledge base*

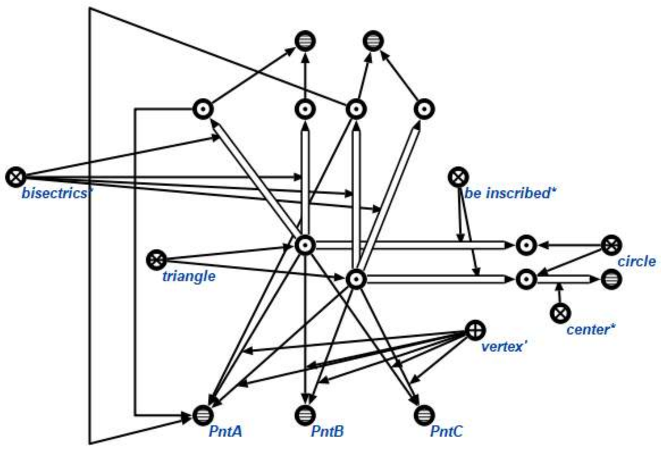


Figure 8. The result of gluing *sc-elements* by the coincidence of their main names

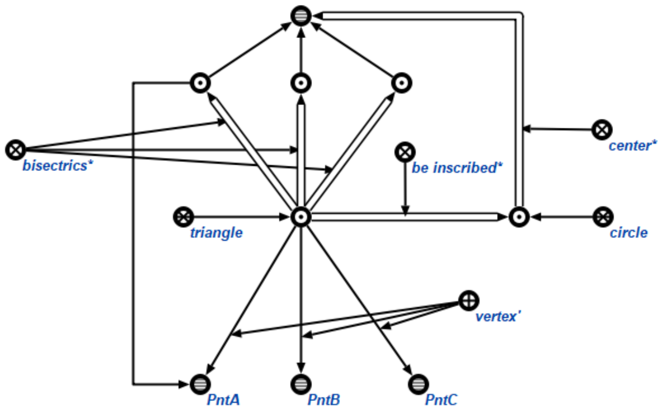


Figure 9. The result of gluing unnamed synonymous *sc-elements* (the final result of immersing introduced sc text into *knowledge base*)

B. Ability to acquire new knowledge and skills coming from outside in formal languages

As the main languages providing the representation of *external texts* for *ostis-system* are used:

- ***SCg-code*** (Semantic Computer graphical Code), which provides graphic representation of *SC-code* texts, in which *sc-nodes* are represented as circles, squares, closed lines representing signs of sc-structures whose images are limited by these lines, and *sc-connectors (sc-arcs* and sc-edges) are represented as lines connecting *sc-node* images;
- ***SCs-code*** (Semantic Computer string Code), which provides string (linear) representation of *SC-code* texts and used by *ostis-system* for input and output of relatively short messages;
- ***SCn-code*** (Semantic Computer natural Code), which provides intuitively understandable formatted representation of large *source code*.

The syntax and declarative semantics of these languages are close to the syntax and semantics of *SC-code*. Thanks to this, the syntactic and semantic analysis of input texts is substantially simplified, as well as the translation of *sc-texts* in specified external languages when messages addressed to users are generated. Proximity of the mentioned external languages to *SC-code*, first of all, is that all used names are the *main names** (main *external signs**) of the corresponding *sc-elements*. Stress that the list of used external formal languages for *ostis-systems* is easily extended, because it is not very difficult to describe syntax and semantics of any external formal language using the *means of SC-code* within ***interface knowledge base***, and it is not difficult to develop a family of *sc-agents* that provide:

- making ***syntactic analysis*** of external source code based on the description of syntaxis of the corresponding external language. The result of this analysis is *sc-text* describing the *syntactic structure* of the external source code with such a degree of detailing that is sufficient for the subsequent ***semantic analysis*** of this code;
- making *semantic analysis* of external source code based on the description of semantics of the corresponding external language.

The result of this analysis is construction of *sc-text* semantically equivalent to external source code. After that, the immersion of *sc-text* described above into the current state of *knowledge base* is implemented.

But a better understanding by *ostis-system* of the accepted external message implies the implementation of the following additional actions:

- Coordination of *the concepts* used in the accepted external message with the concepts used in the current state of *knowledge base* of *ostis-system* (see example 3);
- Inclusion of acquired *knowledge* into general system of accumulated *knowledge*, in particular, into hierarchical system of the selected *subject domains* and corresponding *ontologies*.

C. Ability to acquire new knowledge and skills coming from outside in non-formal languages

Acquisition (understanding) of *knowledge* entering *ostis-system* in ***natural languages*** consists of the same stages as the process of understanding external texts of formal languages considered above, but fundamentally differs from the understanding of external texts of formal languages due to the fact that the syntax and semantics of natural languages have a significantly higher level of complexity compared with the syntax and semantics of formal languages. It means that the development of ***knowledge base of ostis-system natural-language interface***, as well as the development of system of *sc-agents* of the interface is very complex problems.

Nevertheless the principles which are the cornerstone of development of ***verbal interfaces of ostis-systems*** have the serious advantages which are that external (including, and natural language) texts are considered by *ostis-system* as a part of her external environment and can be described in memory of *ostis-system* in her internal language (in a *SC-code*). Thus, the structure of the external text can be described with any extent of specification means of a *SC-code*. Therefore, ***semantics*** of the external text will represent a morphism between the *sc-text* describing structure of the source text with the required extent of specification, and the *sc-text* which is semantic equivalent to the specified source text.

That fact, as the description of structure of the source text, and the *sc-text* semantic equivalent to him, and the description of a morphism between them are presented in one language (*SC-code*) and are in one memory (*sc-memory*), creates good prerequisites for effective realization of semantico-syntactical approach to the analysis of the source text assuming lack of clear split of process of the analysis of the source text in time for a stage of parse and a stage of the *semantic analysis*. Such approach allows to reduce labor input and the syntactic and *semantic analysis* thanks to preliminary results of the semantic analysis when performing parse and vice versa.

D. Ability to acquire knowledge coming from outside through sensors of ostis-systems

Advantage of *ostis-systems* in the analysis of primary sensor information, as well as in the analysis of external texts, is defined by that, as intellectual (in that number, semantic) the analysis of this information it is easy to present primary sensor information, secondary sensor information and result in memory of *ostis-system* in the form of *sc-texts* that allows to pass easily from the "syntactic" level of sensor information to "semantic" level. It is relevant and for different identification of emergency situations in difficult, for example, production, systems to which it is necessary to react quickly, and for ***the semantic analysis*** of images and scenes.

It is necessary to pay attention to analogy of transition from the description of *syntactic structure* of the source external text to sense of this text (that is to semantic equivalent *sc-text*) and transition from the description of structure of primary and secondary sensor information to the description of communications between those objects which are reflected in this sensor information.

Thus, recognition of various objects in a flow of sensor information is an analog of transition from the phrases which are a part of the source external text, and, more precisely, from the *sc-texts* describing structure of these phrases to the *sc-elements* designating the same *entities*, as the specified phrases.

E. Ability to detect and eliminate contradictions in knowledge base

Associative search for sc-structures that satisfy statements about non-existence and uniqueness is generalized description of errors.

Example 5. Here is an example of contradictions that are the consequence of the actions of *knowledge base* developers.

Suppose that in the fragment of *knowledge base* shown in Figure 10, *knowledge base* developer introduced updated

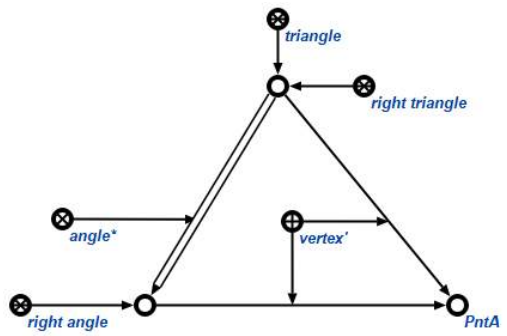


Figure 10. *Knowledge base* fragment about triangle

information about the magnitude of the considered angle with the vertex at point A and let this magnitude be 88^0.

Then:

- the considered angle ceases to be right;
- the considered triangle ceases to be right;
- if the magnitudes of remaining angles of the considered triangle are known, then the sum of the angles of the triangle becomes less than 180^0, and this contradicts the well-known theorem. It means that the magnitudes of the other angles of this triangle should be clarified.

F. Ability to search and eliminate information holes in knowledge base

Search and elimination of *information holes* in *knowledge base* is implemented on the basis of:

- generalized description of required typical specifications for different classes of *entities*;
- generalized quasi-plicative statements describing not what must exist in given infinite *subject domain* in given conditions, but what must be present in *knowledge base* in given conditions.

The following provisions are the basis of detection and specification of *information holes* in *knowledge base*:

- any information should be included in any *subject domain* and/or any specific ontology;
- there is typical structure of mandatory and desirable specification of each instance of this class for each entity class;
- these classes have hierarchy (!) with inheritance of typical structures of specifications of their superclasses' instances;
- unambiguous specifications play important role (for example, representation of numbers in different systems of calculus); classes of *entities* that have typical specifications, in particular, include:
  - •• *entity*,
  - •• *set*,
  - •• *structure*,
  - •• *subject domain*,
  - •• *concept*,
  - •• *relation*,
  - •• *binary relation*,
  - •• *external information construction*,
  - •• *person*,
  - •• *number*,
- any ***concept*** should
  - •• included in any ***subject domain*** as the researched (class of researched objects, researched characteristic, researched relation);
  - •• be defined or explained;
  - •• have set of propositions describing the regularities which are satisfied by instances of this concept;
  - •• have typical examples (instances).

G. Ability to search and eliminate information waste in knowledge base

The specified ability is realized, first of all, by each sc-agent, because each of them after the completion of each actions is obliged to "clean up after itself", that is, to remove all auxiliary structures created only for implementation of corresponding action. In addition, special sc- *agents* are used to eliminate *information waste*. Activity of these sc-*agents* is the opposite of activity of sc-*agents* for removing *information holes*, but it is organized similarly on the basis of generalized quasi-replicative statements which describe that should be removed from the current state of *knowledge base* in given conditions.

H. Ability to systematize the acquired knowledge and skills

As a result of such systematization, each acquired *knowledge* and, in particular, skill should be included in one or several structures that are the results of decomposition of *knowledge base* of *ostis-system* on various grounds. One of the main types of such decomposition is logical-semantic decomposition of *knowledge base*, the result of which is hierarchical system of *subject domains* and corresponding ontologies.As a result of such systematization, each acquired *knowledge* and, in particular, skill should be included in one or several structures that are the results of decomposition of *knowledge base* of *ostis-system* on various grounds. One of the main types of such decomposition is logical-semantic decomposition of *knowledge base*, the result of which is hierarchical system of *subject domains* and corresponding ontologies.

X. Unlimited learnability of ostis-systems

Unlimited learnability of ostis-systems is provided by:

1) Universality of *SC-code*, that is, ability to represent any (!) types of *knowledge* and *skills* in *SC-code*.
2) Universal character of the ***agent-oriented organization of knowledge processing***, which allows on its basis to interpret any (!) models of information processing (sequential and parallel, synchronous and asynchronous, functional, production, genetic, neural network, logical - clear and fuzzy, accurate and plausible , deductive, inductive, abductive, argumentative).
3) ***Compatibility*** of all *knowledge*, *skills*, models, information processing used in *ostis-system* in appropriate coordination with used concepts.

All variety of *knowledge* types and variety of problem-solving models types can be formally described in the form of hierarchical system of interrelated *subject domains* and ontologies. In other words, any kind of *knowledge* can be treated as a language that is sublanguage of *SC-code*, the union of the texts (*sc-texts*) of which is the corresponding *subject domains*, and the description of semantics of which is ontology of the specified *subject domain*. As for problem-solving models, each of them is formally defined by (1) a ***language of problems***, which is sublanguage of *SC-code*, and in particular, describing the ***decomposition of problems into subproblems*** (the union of the texts of such problems language is ***subject domain of problems*** solved within the corresponding *subject domain*) and (2) a set of sc-*agents* that ensure all problems solution within the specified *subject domain* of problems, that is, ensure interpretation of the specified problems language and define ***the operational semantics*** of this language.

Since all kinds of *knowledge* and problem-solving models can be represented using *SC-code* as a part of integrated *knowledge base* of *ostis-system* stored in *sc-memory* of this system, and since different kinds of *knowledge* and different kinds of problem-solving models will intersect if they contain signs with the same *denotations*, the joint use of different types of *knowledge* and problem-solving models is not very difficult. Ensuring the joint use of different types of *knowledge* and different types of problem-solving models makes it possible not only to compensate for disadvantages of some models with the advantages of others, but also to significantly expand the range of solved problems, firstly, due to different problem-solving models provide solutions of different types of problems in general (sometimes even intersecting), and, secondly, due to possibility of solving complex problems which can not be solved in principle with the help of one model and can be solved only by the joint use of different problem-solving models.

XI. Training and self-learning of ostis-systems

Training ostis-systems is a process of improving (designing) *ostis-system*, implemented during its operation.

The efficiency and, first of all, the speed of *training of ostis-systems* is defined not only by the high level of *learnability of ostis-systems*, but also efficiency of ***the ostis-systems training technique*** realized by group of *developer teachers*. In spite of the fact that there are general principles of the organization of project management, accounting of specifics of designed projects (in this case, *ostis-systems*) will allow to increase efficiency of such management significantly. In particular, when training *ostis-systems* there can be very useful that experience which the mankind had accumulated when training children.

When training *ostis-systems* increase in level of activity of the trained *ostis-system* in the course of the training is perspective. It has to be carried out (1) in the direction of development of means of self-training and (2) in the direction of increase in level of automation and intellectualization of management of activity of *developer teachers*.

In this case the developer of *ostis-system* becomes the external agent working on the *knowledge base* of this system and interacting with other developers through the same memory in which the specified *knowledge base* is stored.

XII. The relationship between flexibility and learnability of intelligent systems

Flexibility is a system attribute that characterizes the ability (freedom) to make the system changes that are being implemented or t *Flexibility* is characteristic of system that characterizes possibility (freedom) of introducing certain changes into system, which is implemented either by system's teacher-developers or by the system itself.

Learnability is characteristic of system that characterizes ability to use its own flexibility (that is, freedom of self-change) so that quality of system increases. In particular, learnability is ability to compensate for the negative consequences of teacher-developers actions, fraught with destruction of system integrity.

The simplest form of training *intelligent system* is loading new *knowledge* and / or *skills* into its memory, or "manually" editing *knowledge* and *skills* which are already in memory. But even such simple form of training requires from *intelligent system* not only appropriate level of *flexibility*, that is, providing quite convenient possibility of "manual" replenishment and editing of *knowledge* and *skills* of *intelligent system*, but also appropriate level of their *learnability*.

Principle possibility "manually" and without any restrictions to replenish and correct programs and processed data stored in memory is present in modern computer systems. But, firstly, you can not follow every byte in large systems; secondly, "manual" intervention in large system (for example, changing relational database schema) can have difficult predictable consequences and is powerful source of errors introduced into system.

Consequently, it is important not only to ensure high level of *flexibility* in large computer systems and, in particular, to automate changes introduced to system, but also to ensure sufficient level of *learnability* to prevent negative consequences of these changes.

This is the transition from systems with high level of *flexibility* to *trained* systems. Stress that transition to trained systems is natural standard for evolution of computer systems when their level of complexity increases (with increase in amount of processed data and number of used programs).

Here are some examples of the relationship between the features of *flexibility* and the features of *learnability* of *intelligent systems*:

1) Artificial neural network:
   - *flexibility*: easy ability to change parameters of artificial neural network;
   - l*earnability*: the presence of procedure that changes parameters of artificial *neural network* so that the efficiency of solving problems for which it is intended increased (for example, procedure for back propagation of errors);

2) *ostis-system*:
   - *flexibility*: there are ample opportunity for editing *knowledge base* (including its restructuring, reconfiguration) and high level of such editing automation;
   - *learnability*: there are developed means that ensure the integrity of *knowledge base* (its consistency, completeness, clarity);
3) *ostis-system*:
   - *flexibility*: there are advanced forms of associative access to the searched (askable) fragments of *knowledge base* due to the lack of information duplication in *knowledge base* (that is, the lack of synonymy of internal *skills* and semantic equivalence of internal *sign constructions*);
   - *learnability*: significant simplification of means for analyzing quality of the current state of *knowledge base*;
4) *ostis-system*:
   - *flexibility*: local character of all changes made in *ostis-system* (and, first of all, in its *knowledge base*) due to clear stratification of *ostis-system*, which ensures high degree of independence of changes in different parts of *ostis-system*;
   - *learnability*: significant simplification of means ensuring integrity of *ostis-system*;
5) *ostis-system*:
   - *flexibility*: possibility within internal sense code of unlimited transition from *knowledge* to *meta-knowledge*;
   - *learnability*: ability to specification and systematization of acquired *knowledge* and *skills* and, as a result, ample opportunities for decomposition and structuring of *knowledge base* and *skills* on various grounds;
6) *ostis-system*:
   - *flexibility*: unification of internal representation of *knowledge* and *skills* in memory of *ostis-system* on the basis of hierarchical system of *subject domains* and ontologies, and as a result, simplification of procedures for editing *knowledge base* due to significant reduction in the number of pairs of logically equivalent fragments of *knowledge base*, and also due to the rule of inheriting characteristics with the construction of specifications for instances of concepts that are subclasses of other concepts known to the system;
   - *learnability*: ability to ensure compatibility of *knowledge bases* of various *ostis-systems* and, as a result, to ensure compatibility of these systems;
7) *ostis-system*:
   - *flexibility*: possibility of detailed logging in ostis-system memory of all internal processes occurring in this memory;
   - *learnability*: significant simplification of means to establish causes and sources of emergence of various kinds of contradictions (including conflicts between internal information processes);
8) *ostis-system*:
   - *flexibility*: possibility to change parameters of *integrated problem solver* (in particular, level of each internal agent of *ostis-system* activity);
   - *learnability*: constantly improved procedure aimed at improving performance and quality of functioning of *integrated problem solver* of *ostis-system* (in particular, by redistributing levels of internal *agents* activity).

XIII. Flexibility and learnability of OSTIS technology

Flexibility and unlimited learnability of ***OSTIS Technology*** are determined by the fact that the specified technology is implemented in the form of *ostis-system*, that is, it is implemented using the same *OSTIS Technology*, the support of which it implements. The specified *ostis-system* is called ***metasystem IMS.ostis*** (Intelligent Meta System Open Semantic Technology *Intelligent systems*).

Thus, we focus not only on creating *OSTIS Technology*, but rather on creating process that supports high rates of permanent evolution of this technology.

XIV. Training and self-learning of OSTIS technology

Training and *self-learning* of *OSTIS Technology* (that is, permanent improvement of metasystem *IMS.ostis* during its operation) have a number of features:

1) *IMS.ostis* training (improvement) is implemented within open project that lets anyone to enter the number of developers of this *metasystem*.
2) ***metasystem IMS.ostis*** *self-learning* can be implemented on the basis of analysis of ***child ostis-systems*** operation and evolution building with the help of this metasystem. But it requires the organization of effective interaction between metasystem ***IMS.ostis*** and the child *ostis-systems* generated by it.
3) Rapid evolution of *OSTIS Technology* requires coordination's automation of versions of various ***ostis-systems components*** with current version of *OSTIS Technology* and, in particular, automation of outdated reusable components replacement with current versions of these components in child *ostis-systems*.

XV. Conclusion

Stress that transition from traditional computer systems is the final stage in the evolution of computer systems in the direction of increasing their flexibility, learnability, providing their unlimited (universal) learnability and logical-semantic compatibility.

Learnability of *intelligent systems*, being an important characteristic of *intelligent systems*, creates good background for significant expansion of their life cycle and ensuring high level of ***competitiveness*** in comparison with traditional computer systems.

But to ensure competitiveness of *intelligent systems*, in addition to their learnability, a well-thought-out ***methodology for their training*** (improvement) is needed. ***Management of project aimed at training intelligent system*** has serious features (primarily due to the specifics of object of training). Especially when this is about open projects.

In addition, one of the subjects of *intelligent system* training project can be intelligent metasystem *IMS.ostis*. But for this it is necessary to develop principles of interaction of *intelligent systems* (in the transition to collectives of *intelligent systems*). At the same time, if high rate of *IMS.ostis* evolution is ensured, then the result is high competitiveness not only of the technology, but also of *intelligent systems* developed on its basis.

High level of *ostis-system learnability* is the basis for ensuring high rate of *OSTIS Technology* evolution, as this technology is also implemented as *ostis-system* (*IMS.ostis*). It is essential that high rate of *OSTIS Technology* evolution is factor in accelerating (!) rate of various ostis systems evolution.

The most important direction of evolution of *OSTIS technology* is expansion and improvement of structure of ***library of reusable components*** of *ostis-systems* that provides essential decrease in labor input of development of ostis-systems.

Competitive technology for development of *intelligent systems* should have high rate of unlimited evolution (improvement, training) and should be oriented towards development of *intelligent systems* that have high level of learnability. But for this, ***general theory of compatible (joint used) knowledge and problem-solving skills***, as well as ***general theory of unlimited trained intelligent systems***, should be developed.

High rate of *OSTIS Technology* evolution, open nature of this technology, and open nature of participation in its development (within open-source project *IMS.ostis*) ensure high competitiveness of *OSTIS Technology.*

Acknowledgment

The work is executed at support of the grant BRFFR-RFBR No. Ф16Р-101 "Development of intelligent learning systems, and synergetic training organizations based on open semantic technologies, ontology engineering and patterns of understanding", and No. Ф16Р-102 "Formalization of temporal reasoning in intelligent systems".

References

[1] (2017, Dec.) The IMS.OSTIS website. [Online]. Available: http://www.ims.ostis.net

[2] E. F. Skorokhod'ko, *Semanticheskie seti i avtomaticheskaya obrabotka teksta [Semantic networks and automatic text processing]*. Kiev: Nauk. dumka, 1983, (in Russian).

[3] L. S. Bolotova, *Sistemy iskusstvennogo intellekta: modeli i tekhnologii, osnovannye na znaniyakh: uchebnik [Systems of artificial intelligence: the models and technologies based on knowledge: textbook]*. M.: Finansy i statistika, 2012, (in Russian).

[4] G. S. Osipov, *Metody iskusstvennogo intellekta [Methods of artificial intelligence]*. M.: FIZMATLIT, 2015, (in Russian).

[5] V. V. Golenkov, N. A. Gulyakina, and T. A. Gavrilova, "Unifitsirovannyi sposob abstraktnogo kodirovaniya semanticheskikh setei [unified way of abstract coding of semantic networks]," in *X mezhdunarodnaya nauchnaya konferentsiya imeni T.A.Taran «Intellektual'nyi analiz informatsii IAI-2010» [Intellectual information analysis]*, Kiev, May 2010, p. 46–52, (in Russian).

[6] V. V. Golenkov and N. A. Gulyakina, "Strukturizatsiya smyslovogo prostranstva [structurization of semantic space]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, Minsk, Feb. 2014, p. 65—78, (in Russian).

[7] V. B. Tarasov, *Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam [From multi-agent systems to intelligent organizations]*. M.: Editorial URSS, 2002, (in Russian).

[8] S. I. Nikulenkov, *Tulup'ev A.L. Samoobuchayushchiesya sistemy [Self-learning systems]*. M.: MTsNMO, 2009, (in Russian).

[9] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction (2nd Edition, Complete Draft)*. London: The MII Press, 2017.

[10] A. P. Eremeev and A. A. Kozhukhov, "Implementation of reinforcement learning tools for real-time intelligent decision support systems," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, Minsk, Feb. 2017, pp. 183–186.

[11] P. Flakh, *Mashinnoe obuchenie. Nauk i iskusstvo poctroeniya algoritmov, kotorye izvlekayut znaniya iz dannykh [Machine learning. The sciences and the art of constructing algorithms that extract knowledge from data]*. M.: LVR Press, 2015, (in Russian).

[12] A. A. Barsegyan, M. S. Kupriyanov, V. V. Stepanenko, and I. I. Kholod, *Tekhnologii analiza dannykh: Data Mining. Text Mining, OLAP [Data analysis technologies: Data Mining. Text Mining, OLAP]*. Spb.: BKhV-Peterburg, 2007, (in Russian).

ОТ ОБУЧЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ К ОБУЧЕНИЮ СРЕДСТВ ИХ РАЗРАБОТКИ

Голенков В.В., Гулякина Н.А.,
Гракова Н.В., Давыденко И.Т., Никуленко В.Ю.
Белорусский государственный университет
информатики и радиоэлектроники
Минск, Беларусь
Еремеев А.П.
Национальный исследовательский университет
«МЭИ», Москва, Россия
Тарасов В.Б.
Московский государственный технический
университет им. Н.Э.Баумана, Москва, Россия

В статье рассматривается подход к реализации перехода от обучения интеллектуальных систем, основанных на знаниях, к обучению средств их разработки. При этом архитектура такой интеллектуальной системы рассматривается как основа обеспечения её гибкости и обучаемости. Кроме того, в статье рассматриваются направления обучения и самообучения интеллектуальных систем, а также их способность приобретать знания и навыки из различных источников. Даётся обоснование применения технологии OSTIS для разработки интеллектуальных систем, основанных на знаниях.

Semantic models, method and tools of knowledge bases coordinated development based on reusable components

Irina Davydenko
Belarusian State University
of Informatics and Radioelectronics
Minsk, Belarus
ir.davydenko@gmail.com

***Abstract*—The paper considers the approach to the development of knowledge bases based on semantic models. The problem of collective use of various types of knowledge applied for solving complex problems is considered. To solve the problem was proposed the knowledge base and methods and tools for developing knowledge bases. Feature of this approach is the coordinated development of knowledge bases of developers, as well as the use of reusable components.**

***Keywords*—semantic network; knowledge base; knowledge base structuring; coordinated development; intelligent system**

I. Introduction

The evolution of information technologies has led to the accumulation of large volumes of heterogeneous, weakly structured information. For effective use of this information to solve various probelms, it is necessary to create knowledge bases, which allow to systematize the knowledge stored in them, and also provide an opportunity to work effectively with them for both computer system and human. The emergence of knowledge bases led to the emergence of specialized models and knowledge representation languages, as well as software tools that allow to develop knowledge bases, including distributed development teams. The volume and heterogeneity of the information stored in the knowledge base leads to the need for its structuring, i.e. the allocation of interrelated fragments in it in order to improve the efficiency of its processing, as well as for didactic purposes.

The knowledge base is a key component of such class of computer systems, as knowledge-based systems, the development of which is one of the promising areas in the field of artificial intelligence [1]. The quality of the developed systems of this class is determined, inter alia, by the quality of the knowledge base and by the variety of types of knowledge stored in it. Knowledge-based systems are now used in a wide variety of areas of human activity - medicine, automation and production management, training, design automation tools and many others [2].

In this article, problems in the knowledge bases development from the perspective of providing the learnability of knowledge-based systems as a necessary property of intelligent systems will be considered. To solve these problems, a semantic model of the knowledge base is proposed that provides the coordinated use of different types of knowledge and models of their representation, the ability to present multilevel meta-knowledge and the knowledge base structuring, as well as the method and tools for knowledge bases development based on this model.

II. Problems in the field of knowledge engineering

The value and quality of the processed information is determined by the possibility of its effective use in a variety of models of problem solving. In turn, the problem solving models themselves should be able to use any information stored in the knowledge base, which can be useful for solving each specific problem.

Expanding the scope of knowledge-based systems has led to the need to support the solution of complex problems. Under the complex problem we will understand the problem, the solution of which involves the use of formalized knowledge of various types and different models of their processing, which in turn requires the compatibility and integration of the knowledge used, as well as models for their processing.

Examples of complex tasks include the following:

- the task of understanding of the meaning of the text of a natural language;
- the task of understanding of the meaning of the handwritten natural language text;
- the task of understanding of the meaning of a voice message;
- the task of semantic analysis of the image;
- the task of automating adaptive learning;
- the task of planning behavior in intelligent robots;
- the task of complex automation of various enterprises.

Nowadays the main approach to the implementation of systems capable of complex problems solving is the development of hybrid systems, one of the main problems in the development of which is the collective use of different types of knowledge in the interests of compensating for deficiencies and combining the advantages of heterogeneous models to solve the problem [3]. In addition, the use of various types of knowledge in problems solving allows to increase the number of classes of problems solved in comparison with the number of classes of problems solved by systems using a limited set of knowledge types. Such an increase is achieved through the combination of different approaches in solving the subproblems within the problem, which becomes possible due to the unification of different types of knowledge representation within the same system. The relevance of this problem in the case of databases is considered in [4]. As a consequence, the actual task is to develop a common unified formal basis for different types of knowledge representation within a single system and ensuring their collective use in complex problems solving. This work focuses on the problem of ensuring the compatibility of various types of knowledge, including metaknowledge, which is proposed to be solved by unifying the representation of different types of knowledge within the same knowledge base [5], [6], [7], [8], [9].

The work is devoted to solving of problems related to the development of models, methods and tools for creating knowledge bases of computer systems capable of complex problems solving.

Let's consider in more detail a complex problem on an example of a problem of automation of the batch production enterprise.

Complex automation of the enterprise requires the following information in the appropriate system:

- equipment nomenclature and configuration (to solve the problems of the current situation analyzing and the planned changes simulating, to evaluate the efficiency of the current equipment configuration, to analyze the possibility of its optimization);
- generalized business processes (description of products, stages of production);
- production plans (list and volume of products);
- current business processes (state of specific equipment, current production stage);
- specifications of emergency situations, ways of their elimination;
- current processes related to logistics;
- information on personnel, accounting
- etc.

The description of all listed types of knowledge in one system is necessary because they are closely related to each other and often describe the same objects, but from different aspects. In order for this description to work within the same system, all fragments must be consistent among themselves, in addition, the description should be structured in various aspects. For example, a description of the same department in terms of current production processes, logistics, personnel, emergency situations - different views on the same complex object.

Various kinds of meta-information about the same objects are used to solve various production tasks. For example, in the event of an emergency situation in the workplace, different users may have various questions to the automation system, the answer requires the presence of a different kind of knowledge in the system (Figure 1).

In order for the company to adapt to the market requirements, modern technologies, the situation with supplies and logistics, changes in legislation, it is permanently required to supplement the knowledge base of the automation system with new types of knowledge, including new concept systems (new workshop, new equipment, new type of manufactured products, new laws, etc.) directly in the operation of such a system (Figure 1). In other words, an intelligent system capable of complex problems solving must be *learnable*. By ***learnability*** in this case we will understand the ability of the system to acquire new knowledge and skills in the process of its operation, while preserving the correctness and integrity of the knowledge base. In turn, the learnability causes the requirement of ***flexibility*** to such a knowledge base, that is, reducing the complexity of making changes to the knowledge base.

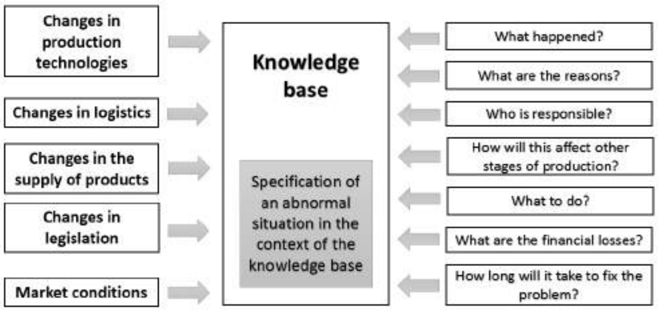


Figure 1. Examples of requests for a production automation system and factors affecting the knowledge base changes

The foregoing allows us to formulate requirements to the knowledge base of systems capable of complex problems solving:

- The possibility of coordinated use of different types of knowledge within the same knowledge base, including, during complex problems solving.
- The knowledge base must have a structure that takes into account the different aspects of the described entities specification.
- Possibility of description in the knowledge base of metaknowledge (systematization and structuring of the knowledge base requires advanced tools of meta-description, i.e. tools of transition from the description of knowledge to the description of metaknowledge).
- convenience of the knowledge base processing, which implies, among other things, the possibility of narrowing and expanding the scope of problem solving, if necessary, as well as the ability to take into account metainformation of various kinds during processing.
- flexibility of the knowledge base, i.e. the possibility of adding to the knowledge base of new fragments, including new types of knowledge, without making changes

to the existing knowledge base structure directly in the process of system operation.
- in order for the system to be able to analyze and, among other things, verify and optimize its own knowledge base and processes of its development, the structure of this knowledge base, the history of its changes and plans for its evolution should be described by the same means as the domain part of knowledge base. The ability of the system to analyze its own components, in particular, the knowledge base, will be called **reflexivity**.

The creation and evolution of a knowledge base that meets these requirements is a labour intensive and time-consuming process that requires a high level of knowledge engineers experience, which entails a deficiency of specialists in the field of knowledge engineering and a high cost of knowledge bases and, consequently, knowledge-based systems.

Taking into account the requirements for knowledge bases, we will also formulate the requirements for the technologies for their development:

- provide the ability to create, using the specified models, tools and methods, knowledge bases that meet the above requirements and are compatible with each other;
- reduce labor intensity and shorten the development time of the knowledge bases;
- ensure high rates of evolution of models, tools and methods for the knowledge bases development;
- reduce the requirements for knowledge bases developers by providing information support.

To date, theoretical studies in the development of formal knowledge representation models and formal models for problems solving have led to a wide variety of such models.

Nowadays, there are dozens of knowledge representation models, but most of them are based on the main four: semantic networks, frames, production and logical models [10].

Each of the described models has its advantages and disadvantages, and when developing knowledge bases, the knowledge engineer must make a choice in favor of the most suitable model, because different types of knowledge require the use of different models of knowledge representation. The existing tools for knowledge bases development, as a rule, are focused on the use of one of the above-mentioned models. However, when creating knowledge-based systems, especially when solving complex problems, it becomes necessary to represent different types of knowledge within the same knowledge base, especially knowledge at the intersection of sciences, which none of the above-mentioned models taken separately can provide [7], [6]. In addition, the choice of the knowledge representation model limits a variety of problem solving models that can be used in the developed system.

Due to that, there is a need to create a universal model of knowledge representation that would make it possible to represent any kind of knowledge in a unified, easy-to-process form. Approaches to solving this problem are considered in the papers [6], [11], [12].

However, for coordinated use of the different types of knowledge in solving complex problems, it is not enough to develop only a knowledge representation model itself, it is necessary to develop unified methods of representation for each kind of knowledge that are agreed upon within the chosen model of knowledge representation.

Each knowledge representation model corresponds to a set of knowledge representation languages that implement these models.

In the knowledge representation languages, as a rule, the syntactic and semantic components are divided. The syntax specifies the rules by which the construction of the given language are built, and the semantics determines the rules for the specified constructions interpreting.

At the moment, different languages for knowledge representation are used: CycL [13] (the knowledge representation language based on ontologies and used within the Cyc project), IDEF5 [14] (Integrated Definitions for Ontology Description Capture Method - an ontological research standard for visualizing data, obtained as a result of processing ontological queries in a simple natural graphical form), Prolog [15] (predicate based language of mathematical logic of Horn clauses, which is a subset of the first order predicates logic), CLIPS [16](C Language Integrated Production System - a language of knowledge representation based on logical rules that use the same name program shell for expert systems creating), and others.

Particular attention should be paid to the tools offered in the scope of the Semantic Web, because of their thoroughness and prevalence. This direction is actively developed by the W3C consortium, whose main task is the development of standards for the Semantic Web [17], [18].

The means of the Semantic Web approach are a set of methods and technologies designed to present information in a form suitable for machine processing. Information is presented in the form of a semantic network, specified by means of ontologies. Standardization of information representation allows a computer system to obtain various factual information and make logical conclusions based on it. The use of W3C standards in the development of intelligent applications in recent decades has become very popular.

As part of the Semantic Web for the presentation of knowledge, the following have been developed:

- Resource Description Framework (RDF) and languages that provide RDF data representation;
- RDF Schema metadata presentation tools;
- principles of representation of knowledge in the form of ontologies and ontology description languages (OWL Lite, OWL DL, OWL Full);
- the SPARQL query language for RDF data storing;
- and a number of other standards.

Efficient knowledge storages based on RDF are also developed, such as Sesame, HyperGraphDB, Neo4j, Virtuoso, AllegroGraph, which provide storage and access to data using the SPARQL query language. To edit ontologies created on the basis of W3C standards, a large number of editors have been created that have a fairly wide functionality [19].

The knowledge representation tools offered in the Semantic Web approach have some disadvantages, one of which is their limitations, in particular, the lack of the ability to describe metacommunications, the means of fuzzy knowledge describing, the inability to describe the properties of entire entity classes, the impossibility of exceptions to certain rules describing and other [20].

Despite the successes achieved in the field of creating knowledge bases, there are the following problems:

- laboriousness of simultaneous use of models representing various types of knowledge;
- incompatibility of already developed components of knowledge bases leads to the need to re-develop existing solutions;
- changes to the knowledge base may necessitate significant changes in the structure of the knowledge base, especially in the case of dynamic knowledge bases;
- despite the existence of sufficiently developed tools of knowledge bases development, they do not fully provide comprehensive support (including information) to the team of developers at all stages of the knowledge base development, and also do not have sufficient flexibility and extensibility;
- existing facilities are oriented, as a rule, to a specific knowledge storage format, which makes it difficult to transfer the already developed knowledge base to another model interpretation platform.

The main reason for all these problems is the lack of unification of the representation of various types of knowledge, including meta-knowledge, within the same knowledge base, and the lack of a knowledge base structuring model that allows to structure the knowledge base on various aspects simultaneously.

III. The proposed approach

Among the approaches to solving these problems are the following:

- development of models to represent different types of knowledge, development of standards for knowledge representation;
- development of knowledge base structuring models that would allow to separate independent fragments of the knowledge base in such a way as to minimize the number of approvals in the collective development of knowledge bases and, as a result, to reduce the complexity of developing and improving of the knowledge base;
- ensuring the compatibility of the already developed knowledge base components for their reuse, and the creation of libraries of such components in order to reduce labor costs and shorten the development of knowledge bases;
- use of automation and information support tools for developers, focused on the collective development of knowledge bases, providing integrated support for all development stages to reduce labor costs, shorten development time, and reduce requirements for knowledge base developers.

However, existing implementations of these approaches are usually aimed at solving of any one of these problems and do not take into account the need to solve all of these problems in a complex.

The approach proposed in this paper is based on the idea of systems constructing based on semantic networks, proposed in [21]. These ideas served as the basis for the creation of ***OSTIS Technology***, which is a complex of models, tools and methods for intelligent systems development, as well as for the permanent updating and improvement of this technology.

OSTIS technology is based on using as a method of encoding information of unified semantic networks with a basic set-theoretic interpretation of their elements. This method is called ***SC-code*** (*Semantic Code*), and semantic networks represented in *SC-code* are called *sc-graphs* (*sc-texts* or texts of SC-code). Elements of such networks, represented in the SC-code, will be called *sc-elements*, in turn, the nodes of such networks will be called *sc-nodes*, the connections between them are *sc-connectors* (sc-arcs, sc-edges).

A key feature of *SC-code* is the joint use of the mathematical apparatus of graph theory and set theory. This allows, on the one hand, to ensure the strict and versatility of formalization tools, on the other hand - to ensure the convenience of storing and processing of information presented in this form.

For visualization of *SC-code* texts, external languages such as *SCg* (Semantic Code graphical), *SCn* (Semantic Code natural) and *SCs* (Semantic Code string) are used [**?**].

The model of an entity recorded by *SC-code* means is called the *sc-model of the specified entity*.

One of the features of the *sc-model of the knowledge base* is the lack of synonymy between sc-elements, which should be identified and eliminated by gluing together synonymous sc-elements, or by explicitly indicating the fact of their synonymy. [12]

To solve the problem of coordinated use of different types of knowledge within the knowledge base, compatibility and flexibility of knowledge bases, as well as high labor costs and long terms of their creation, it is proposed to develop:

- A unified semantic model of the knowledge base that ensures the unification of the representation of various types of knowledge and the possibility of using a wide range of knowledge base structuring types through:
  - internal representation of the knowledge base in the memory of the intelligent system in the form of a semantic representation using formalized semantic network;
  - separation of a hierarchical system of subject domains and an explicit representation of ontologies that describe the semantics of all subject domains and corresponding languages in the knowledge base.

  Use of these principles will ensure the possibility of coordinated use of different types of knowledge, the compatibility of the knowledge bases being developed among themselves, as well as their flexibility.

- The method of coordinated development of knowledge bases built on the specified model, based on the formal ontology of the design actions of knowledge base developers, implementing a component approach and oriented towards the collective development of knowledge bases. The presence of such a method will ensure the correctness and consistency of the project activity of developers directly in the process of the knowledge base use.
- a library of reusable components of knowledge bases, including components search tools based on their specifications. The presence of such a library will significantly reduce the labor costs for the development of knowledge bases.
- tools of automating the activity of knowledge base developers, as well as their information support, including the knowledge base development automation system and the developer consulting services subsystem within the IMS metasystem that implement the proposed method and provides coordination, verification and editing of knowledge base fragments directly in the process of using it. The availability of such tools will reduce the time required to create knowledge bases and requirements for their developers. These tools are expected to be developed using OSTIS Technology, which in turn will ensure flexibility of the tools themselves.

IV. Semantic model of the knowledge base

To implement the proposed approach, it is necessary:

1) to provide a unified basis for the presentation of various types of knowledge, which involves the development of entity specification tools in the knowledge base within a certain ontology of representation;
2) to provide the possibility of an unrestricted transition from knowledge to meta-knowledge due to the possibility of allocating of some integral fragment of the knowledge base and considering it as a specification object;
3) to construct a formal model of knowledge, identify the typology of knowledge;
4) to construct formal models of the most important types of knowledge, which are the basis for the specification of different kinds of entities.

From a formal point of view, the sc-model of the knowledge base is a set of sc-elements.

In order to transform a different kind of knowledge stored in the memory of a computer system into a single well-structured knowledge base, it is necessary to bring all these diverse types of knowledge to a *common syntactic and semantic foundation* based on some *universal ontology of representation*. In this paper, the ***ontology of sc-elements*** is the role of such an ontology of the representation, within the framework of which the typology of entities described in the knowledge base is specified, as well as the typology of signs included in the knowledge base, which reflects the character of the relationship of these signs with the current state of the knowledge base. An example of the specification of entities using the *ontology of sc-elements* in the knowledge base is shown in Figure 2.

The development of such an ontology allows us to describe the syntactic and semantic properties of *sc-elements* (that is, the signs of entities described in the knowledge base) within a single knowledge base, which in turn allows us to provide a property of *reflexivity* for systems based on the proposed approach. In particular, the ontology under consideration allows us to describe the properties of not only the objects of the external world, but also of the internal signs (*sc-elements*) themselves. For example, when describing objects in a dynamic domain, it is necessary to describe, on the one hand, the temporal properties of the entities themselves (*a past entity*, *a present entity*, *a future entity*), on the other hand, the temporal properties of the signs relative to the current state of the knowledge base (*a sign represented in the current state of the knowledge base*, *a sign not represented in the current state of the knowledge base*, etc.).

Existing approaches to the development of knowledge bases are based on the examination of specific elements of the knowledge base (classes, instances, relations, etc.) as objects of the specification. However, with the accumulation of large amounts of information in the knowledge base, it becomes necessary to allocate entire fragments of the knowledge base and be able to specify them, treating them as separate entities. This is necessary to ensure the possibility of an unrestricted transition from knowledge to meta-knowledge. In the framework of this paper, such a fragment of the knowledge base is called a ***structure*** (*sc-structure*). Each ***structure*** is a *sc-element*, denoting some text of *SC-code*, which can later be a specification object, including being part of other structures, be connected with other entities by different relations.

From a formal point of view, the *structure* is treated as a set, the elements of which are all the *sc-elements* that make up the fragment of the knowledge base designated by the given structure.

For the specification in the knowledge base of various structures, their typology was developed and the roles of the elements that make up the structure were determined.

The fact that as a formal basis for representation of knowledge in the *SC-code* the theory of graphs and set theory are used, allows us to analyze not only the external connections of the considered fragment with other elements of the knowledge base, but also to analyze the internal structure of these fragments with the necessary degree of detail, i.e. identify in the knowledge base the analogies, similarities, differences, build different types of correspondence between fragments.

In one of the *SC-code* external representation languages - the *SCg* - *structure* can be represented by explicitly indicating all the pairs of elements membership to the *structure* (see Figure 3a), and also as a contour containing all the elements that make up the structure (see Figure 3b).

The concept of *structure* is a formal basis for the *semantic*

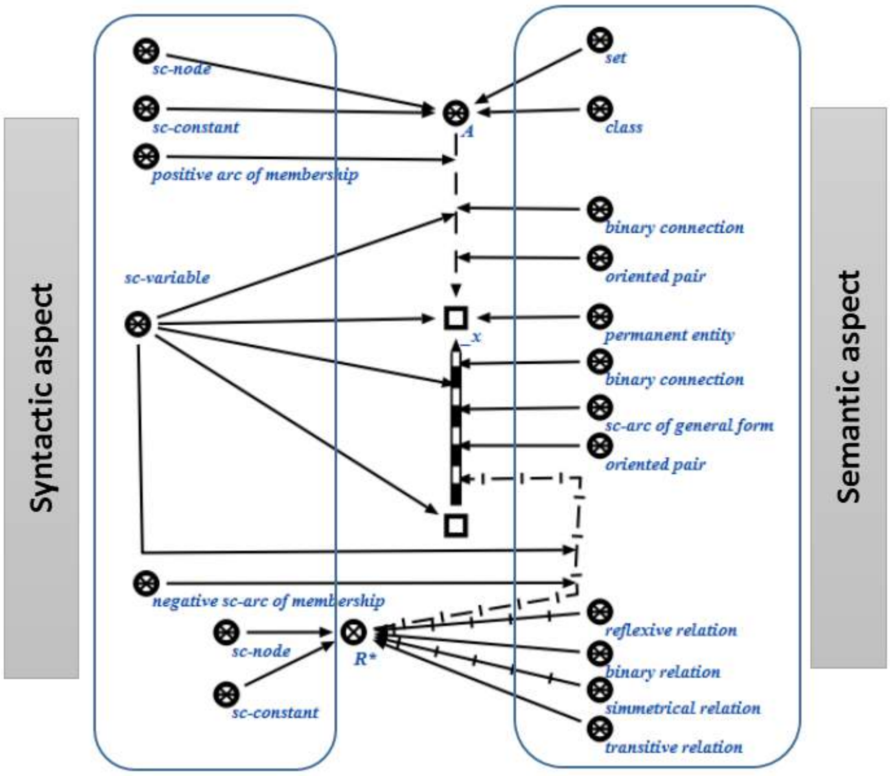


Figure 2. Syntactic and semantic aspects of the description of entities stored in the knowledge base

model of the knowledge base and models for its structuring in various aspects. In the works [22], [23], various relations defined on the structures are considered.

On the basis of the concept of structure, the concept of ***knowledge*** as the most important kind of entities described in the knowledge base is clarified.

Within the proposed knowledge base model, each *structure* will be called ***knowledge*** if and only if:

- for any connection (including the *sc-connector*) that is part of this structure, this structure includes all components of this connection with explicit indication of the membership of the specified components in the specified connection;
- for any sign of the final structure included in the structure under consideration, this structure also includes all components of this finite structure with explicit indication of the membership of these components in the specified final structure;
- for any sc-element that is part of the structure under consideration, this structure includes the signs of all concepts for which the indicated sc-element is a member, with explicit indication of this membership.

The most commonly used types of knowledge within the proposed knowledge base model are the *semantic neighbourhood*, *factual knowledge*, *comparison*, *knowledge base section*, *subject domain*, *ontology*, *task*, *program*, *plan*, *solution*, *statement*, *definition*, *reasoning*, etc.

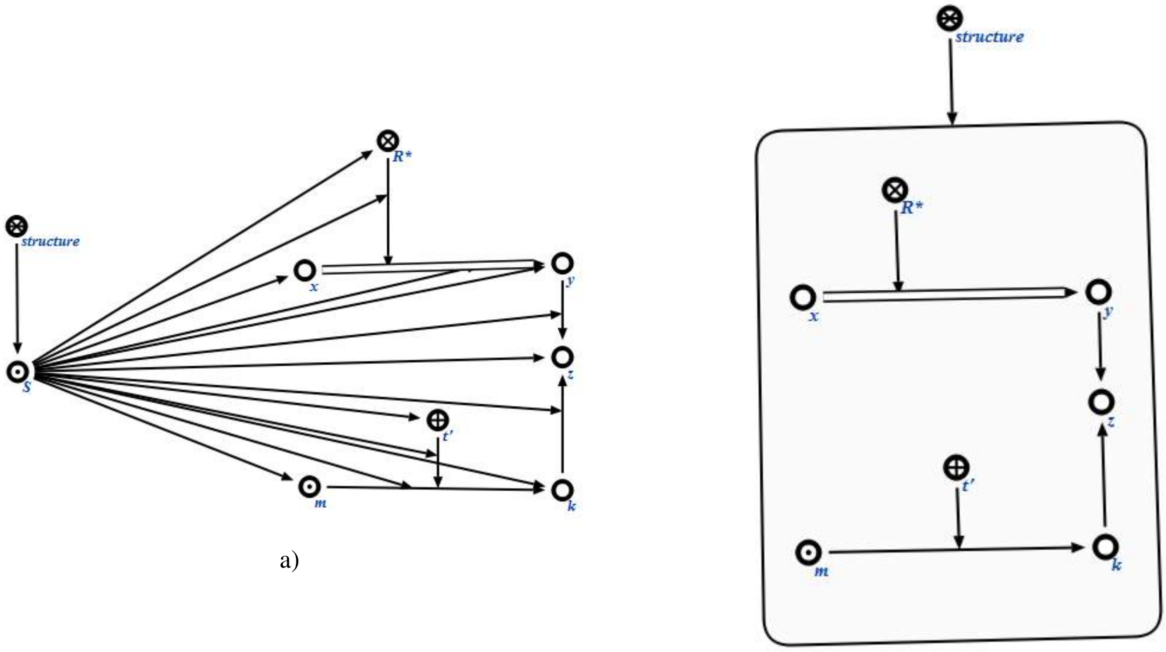
The formal refinement of the concepts of *structure* and *knowledge* makes it possible to provide the possibility of an unrestricted transition from knowledge to the corresponding *metaknowledge*. This property is achieved due to the ability to designate some fragment of the knowledge base with one sc-element, and, accordingly, consider it as a single whole, specify its properties and connections with other fragments. In turn, such a specification can be further considered as an object of formal description (Figure 4).

For the specification of individual entities within the knowledge base, the concept of a ***semantic neighbourhood*** is introduced.

A ***semantic neighbourhood*** is a specification of some entity for a specific set of characteristics, which is essentially a collection of metainformation. Formally, the set of such attributes is determined by the set of relations and classes to which the described entity belongs and is the basis for semantic neighbourhoods classifying.

In general, a set of characteristics specifying entities belonging to certain classes will be different. In addition, it often becomes necessary to specify the same entity in various aspects and explicitly capture these aspects in the knowledge base.

For example, some person can be described from a professional (job, position, professional skills), medical (sex, weight, height, illness), civil (family status, nationality, age, attitude



a)

b)

Figure 3. An example of a complete and abbreviated representation of a structure in the SCg-code

to military service), etc. points of view (see Figure 5). The ability to describe the various properties of the same object is achieved through the allocation of various classes of semantic neighbourhoods and the selection of a set of characteristics that determine a particular class of semantic neighbourhoods.

An example of a semantic neighbourhood is shown in Figure 5.

The most important types of knowledge in terms of the process of knowledge bases development are ***subject domains*** and ***ontologies***. In the framework of the proposed approach, they are the basis for creating formal means of describing different types of knowledge in the knowledge base.

The concept of the *subject domain* is the most important in the field of knowledge engineering, allowing you to focus on a particular class of entities being researched and a specific family of relations that are specified on the specified class. Thus, abstraction is carried out from the rest of the investigated world.

Consideration of the structure of the knowledge base in relation to the *subject domain* allows us to consider the objects under study at different levels of detail. Detailed analysis of the researched objects can be carried out both within the original subject domains and within a system of independent, but related subject domains.

Each subject domain focuses on the description of the relations of the corresponding class of researched objects. Formally, the subject domain model is defined as follows:

$$M_{SD} = (SD_N, SD_E, SD_R, SD_O), \qquad (1)$$

where SD_N – non-empty set of elements of the subject domain (carrier);
SD_E – set of roles of elements within the subject domain;
SD_R – set of relations between subject domains;
SD_O – set of classes of ontologies specifying subject domains.

An example of a formal specification of a subject domain is shown in Figure 6.

For a formal specification (description of properties) of the relevant subject domain, oriented to describing the properties and interrelations of the concepts that make up the specified subject domain, such type of knowledge is used as ***ontology***.

Within the framework of the proposed approach, the concept of *ontology* is refined and their typology based on the selected typology of *semantic neighbourhoods* is introduced. This approach to the allocation of ontology classes is based on the approach to the classification of ontologies, depending on the set of the relations used to describe the entity properties.

From the formal point of view, within the framework of the proposed knowledge base model, ***ontology*** will be interpreted as the result of the set-theoretical union of semantic neighbourhoods of one type. Depending on the properties of the subject domain concepts under consideration, which are described in the ontology, i.e. type of unioned semantic neighbourhoods, the following types of ontologies distinguished: *structural specification of the subject domain*, *set-theoretic ontology*, *logical ontology*, *terminological ontology*, *ontology of problems and problems solutions*, *ontology of classes of problems and methods for problems solving*, etc. In more detail this typology was considered in [22], [23].

Explicit allocation of *ontologies* in knowledge bases of

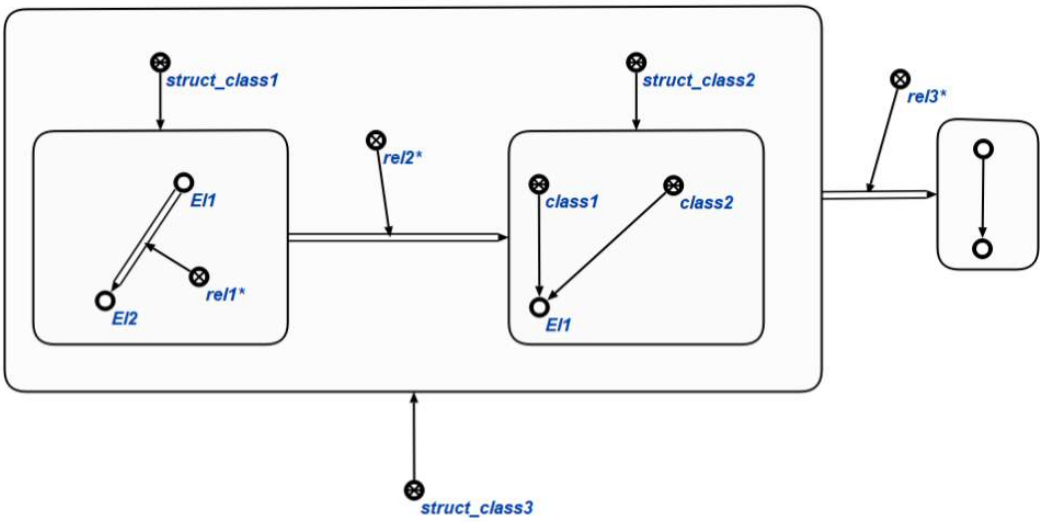


Figure 4. Example of metaknowledge description in SC-code

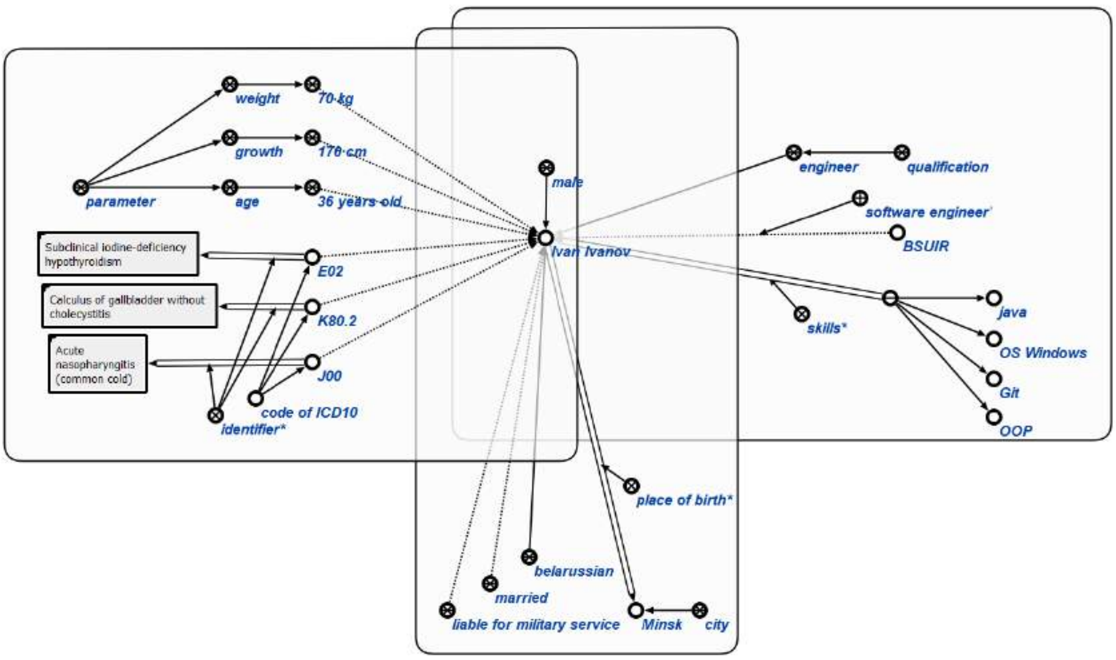


Figure 5. Example of the description of the semantic neighbourhood in the SC code

intelligent systems is necessary in order to:

- commit the agreed current version of the interpretation (clarification) of all the concepts used;
- ensure a clear organisation of the continuous process of development and harmonisation of the system of concepts used. This, in turn, requires a fairly detailed documentation (logging) of all changes in the system of concepts.

The main goal of structuring is to divide the specification of the described world into parts. This allows us to abstract each of the parts from those details that are not essential for solving the current problem, by analogy with experiments in various natural sciences.

One of the structuring way, which makes it possible to localise the domain of finding ways to solve problems, is the structuring based on the hierarchy of subject domains and their ontologies. In this case, the search area for solving a problem can be one or more subject domains that are sufficient to solve a given class of problems. At the same time, if necessary, the scope of the problem solution search can expand up to the

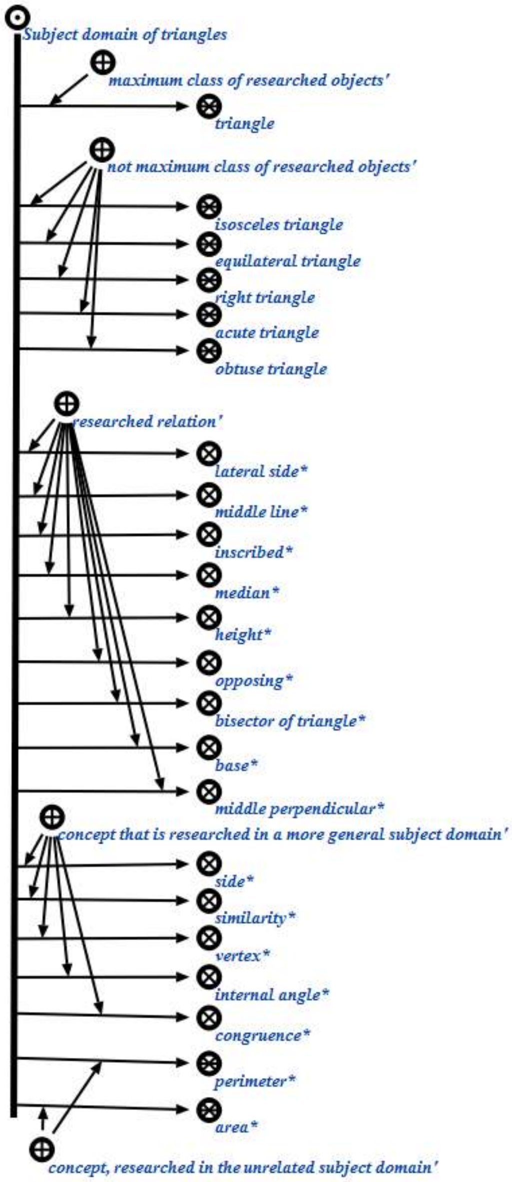


Figure 6. The example of a subject domain specification in the SC code

whole knowledge base.

Another important structuring way is the structuring of the knowledge base from the perspective of the development process (Figure 7), within which sections are identified that describe the part of the knowledge base available to the end user and parts of the knowledge base available to different categories of developers.

In addition, in the knowledge base it is necessary to describe the dynamics of the specified domain itself as a model of the external world fragment (i.e., the description of the past, present and future state of various external entities), and the dynamics of the knowledge base itself (i.e., the history of its development, plans for completion, current state), which is also one of the ways of the knowledge base structuring (Figure 7).

Using the models discussed earlier, a ***formal model of the knowledge bases structuring*** has been developed:

$$M_{STR} = (STR_S, STR_C, STR_R), \quad (2)$$

where STR_S – set of knowledge base sections;
STR_C – set of allocated classes of knowledge base sections, determined by a set of structuring characteristics;
STR_R – set of relations specifying the knowledge base sections, including - decomposition of sections into subsections.

This model is considered in more detail in [22]. An example of the structuring of the knowledge base on various aspects is shown in Figure 8.

The proposed model of structuring based on the allocation of sections and the formation of their hierarchy makes it possible to structure the knowledge base on the basis of an arbitrary set of characteristics, i.e. structuring the knowledge base from different points of view, while combining them within a single knowledge base.

An important feature of this approach is the description of the entire structure of the knowledge base with the help of SC-code in the same knowledge base, which in turn ensures the reflexivity of the intelligent system, that is, the ability to analyse the structure of its own knowledge base, for example, to identify various contradictions.

On the basis of the above results, a ***semantic model of the knowledge base*** is constructed that satisfies the above requirements:

$$M_{KB} = (M_S, ONT_R, ONT_{HL}, \{M_{STR1}, M_{STR2}, ...M_{STRn}\}), \quad (3)$$

M_S – set of structures stored in the knowledge base;
ONT_R – the ontology of sc-elements, which is the ontology of representation in the framework of the proposed approach;
$ONT_{HL} = \{ONT_{STR}, ONT_K, ONT_{SN}, ONT_{SD}, ONT_O\}$ – set of top-level ontologies,
ONT_{STR} – the ontology of the Subject domain of structures,
ONT_K – the ontology of the Subject domain of knowledge,
ONT_{SN} – the ontology of the Subject domain of semantic neighbourhoods
ONT_{SD} – ontology of the Subject domain of subject domains
ONT_O – ontology of the Subject domain of ontologies,
M_{STRi} – model of the knowledge base structuring with the i-th way.

In accordance with the proposed model, adding a new entity to the knowledge base requires specification of this entity using one or more concepts from the ontology of representation or the set of top-level ontologies. The specification of the entity in this case means the indication of the inclusion or belonging to any of the classes of the mentioned ontologies.

In turn, the addition of a new kind of knowledge is reduced to:

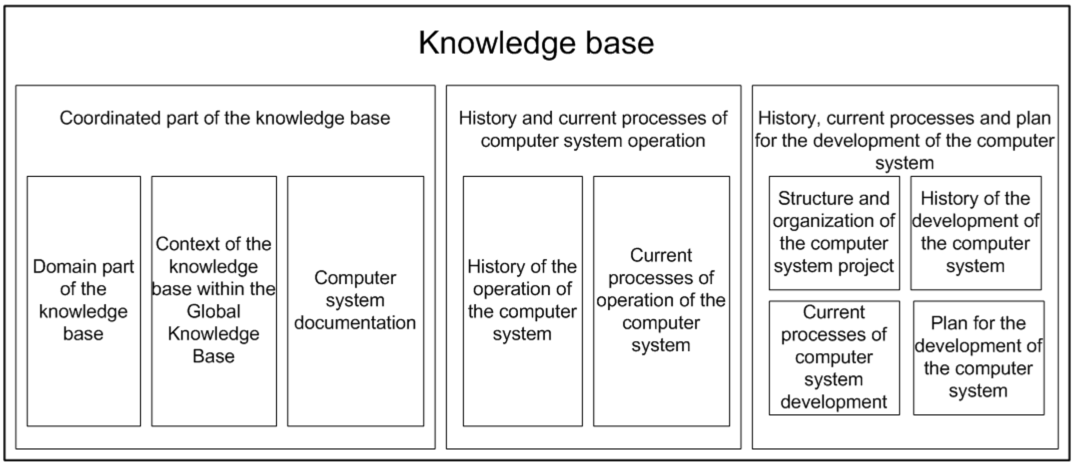


Figure 7. Structuring the knowledge base in terms of the development process

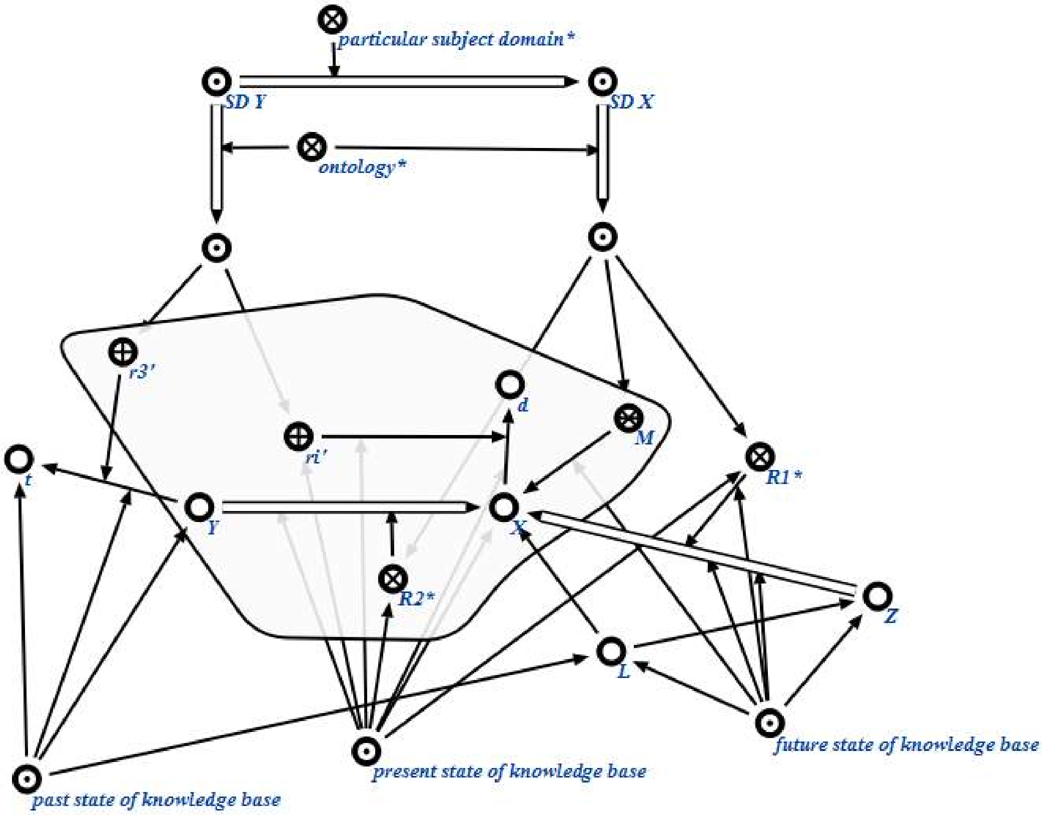


Figure 8. An example of different structuring ways combining in one knowledge base, represented in SC-code

- construction of a model of the relevant subject domain (or several subject domains), within the framework of which the introduced concepts are reseached - i.e. constructing of a structural specification of the subject domain, including indicating the relations of the considered subject domain with others already existing in the knowledge base model;
- building a family of ontologies of various types, specifying the developed subject domains.

The knowledge bases developed with the help of the proposed model have the following main advantages:

- the consistent use of different types of knowledge by using as a formal basis for information encoding of unified semantic networks with a basic set-theoretic interpretation, the syntax and semantics of which are specified within the *ontology of sc-elements*, which is the ontology of representation in the framework of the proposed approach, and use of the developed set of *top-level ontologies* (ontology of the subject domain of structures, ontology of the subject domain of sets, ontology of subject domain of connections and relations, ontology of subject domain of subject domains, etc.), specifying the most common kinds of knowledge, on the basis of which other types of knowledge may be described;
- the possibility of structuring from the point of view of various aspects and description in the knowledge base of meta-knowledge through the use of a *structure model* that allows to consider and specify a fragment of the knowledge base as a single whole, which in turn enables the transition from knowledge to meta-knowledge to an unlimited number of levels;
- convenience of the knowledge base processing due to the possibility of metainformation use during the processing, as well as the possibility of localising the area of finding ways to solve problems within one or several subject domains explicitly specified in the knowledge base;
- the flexibility of the knowledge base, i.e. its ease of modification, due to the explicit specification in the knowledge base of subject domains and their ontologies, which allows you to localise those fragments of the knowledge base in the process of the knowledge base developing and changing, which will also need to be changed.
- the ability to analyse, including, verify, adjust and optimise the structure of the knowledge base, history and plans for its development by the same means as the domain part of the knowledge base by using unified tools to describe all the listed information. At the same time, the possibilities of such an analysis can be easily extended by specifying in the knowledge base the information needed for analysis, for example, the specification of incorrect structures classes.

The principles of representation of various types of knowledge, using the ontology of sc-elements and a set of ontologies of the upper level, developed within the framework of the *knowledge base semantic model* allowed to solve a number of problems related to the presentation of knowledge within the Semantic Web, in particular, using the standards RDF and OWL 2. Let us consider several examples in more detail.

Example 1. In RDF and OWL 2, the principle of set normalisation is not used, which is one of the fundamental principles of the SC-code. The principle of the sets normalisation assumes that each sc-element (including the sc-arc) is treated either as a sign of the terminal entity (for example, the sign of a concrete number or the sign of a concrete material entity) or as a sign of a set whose elements are, in turn, only sc-elements. This principle ensures strict formalisation and unambiguous interpretation of each element of the knowledge base, which in turn increases the convenience of information processing and reduces the number of concepts.

Figure 9 shows an example of using the "color" concept in RDF (Figure 9a) and SCg (Figure 9b). As can be seen from Figure 9a, the specific color (Red) is denoted by the node of the semantic network and is associated with the object being characterised by the relation "have color". When trying to apply the principle of normalization, it becomes obvious that a particular color is not a terminal entity (there is no "red" entity), and should be treated as a set of all elements with this attribute, in this case, having a red color (Figure 9b). Thus, the relation "have color" is redunant. Analogous arguments lead to the fact that the relations "sex", "length" and other relations that describe the properties of objects in a certain parameter space turn out to be redundant. In turn, the use of the sc-arc of membership instead of the arc of the relation "have color" simplifies the processing of such a construction, since the semantics of such an arc is unambiguously interpreted by the processing means.

Example 2. The RDF and OWL2 tools allow you to describe relation properties, but do not allow you to specify individual connections of a specific relation, although sometimes such a need arises. Let us consider the example given on the official website of the W3C consortium [24], in which it is necessary to characterise the degree of certainty when exhibiting a certain diagnosis to the patient. As can be seen from Figure 10a, to record this information by means of RDF, it is necessary to introduce an additional node of the semantic network, denoting a specific diagnosis (in fact - a connection of the "diagnosis" relation) and relations "have diagnosis", "have value", "have probability". In addition, that such a description is cumbersome, the description of the diagnosis in the case where it is necessary to indicate the probability and in the case when it is not necessary to do so is fundamentally different (in the second case, only the "diagnosis" relationship between the patient and the value of the diagnosis is used). From the point of view of the set-theoretic interpretation underlying the SC-code, each relation connection (for example, the sc-arc) is treated as a set of related elements, which in turn is an element of the corresponding relation. This approach allows us to specify both the relation itself and separately each connection of the relation, while adding any specification does not lead to changes in the representation of the original relationship

Figure 9. An example of the "color" concept use

(Figure 10b). In this case, as in the previous example, the relation "have probability" turns out to be redundant.

Example 3. The set-theoretic approach, underlying the SC-code, allows, if necessary, to specify in the knowledge base not only binary connections, which are usually represented by sc-arcs or sc-edges, but also various capacity links, denoted by sc-nodes. However, such sc-nodes do not need to be named, unlike nodes in RDF-triples. Each such link can then be specified in the necessary way.

This approach solves one of the problems described on the W3C website [25], related to the presentation of the fact that some set is the union of a family of pairwise disjoint sets. In the variant of representation on OWL 2 (Figure 11a), the "union" relation is used for this, but this relation cannot be described by RDF triples, in addition, in this case it is necessary to indicate the fact of disjointness for each pair of sets. The SC-code means allow to introduce quasi-binary relations connecting some sc-element and some set (connection) of other sc-elements. For example, to solve the problem under consideration, we can define the relation "subdividing*" connecting some set and a family of pairwise disjoint sets (Figure 11b). A formal definition of this relation through other relations (for example, "union" and "pair of disjoint sets") can also be written in the SC-code, and not be specified separately for each use of this relation.

Next, we will consider some advantages of different types of knowledge representation using the models proposed in this work in relation to the classical models of representation (frames, products, logical models). Figure 12 shows an example of representing of the knowledge base fragment using frames (Figure 12a) and the first-order predicate logic language (Figure 12b) [26].

Figure 13 shows an example of a SC-code construction, semantically equivalent to the fragments shown in Figure 12.

Figure 14 shows an example of formalisation in SC-code of the following production rules:

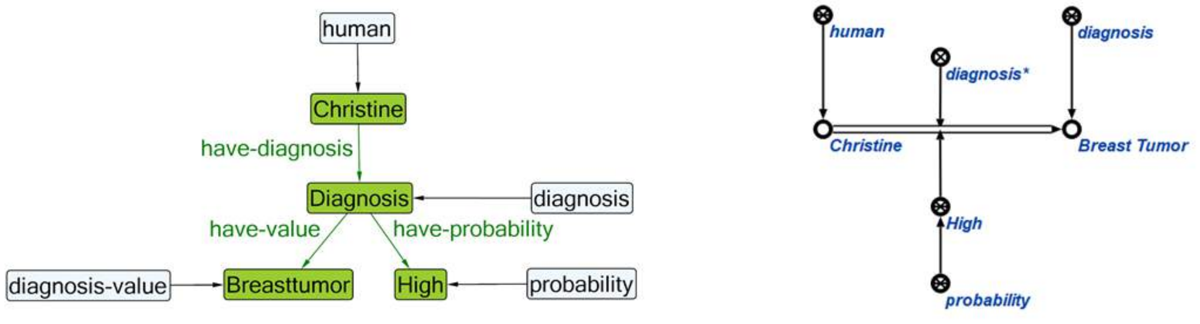
- IF X is a bird, then X is an animal;
- IF X is a bird, then X has two legs;
- IF X is an animal, then X has two legs OR X has four legs.

These examples allow us to formulate the following shortcomings of classical models:

- each classical model is heavily oriented to presenting knowledge of a particular types. So, with the help of frames it is inconvenient to represent logical rules and patterns; the product model is convenient for the presentation of rules, but requires the presence of some additional language for describing the factual information, in addition, the production model does not allow describing complex logical statements containing several quantifiers at different levels; the logical model is oriented to the representation of strict facts and statements, and in the classical version it does not allow to describe knowledge that does not have sufficient completeness, accuracy and correctness;
- the form of presentation of some information in the classical models is largely determined by the syntax of the chosen model, not by the sense of the information presented. So, for example, the presentation of even simple factographic information in the production model obliges the knowledge engineer to formulate all knowledge in the form of rules like «IF ..., THEN ...», although this is not always convenient for the developer and especially for the domain expert;
- the syntax of some classical models in some cases (for example, in the case of frames) is oriented on its perception by a person, not by a machine, in others (for example, in logical models) on the contrary. Thus, classical models do not allow to provide knowledge representation in the form of both human-friendly and convenient for storage and processing by a machine.
- in many cases, classical models (for example, frames) do not have sufficient strictness of representation and allow you to write the same information in different ways, which in the future can lead to incompatibility of fragments of knowledge bases developed by different developers.

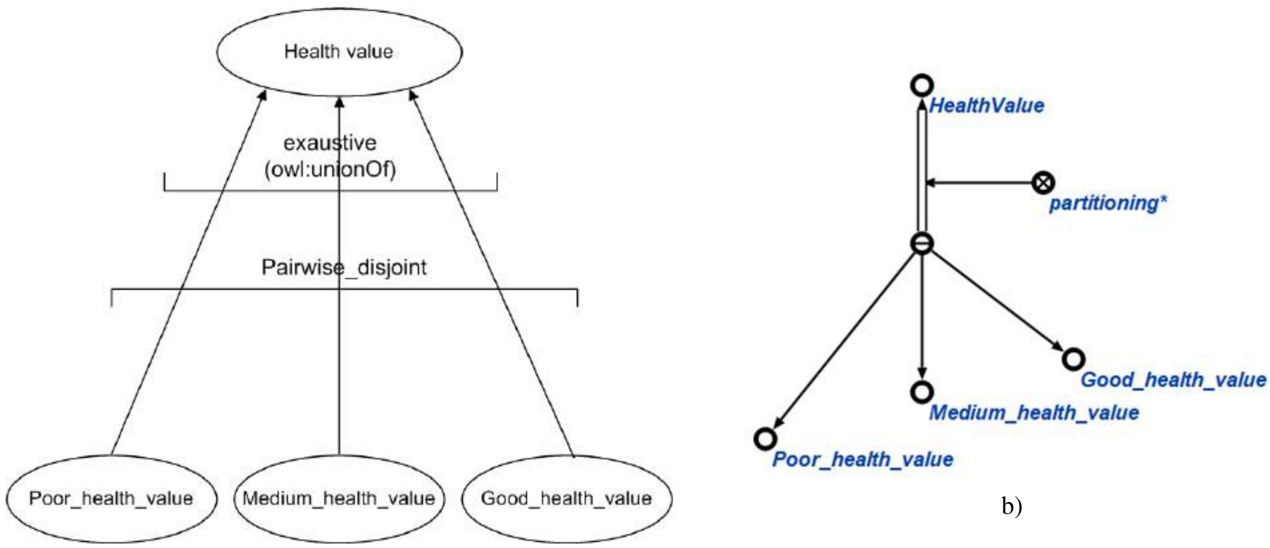
In turn, the models constructed on the basis of the SC-code, proposed in this paper, allow to eliminate the indicated shortcomings, in particular:

- the use of the apparatus of graph theory and set theory as a basis for representation of knowledge allows to ensure, on the one hand, strictness in presentation, on the other hand, presentation visibility and convenience



a) b)

Figure 10. Connection specification example



a) b)

Figure 11. An example of a description of disjoint subsets of a set

of information storing and processing in the computer system memory;

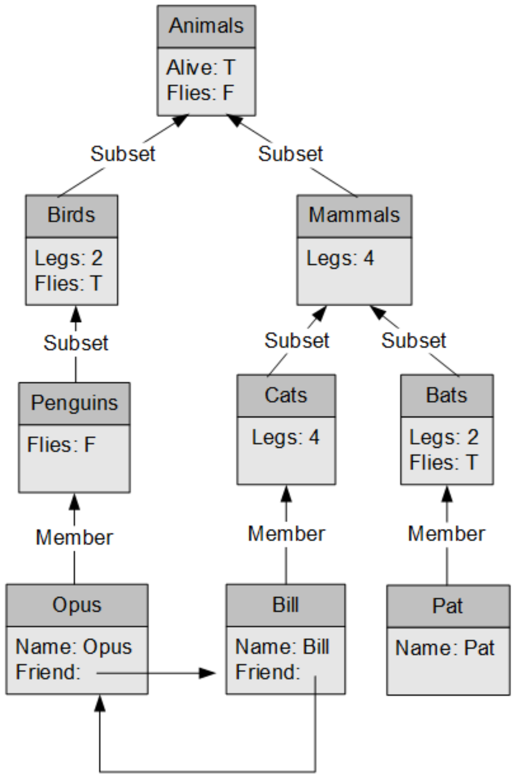
- the proposed models allow to represent different types of knowledge in a unified form (see Figures 13, 14);
- SC-code has a relatively small alphabet and simple syntax, all models built on its basis use the same basic alphabet and syntax, only a set of key nodes is expanded, thus the convenience of storage and processing is not violated. The meaning of the stored information is fixed by the configuration of the connections between sc-elements that are built in accordance with the mentioned syntax;
- The formalism used does not require the developer to bring the information presented to any special kind, determined by the syntactic features of the language. The only condition is the interpretation of all the described entities as particular for some of concepts from the representation ontology considered, based on the formalism of set theory, which in turn ensures the unambiguity of the semantics of each entity described in the knowledge base.

V. Method and tools for development of knowledge bases

Developing a knowledge base is a labour intensive and time-consuming process. Among the ways to reduce the timeframe for creating knowledge bases, the main ones are to ensure the joint development of knowledge bases by a distributed team of developers, automating their activities, and also reusing the already developed knowledge base components. [27], [28], [29], [30], [31], [32], [33], [34], [35], [36].

However, it is impossible to fully automate the process of developing the knowledge base, because some stages, such as the formation of a system of concepts, require the concerted efforts of a number of developers and experts and are subjective.

In addition, in the process of using any knowledge-based system, there is a permanent need to improve its knowledge base: the addition of new knowledge, the removal of irrelevant information, the search for and correction of errors and inaccuracies. The actualisation of information stored in



a)

Rel(Alive,Animal,T)
Rel(Flies,Animal,F)

Birds ⊂ Animals
Mammals ⊂ Animals

Rel(Flies,Birds,T)
Rel(Legs,Birds,2)
Rel(Legs,Mammals,4)

Penguins ⊂ Birds
Cats ⊂ Mammals
Bats ⊂ Mammals
Rel(Flies,Penguins,F)
Rel(Legs,Bats,2)
Rel(Flies,Bats,T)

Opus ∈ Penguins
Bill ∈ Cats
Pat ∈ Bats
Name(Opus,"Opus")
Name(Bill,"BIll")
Friend(Bill,Opus)
Friend(Opus,Bill)
Name(Pat,"Pat")

b)

Figure 12. An example of a knowledge base fragment presented with the help of frame and logic models

the knowledge base in accordance with the current state of the described subject domains (especially dynamic ones) may require such significant changes as the replacement of one conceptual system with another, including the introduction of new concepts, the removal of outdated concepts, the redefinition of existing concepts, etc. Due to the fact that standards, existing in various fields of application of knowledge-based systems, and also requirements and technologies are constantly changing, the process of evolution of the knowledge base must be carried out continuously directly during the operation of such a system [37]. The vast majority of knowledge base development tools do not provide such an opportunity, while strict dividing the processes of developing, improving and maintaining knowledge bases of knowledge-based systems.

A. Library of reusable components

To reduce labour costs in the development of knowledge bases of knowledge-based systems, it is proposed to use already developed fragments of knowledge bases or the whole knowledge bases of any systems. To organise the storage and search of such components, a ***library of knowledge bases reusable components***, which is part of the *IMS metasystem* [**?**], is proposed in this paper.

The *library of knowledge bases reusable components* includes a set of such components, means of such components specification and tools of automating the search for components based on their specifications (Figure 15).

Each ***knowledge bases reusable component*** is a structure either explicitly represented in the current state of the sc-memory or an incompletely formed structure which, if necessary, can be completely formed by combining its parts indicated by any decomposition relation, for example, a subdividing, or the inclusion relation, and which can be used within another knowledge-based system.

Each *knowledge bases reusable component* has a formal specification, that is, some semantic neighbourhood that characterizes this component. On the base of the formal specification, search for a suitable component in the library is carried out, comparation of it with other components, and so on.

The main semantic classes of knowledge bases reusable components stored in the library of knowledge base components are:

- semantic neighbourhoods of different entities;

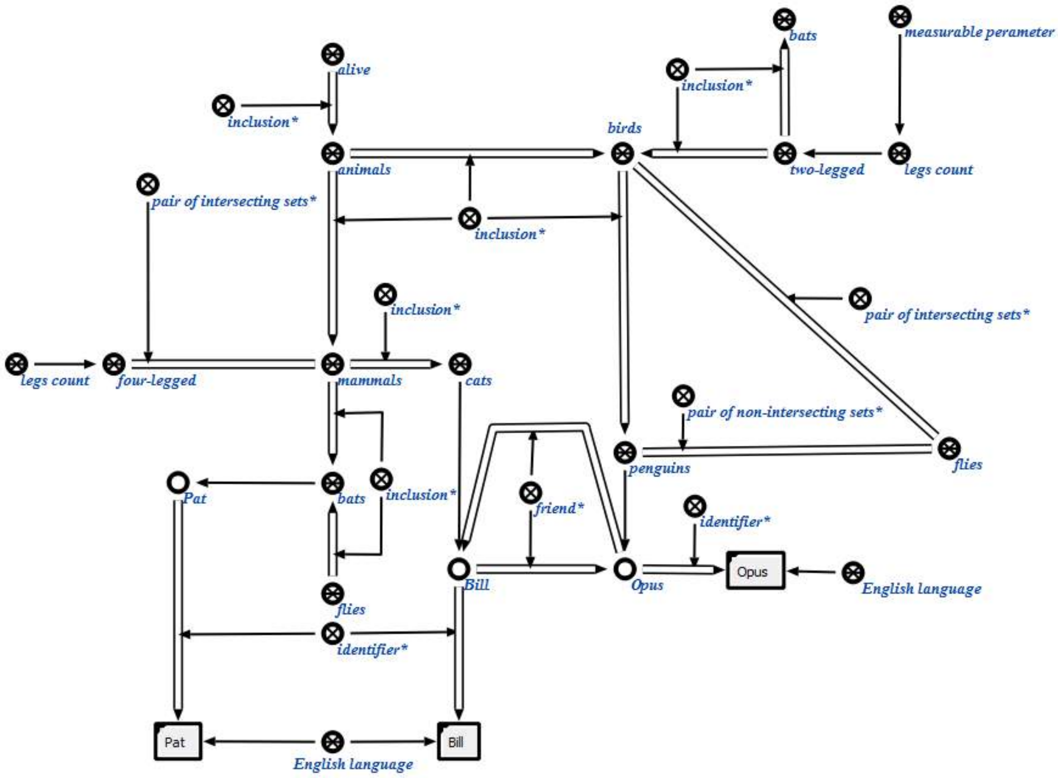


Figure 13. SC-code construction, semantically equivalent to the given examples

- ontologies of different subject domains;
- specifications of formal languages for describing different subject domains;
- sections of the knowledge base of various semantic types (including non-atomic ones);
- knowledge bases of entire subsystems that provide solutions to various tasks;
- and etc.

Integration of a *knowledge bases reusable component* into a system is reduced to merging key nodes by identifiers and eliminating possible duplications and contradictions that could arise if the developer of the system manually made any changes to its knowledge base.

The components search automation tools include tools for finding dependencies between components, searching for components in which the specified concepts are described, as well as searching for components by a fragment of their specification, etc.

More detailed knowledge bases component development based on *OSTIS Technology* is considered in [38].

B. The method of coordinated development of knowledge bases

When developing knowledge bases, it is necessary at every stage to ensure the semantic compatibility of knowledge bases and their components, that is, all concepts used must be treated equally in different components. Especially this task is relevant in terms of collective development, as well as in situations where the system of concepts used is changing. To ensure such compatibility it is necessary to use ontologies, as well as to commit protocols of coordinated changes in the knowledge base.

To solve these problems, a ***method of creating knowledge bases by a team of developers*** based on a formal model of the project activity of various knowledge base developers, each of which can play a certain role in the development process, is proposed. The construction of such a model and the clear allocation of classes of such actions allows to automate the process of collective development, as well as minimize the overhead costs for coordinating the activities of various developers. The main attention in the proposed method is given to the processes of harmonising the interpretation of certain concepts within the framework of the created knowledge base.

This method assumes two main stages - the *creation of a start-up version of the ostis-system* (including its knowledge base) and the stage of the knowledge base development itself.

The process of creating the start-up version of the ostis-system (system, built with OSTIS Technology) can be divided into four main stages:

- selection and installation of an interpretation platform for the ostis-system model;
- installation of the *Core of sc-models of knowledge bases* from the library of knowledge bases reusable components, which contains ontologies of the most common

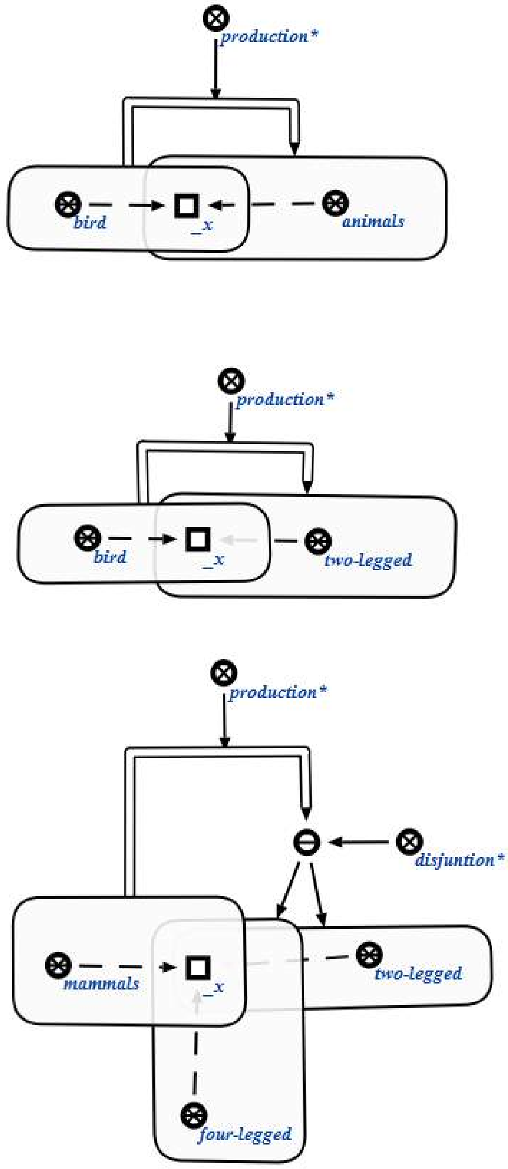


Figure 14. An example of formalisation of productions in SC-code

subject domains, i.e. top-level ontologies (for example, the *Subject domain of numbers and numeric structures*, the *Subject domain of ontologies*, the *Subject domain of logical formulas and logical ontologies*, the *Subject domain of connections and relations*, the *Subject domain of parameters and quantities*, etc.);

- installation of the *Core of the knowledge base processing machine* from the library of the reusable components of the knowledge base processing machines [39], that is, a set of basic reusable components of the knowledge base processing machines required to run the start-up version of the ostis-system;
- installation of the Core of the sc-models of interfaces [40], that is, a set of basic reusable components of the sc-models of interfaces required for running the start-up version of the ostis-system;
- installation of a system for the collective development of knowledge bases support.

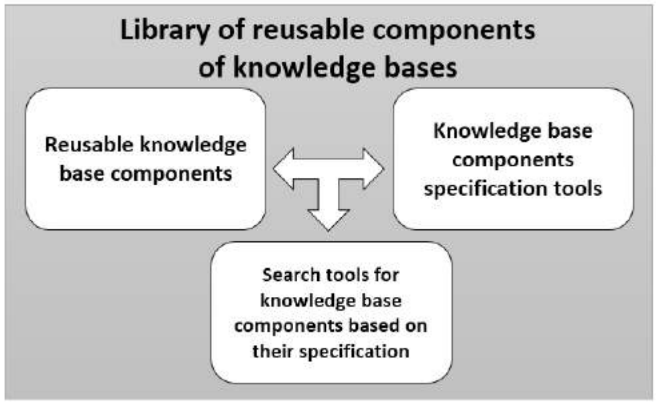


Figure 15. The structure of the library of reusable components of knowledge bases

After the basic configuration of the initial version of the ostis-system is build, the stage of development of the knowledge base begins, which includes the following stages:

1) formation of the initial structure of the knowledge base, which assumes:
   - formation of the structure of knowledge base sections;
   - identification of the described subject domains;
   - construction of a hierarchical system of described subject domains;
   - building a hierarchy of knowledge base sections within the subject domain of the knowledge base, taking into account the hierarchy of subject domains built at the previous stage.
2) identification the components of the knowledge base that can be taken from the library of reusable components of knowledge bases, and their inclusion into the developed knowledge base.
3) formation of project tasks for the development of missing fragments of the knowledge base and distribution of tasks among developers.
4) development and coordination of knowledge base fragments, which, in turn, can be included in the library of knowledge bases reusable components.
5) verification and debugging of the knowledge base.

It should be noted that during the development of the knowledge base, steps 3-5 are performed cyclically.

To ensure the property of the intelligent system's reflexivity, in particular, the possibility of the analysis automating of the history of the knowledge base evolution and generating plans for its development, all activities related to the development of the knowledge base must be specified in the same knowledge base by the same means as the domain part. To solve this problem, a ***formal ontology of the developers activity*** aimed at the development and modification of knowledge bases has been developed. This model is based on the *model of activity of various subjects* proposed in [41].

To organise the coordinated project activity on the creation of knowledge bases within the framework of this ontology, the roles of the participants in the development process (administrator, manager, expert, developer), the classes of actions performed by them, as well as the means for specifying the mentioned actions are introduced. This model is considered in more detail in [22].

The process of creating and editing the knowledge base of the ostis-system comes to the formation of *proposals for editing* a particular section of the knowledge base by developers and the subsequent consideration of these proposals by the knowledge base administrators. In addition, it is assumed that, if necessary, experts can be involved to verify the incoming proposals for editing the knowledge base, and the management of the development process is carried out by the managers of the relevant knowledge-base development projects (see Figure 16). At the same time, the formation of project tasks and their specification is also carried out with the help of the mechanism of proposals for editing the relevant section of the knowledge base. Thus, all information related to the current processes of developing the knowledge base, history and plans for its development is stored in the same knowledge base as its domain part, i.e. the part of the knowledge base available to the end user.

This approach provides wide possibilities for automating the process of knowledge bases development, as well as subsequent analysis and improvement of the knowledge base.

The developed method of coordinated construction and modification of knowledge bases on the basis of a formal ontology of project activity using reusable components allows to ensure the correctness and consistency of the project activity of developers directly in the process of using the knowledge base. At the same time, the discussion and the reconciliation process takes place in the same memory of the computer system where the knowledge base is also stored.

C. Knowledge Base Development Tools

To reduce the complexity of the process of knowledge bases development and reducing requirements for developers, *tools for automating the development of knowledge bases and information support for knowledge base developers* have been built.

The *tools of information support for developers* are implemented in the form of intelligent *metasystem IMS* (Intelligent MetaSystem) [**?**], which is also built using OSTIS technology. At each moment of time, the metasystem contains the models, tools and methods of developing computer systems based on OSTIS technology that have been accumulated and formalized to date. All the models, methods and tools presented in this dissertation work are formally described in the knowledge base of this system. In this way, it is possible to continuously update these results.

Tools of knowledge bases development automation are implemented in the form of a ***system for the knowledge bases collective development support***. An important aspect of supporting the creation of knowledge bases is to support the activity of knowledge base developers directly in the process of operating the system being developed. Thus, the system for supporting the collective development of knowledge bases is built in as a subsystem in each system being developed.

Based on the analysis of similar tools (Protégé [42], Co4 [43], NeON [44], etc.), the following additional requirements were formulated for the functionality of the *system for supporting the collective development of knowledge bases*, taking into account the needs of developers of knowledge bases and the identified shortcomings of the analogues considered:

- providing the possibility of both manual and automatic editing of knowledge bases;
- ensuring the possibility of automatic verification of the knowledge base, including the analysis of the correctness and completeness of the knowledge base;
- ensuring the possibility of creating a knowledge base by a distributed team of developers, including a mechanism for reconciling changes to the knowledge base, as well as a mechanism for storing the history of changes introduced with authority specification.

The implementation of these capabilities implies the refusal to work with the source code of the knowledge base. In this case it is assumed that all changes are made directly to the memory of the system, where the entire knowledge base is stored, which makes it possible to develop the knowledge base of the computer system in the process of its operation.

As it was said before, the system of support of collective development of knowledge bases is constructed as an ostis-system and has the appropriate architecture (figure 17)

The semantic model of the knowledge base of the system of support of collective development of knowledge bases includes sections containing all the knowledge necessary to support the development and evolution of the knowledge base:

- a set of top-level ontologies that are necessary for the functioning of the support system for the process of developing knowledge bases and are the basis for building the knowledge bases of the systems being developed;
- formal ontology of the subject domain of activities aimed at the development and evolution of knowledge bases, including a description of the typology of the knowledge bases developers roles, the classification of developers' actions, as well as formal means of specifying proposals for editing the knowledge base;
- ontology of the subject domain of problem structures in knowledge bases, that is, structures that describe incomplete, incorrect or excessive information in the knowledge base;
- tools of specifying changes and transient processes in the knowledge base.

Tools for processing the knowledge base of the system of support of collective development of knowledge bases is

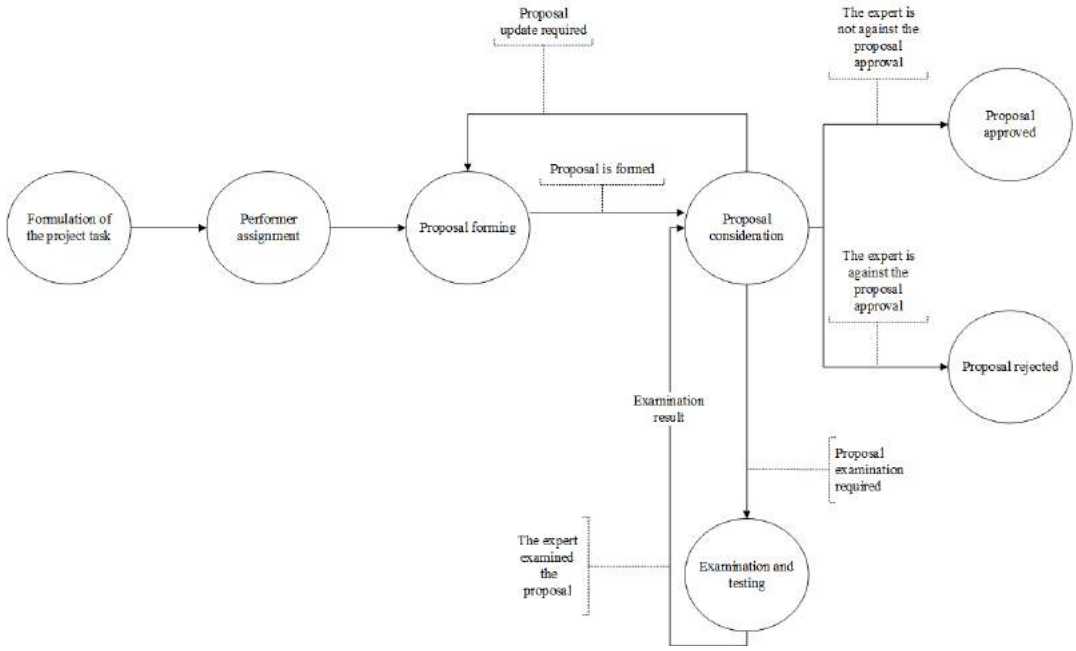


Figure 16. The mechanism for knowledge base fragments coordination

Figure 17. Architecture of the system of support of collective development of knowledge bases

a group of agents [45], each of which automates actions belonging to one of the knowledge bases developers actions classes discussed above. The tools for knowledge processing includes the following agents:

- class of agents for the verification of knowledge bases;
- class of knowledge base editing agents;
- class of automation agents for the activity of the knowledge base developer;
- class of automation agents for the activity of the knowledge base administrator;
- class of automation agents for the activity of the knowledge base manager;
- class of automation agents for the activity of the knowledge base expert;
- class of agents for calculating the characteristics of the knowledge base.

The *user interface of the system of support of collective development of knowledge bases* is represented by a set of interface commands that allow developers to initiate the activity of the required agent that is part of this system [40]. This set completely corresponds to the above set of the knowledge processing agents.

The developed tools for automating the process of building and modifying knowledge bases implement the proposed method and ensure coordination, verification and editing of knowledge base fragments directly in the process of their using.

VI. Conclusion

The proposed models, methods and tools were used to develop the knowledge bases of a number of systems, in particular, the IMS metasystem, as well as a number of applied intelligent reference systems in various subject domain, such as geometry, discrete mathematics, numerical models, chemistry, etc. , as well as in the development of a prototype of the batch production enterprises automation system [46].

Acknowledgement

This work was supported by BRFFR-RFFR (No F16R-101), BRFFR-SFFRU (No F16K-068).

References

[1] T. Gavrilova, D. Kudryavtsev, and D. Muromtsev, *Inzheneriya znanii. Modeli i metody: Uchebnik [Knowledge engineering. Models and methods: Textbook]*. SPb.: Lan', 2016.

[2] G. Alor-Hernández and R. Valencia-García, Eds., *Current Trends on Knowledge-Based Systems*. Springer, 2017.

[3] A. Kolesnikov, I. Kirikov, and S. Listopad, *Gibridnye intellektual'nye sistemy s samoorganizatsiei: koordinatsiya, soglasovannost', spor [Hybrid intelligent systems with self-organization: coordination, consistency, dispute]*. M.: IPI RAN, 2014.

[4] L. Kalinichenko, *Metody i sredstva integratsii neodnorodnykh baz dannykh [Methods and tools for integrating heterogeneous databases]*. M.: Nauka, 1983.

[5] Y. Val'kman, V. Gritsenko, and A. Rykhal'skii, *Model'no-parametricheskoe prostranstvo. Teoriya i prilozhenie [Model-parametric space. Theory and application]*. Kiev: Naukova dumka, 2010.

[6] Y. Zagorul'ko, "O kontseptsii integrirovannoi modeli predstavleniya znanii [on the concept of an integrated knowledge representation model]," *Izvestiya TPU*, vol. 322, no. 5, pp. 98–103, 2013.

[7] A. Kleshchev, *Semanticheskie porozhdayushchie modeli. Obshchaya tochka zreniya na freimy i produktsii v ekspertnykh sistemakh [Semantic generative models. General point of view on frames and products in expert systems]*, Vladivostok, 1986.

[8] V. Rubashkin, "Ontologii: ot informatsionno-poiskovykh tezaurusov k inzhenerii znanii [ontology: from information-search thesauri to knowledge engineering]," in *Desyataya natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2006 (Obninsk, 25-28 sentyabrya 2006 g.)*. Moskva: Fizmatlit, 2006.

[9] G. Shchedrovitskii, *Sintez znanii: problemy i metody [Synthesis of knowledge: problems and methods]*. M.: Nauka, 1984, pp. 67–99.

[10] T. Gavrilova and V. Khoroshevskii, *Bazy znanii intellektual'nykh sistem: Uchebnik [Knowledge bases of intellectual systems: Textboo]*. SPb.: Piter, 2001.

[11] T. H. R. Chan, "Design an integrated knowledge representation (ikr) model," Degree of Master of Philosophy, Department of Manufacturing Engineering and Engineering Management, May 2006.

[12] V. Ivashenko, "Modeli i algoritmy integratsii znanii na osnove odnorodnykh semanticheskikh setei [models and algorithms of knowledge integration based on homogeneous semantic networks]," avtoref. dis… kand. tekhn. nauk: 05.13.17, V.P. Ivashenko ; Uchrezhdenie obrazovaniya «Belorusskii gosudarstvennyi universitet informatiki i radioelektroniki», Minsk, 2014.

[13] (2016) The syntax of cycl. [Online]. Available: http://www.cyc.com/documentation/ontologists-handbook/cyc-basics/syntax-cycl

[14] P. Benjamin, C. Menzel, R. Mayer, F. Fillion, M. Futrell, P. S. deWitte, and M. Lingineni, *IDEF5 Method Report*. College Station : Knowledge Based Systems, Inc., 1994.

[15] *ISO/IEC 13211-1:1995 Information technology – Programming languages – Prolog – Part 1: General core [Electronic resource]*, International Organization for Standardization Std., mode of access: https://www.iso.org/obp/ui/#iso:std:iso-iec:13211:-1:ed-1:v1:en. — Date of access: 08.08.2016.

[16] (2017) Clips - a tool for building expert systems. [Online]. Available: http://clipsrules.sourceforge.net/

[17] (2017, Dec.) The world wide web consortium. [Online]. Available: https://www.w3.org

[18] V. Khoroshevskii, "Prostranstva znanii v seti internet i semantic web (chast' 1) [knowledge spaces on the internet and semantic web (part 1)]," *Iskusstvennyi intellekt i prinyatie reshenii*, no. 1, pp. 80–97, 2008.

[19] "Ontology Tools [Electronic resource]," *Open Semantic Framework*, mode of access: http://wiki.opensemanticframework.org/index.php/ Ontology_Tools. - Date of access: 21.09.2016.

[20] S. Gorshkov, *Vvedenie v ontologicheskoe modelirovanie [Introduction to ontological modeling]*. Ekaterinburg: OOO «TriniData», 2016.

[21] V. Golenkov, "Grafodinamicheskie modeli i metody parallel'noi asinkhronnoi pererabotki informatsii v intellektual'nykh sistemakh [graphodynamic models and methods of parallel asynchronous information processing in intelligent systems]," Dis… d-ra tekhn. nauk: 05.13.11, 05.13.17, Minsk, 1996.

[22] I. Davydenko, N. Grakova, E. Sergienko, and A. Fedotova, "Sredstva strukturizatsii semanticheskikh modelei baz znanii [means of structuring semantic models of knowledge bases]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed., BGUIR. Minsk : BGUIR, 2016, pp. 93–106.

[23] I. Davydenko, "Ontology-based knowledge base design," in *Open semantic technologies for intelligent systems*, V. Golenkov, Ed., BSUIR. Minsk : BSUIR, 2017, pp. 57–72.

[24] "Defining n-ary relations on the semantic web [Electronic resource]," *The World Wide Web Consortium (W3C)*, mode of access: https://www.w3.org/TR/swbp-n-aryRelations/#pattern1. — Date of access: 30.09.2016.

[25] (2017) Representing specified values in owl: "value partitions" and "value sets" in the world wide web consortium (w3c). [Online]. Available: https://www.w3.org/TR/2005/NOTE-swbp-specified-values-20050517/

[26] S. Rassel and P. Norvig, *Artificial Intelligence. The modern approach*. Moskva : Vil'yams, 2006.

[27] I. Efimenko and V. Khoroshevskii, *Ontologicheskoe modelirovanie ekonomiki predpriyatii i otraslei sovremennoi Rossii: Chast' 1. Ontologicheskoe modelirovanie: podkhody, modeli, metody, sredstva, resheniya [Ontologicheskoe modelirovanie ekonomiki predpriyatii i otraslei sovremennoi Rossii: Chast '1. Ontologicheskoe modelirovanie: podkhody, modeli, metody, sredstva, resheniya]*, F. Aleskerov, V. Podinovskii, and B. Mirkin, Eds. M.: Izd. dom Vysshei shkoly ekonomiki, 2011, (Preprint / WP7/2011/08 (ch. 1)).

[28] V. Gribova, A. Kleshchev, F. Moskalenko, V. Timchenko, L. Fedorishchev, and E. Shalfeeva, "Platforma dlya razrabotki oblachnykh intellektual'nykh servisov [platform for developing cloud-based intelligent services]," in *XV natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2016 (3–7 oktyabrya 2016 g., g. Smolensk, Rossiya). Trudy konferentsii. V 3-kh tomakh*. Smolensk: Universum, 2016, pp. 24–33.

[29] E. J. Reddy, C. Sridhar, and V. P. Rangadu, "Knowledge based engineering: Notion, approaches and future trends," *American Journal of Intelligent Systems*, no. 5 (1), pp. 1–17, 2015.

[30] H. Eriksson, Y. Shahar, S. Tu, A. Puerta, and M. Musen, "Task modeling with reusable problem-solving methods," *Artificial Intelligence*, no. 79 (2), pp. 293–326, 1995.

[31] R. Mizoguchi, J. Vanwelkenhuysen, and M. Ikeda, *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*. Amsterdam : IOS Press, 1995, ch. Task ontology for reuse of problem solving knowledge, pp. 46–59.

[32] A. N. Borisov, "Postroenie intellektual'nykh sistem, osnovannykh na znaniyakh, s povtornym ispol'zovaniem komponentov [building knowledge-based intelligent systems, reusing components]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]*, V. G. (otv. red.) [i dr.], Ed. Minsk: BGUIR, 2014, pp. 97–102.

[33] B. Wielinga and A. Schreiber, *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*. Amsterdam : IOS Press, 1994, ch. Reusable and sharable knowledge bases: a European perspective, pp. 110–120.

[34] R. Hartung and A. Håkansson, "Using meta-agents to reason with multiple ontologies," *Agent and Multi-Agent Systems: Technologies and Applications. KES-AMSTA 2008. Lecture Notes in Computer Science*, vol. 4953, pp. 261–270, 2008.

[35] A. Gladun and Y. Rogushina, "Repozitorii ontologii kak sredstvo povtornogo ispol'zovaniya znanii dlya raspoznavaniya informatsionnykh ob"ektov [ontology repositories as a means of re-using knowledge to recognize information about objects]," *Ontologiya proektirovaniya*, no. 1, pp. 35–50, 2013.

[36] C. Debruyne, P. D. Leenheer, and R. Meersman, *On the Move to Meaningful Internet Systems: OTM 2009 Workshops*. Heidelberg : Springer, 2009, ch. Method and Tool for Fact Type Reuse in the DOGMA Ontology Framework, pp. 1147–1164.

[37] V. Lapshin, *Ontologii v komp'yuternykh sistemakh [Ontologies in computer systems]*. M.: Nauchnyi mir, 2010.

[38] I. Davydenko, "Tekhnologiya komponentnogo proektirovaniya baz znanii na osnove unifitsirovannykh semanticheskikh setei [the technology of component design of knowledge bases on the basis of unified semantic networks]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed., BGUIR. Minsk : BGUIR, 2013, pp. 185–190.

[39] D. Shunkevich, I. Davydenko, D. Koronchik, I. Zhukov, and A. Parkalov, "Sredstva podderzhki komponentnogo proektirovaniya sistem, upravlyaemykh znaniyami [tools for supporting the component design of knowledge-driven systems]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. BGUIR, 2015, pp. 79–88.

[40] D. Koronchik, "Semanticheskie modeli mul'timodal'nykh pol'zovatel'skikh interfeisov i semanticheskaya tekhnologiya ikh proektirovaniya [semantic models of multimodal user interfaces and semantic technology of their design]," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. Minsk: BGUIR, 2012, pp. 339–346.

[41] D. Shunkevich, A. Gubarevich, M. Svyatkina, and O. Morosin, “Formal’noe semanticheskoe opisanie tselenapravlennoi deyatel’nosti razlichnogo vida sub"ektov [a formal semantic description of the purposeful activity of various types of sub-objects],” in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual’nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. Minsk: BGUIR, 2016, pp. 125–136.

[42] (2016) Protege. ontology editor and framework for building intelligent systems. [Online]. Available: http://protege.stanford.edu

[43] A. Gomez-Perez and M. Suarez-Figueroa, “Scenarios for building ontology networks within the neon methodology,” in *Proceedings of the Fifth International Conference on Knowledge Capture (K-CAP 2009).* New York : ACM, 2009, pp. 183–184.

[44] J. Euzenat, *10th Knowledge Acquisition, Modeling and Management for Knowledge-based Systems Workshop (KAW’96).* Banff : Alberta, 1996, ch. Corporate memory through cooperative creation of knowledge bases and hyper-documents, pp. 1–18.

[45] D. Shunkevich, “Ontologicheskoe proektirovanie mashin obrabotki znanii [ontological design of knowledge processing machines],” in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual’nykh sistem [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. Minsk: BGUIR, 2017, pp. 73–94.

[46] I. Davydenko, V. Golenkov, V. Taberko, D. Ivanyuk, K. Rusetskii, D. Shunkevich, V. Zakharov, V. Ivashenko, and D. Koronchik, “Proektirovanie predpriyatii retsepturnogo proizvodstva na osnove ontologii [designing enterprises of prescription production on the basis of ontology],” *Ontologiya proektirovaniya*, vol. 7, no. 2(24), pp. 123–144, 2017.

СЕМАНТИЧЕСКИЕ МОДЕЛИ, МЕТОД И СРЕДСТВА СОГЛАСОВАННОЙ РАЗРАБОТКИ БАЗ ЗНАНИЙ НА ОСНОВЕ МНОГОКРАТНО ИСПОЛЬЗУЕМЫХ КОМПОНЕНТОВ

И.Т. Давыденко
Белорусский государственный университет
информатики и радиоэлектроники
г. Минск, Беларусь

База знаний является ключевым компонентом такого класса компьютерных систем, как системы, основанные на знаниях, разработка которых является одним из перспективных направлений в области искусственного интеллекта. Качество разрабатываемых систем такого класса определяется, в том числе, качеством базы знаний и разнообразием видов знаний, хранимых в ней.

Расширение областей применения систем, основанных на знаниях привело к необходимости поддержки решения комплексных задач. Под комплексной задачей будем понимать задачу, решение которой предполагает применение формализованных знаний различного вида и различных моделей их обработки, что, в свою очередь, требует обеспечения совместимости и интеграции используемых знаний, а также моделей их обработки.

Настоящая работа посвящена решению задач, связанных с разработкой моделей методов и средств создания баз знаний компьютерных систем, способных решать комплексные задачи.

Agent-oriented models, method and tools of compatible problem solvers development for intelligent systems

Shunkevich D. V.
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
shunkevichdv@gmail.com

Abstract—**The article is devoted to the development of agent-oriented models, a method and means for developing compatible solvers of problems of intelligent systems capable to solve complex problems. The requirements for such solvers, the model of problem solver that satisfy the requirements, as well as the method and tools for developing and modifying such solvers are considered.**

The main problem considered in the work is the problem of low consistency of the principles underlying the implementation of various problem solving models. As a consequence, it is difficult to simultaneously use different models for solving problems within the same system when solving the same complex problem, it is practically impossible to reuse the technical solutions implemented in a certain system, in fact, there are no integrated methods and tools for problem solvers development.

As a basis for problem solvers design, it is proposed to use the multi-agent approach. The process of any problem solving is proposed to decompose into logical atomic actions, which will ensure compatibility and modifiability of the solvers. The solver is proposed to be considered as a hierarchical system consisting of several interconnected levels, which allows to provide the possibility of independent designing, debugging and verification of components at different levels.

Keywords—**semantic technologies, ostis-system, problem solver, multi-agent system, intelligent agent, knowledge base**

I. Introduction

One of the key components of each intelligent system is the problem solver, which provides the ability to solve various problems, related both to the basic functionality of the system, and to the such a system efficiency ensuring, as well as the development automation of the system itself. The problem solver, ensuring the fulfillment of all listed functions, will be called the **integrated problem solver**.

The capabilities of the problem solver largely determine the functionality of the intelligent system as a whole, the ability to answer on non-trivial user questions and solve various problems in a certain subject domain.

The composition of the solver of each particular system depends on its purpose, the classes of problems being solved, the subject domain and a lot of other factors. Expanding of the scope of intelligent systems applications requires such systems to be able to solve **complex problems**, that is, problems that require the application of a variety of different knowledge representation models and different knowledge processing models. Examples of such tasks are:

- The problem of the natural language texts understanding, both printed and hand-written, understanding of speech messages, semantic analysis of images. In each of the listed cases, it is necessary first to performs the syntactic analysis of the processed file, remove the insignificant fragments, then classify the significant fragments, correlate them with the concepts known to the system, identify those fragments, that the system can not recognize, eliminate duplication of information, etc.;
- The problem of automating adaptive learning of students, suggesting that the system itself can solve various problems from a certain subject domain, and also manage the learning process, create tasks for students and monitor their implementation;
- The problem of intelligent robots behavior planning, including both understanding of various kinds of external information, and the various decision making, using both reliable and plausible methods.
- The problem of complex and rapidly evolving automation of various enterprises;
- and others.

At present, there are a lot of problem solving models of various kinds, including the variety of types of logics (clear, fuzzy, inductive, deductive, temporal, etc.), neural networks and genetic algorithms, various strategies of problem solution search ways (depth-first search, breadth-first search, etc.), various programming languages, both declarative and imperative.

The ability of different models using for problems solving within a single system will allow to decompose the complex problem into subproblems, each of which can be solved by one of the methods, known to the system. Thanks to a combination of different problem solving models, the number of problem classes, which such a system will be able to solve, will be significantly larger than the total number of problem classes, which can be solved by several systems, each of which implements only one of the problem solving models being integrated.

Modern intelligent systems, oriented to the simultaneous use of different types of knowledge and various problem solving models, are built on the principle of hybrid computer systems [1], [2]. Such an approach allows to solve complex problems, however, during hybrid systems design process, it becomes necessary to ensure the interaction interface for various problem solving models. That substantially increases the overhead in the such systems construction. As a result, such systems, as a rule, have a complex monolithic architecture, the introduction of any changes in which requires considerable work. To solve this problem, it is necessary to ensure the compatibility of various problem solvers.

At the same time, the urgent problem is the intelligent system teaching to new knowledge and skills and adapting it to permanently changing requirements. It is necessary to attach new resources, including new approaches to solving of problems of various classes. At the same time, unlike most modern approaches to computer systems learning (machine learning) [3], where the class of tasks being solved is actually fixed (does not change during the learning process) and only the method of solving is optimized, in this case we are talking about the expansion of the number of classes of problems solved by the system, and in the general case - unlimited expansion.

An important way to reduce the complexity of the process of changing the functionality of intelligent systems is the accumulation of libraries of reusable components of solvers that will significantly reduce both the terms of development and modification of solvers, and the level of professional requirements for their developers. At the same time unification of various models of problem solving on a common formal basis will allow to form not only traditional libraries of standard subprograms, but also libraries of entire solvers that implement one or another problems solving model.

Let us consider in detail one of the examples of complex problems listed above, which is related to the enterprise automation.

Within the integrated system of the enterprise automation, the following automation levels can be conventionally distinguished:

- automation of the actual production process at all stages, from the receipt and evaluation of raw materials to packaging and goods delivery to the end user;
- automation of production process management, that is, automation of making changes in the production process, for example, changes in batch quantities, nomenclature or properties of the manufactured product, etc;
- automation of production process control, which involves the use of various methods of the current situation analysis, as well as mechanisms for identifying, classifying and eliminating of emergencies, up to complete elimination of the emergency situation without operator intervention.

Figure 1 describes a schematic diagram of the integrated system of the enterprise automation, showing the application of which problem solving models is actual in each of the subsystems of such a system.

In the figure 2 the conventional scheme of the contingencies handling subsystem is shown in more detail.

This example shows that design of such a complex automation system is impossible without ensuring the consistent use of different types of knowledge and problem solving models within the same system when solving the same complex problem. In addition, the problem of such a system support in a state corresponding to the current production technologies level, supplementing it with more advanced models and methods of problem solving becomes urgent. It is obvious that such a system reconfiguration should be carried out directly during the system operation, and not require a complete stop of the entire production or its individual parts at every time. Thus, it can be said that such a system should be learnable, that is being able to acquire not only new knowledge, but also new skills.

The foregoing allows us to formulate requirements for the problem solver of an intelligent system, which is able to solve complex problems:

- at each point in time the solver must ensure the solution of problems from the specified class for a specified time, and the result of the problem solution must satisfy certain known requirements. In other words, as in the case of modern computer systems, the correctness of the problem solving results at the system development stage should be verified by special methods, including such modern approaches as unit-testing, «black box» testing and others [4].
- the solver should be easily **modifiable**, that is, the complexity of making changes to an already developed solver should be minimal. Ways to increase the modifiability are the ensurance of the introduced changes localness, as well as the availability of ready-made reusable components that can be included into the solver if necessary. At the same time, changes must be made directly during the system operation, and the overhead of new components integrating or replacing existing ones should be minimal.
- in order for the intellectual system to be able to analyze and optimize the existing problem solver, to integrate new components into it (even by ssystem itself), to evaluate the importance of certain components and their applicability for particular problem solving, the specification of the solver should be defined with system understandable language, for example, with the same means as the processed knowledge. The ability of an intelligent system to analyze (verify, correct, optimize) its own components will be called **reflexivity**.
- an additional requirement to the integrated problem solver in relation to the solver in general is its **completeness** (integrity, complexity), that is, such a solver must provide all the functionality of the system, i.e. ensure the solution of all problems, both related to the direct designation of the system, and ensuring the effectiveness of the system operation.

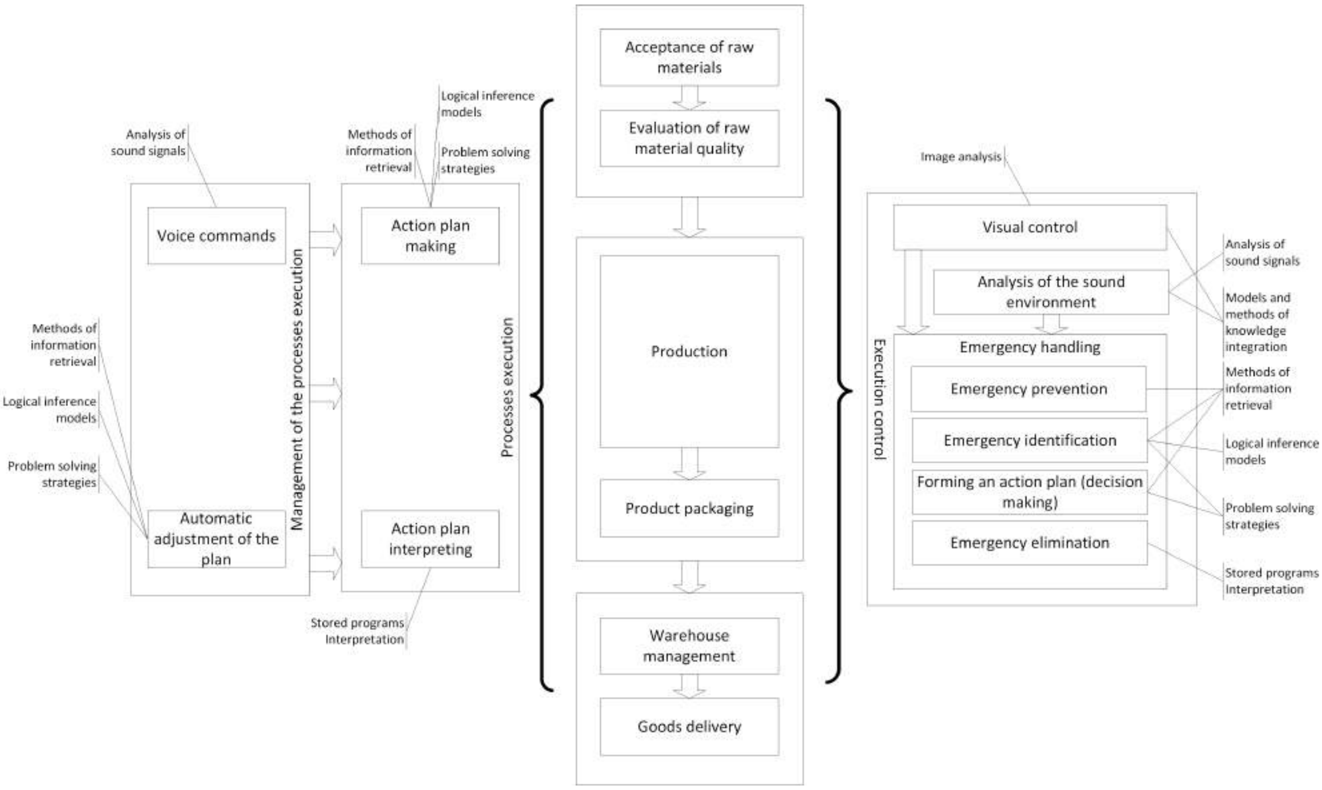


Figure 1. Simplified scheme of integrated system of the enterprise automation

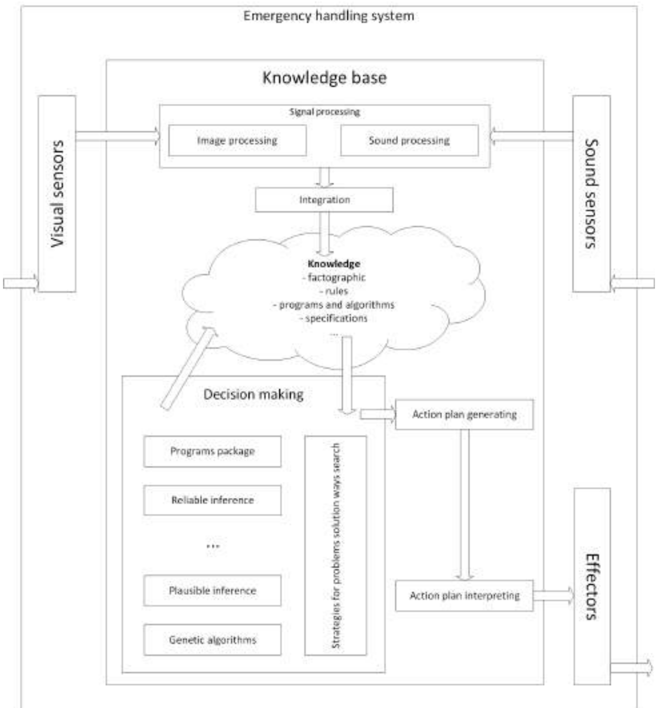


Figure 2. The simplified scheme of the contingencies handling subsystem

A. Problems in the development of problem solvers

Despite the fact that there are currently a lot of problem solving models, many of which are implemented and successfully used in practice in various systems [5], [6], [7], [8], the problem of low consistency of the principles underlying the implementation of different models of this kind, leads to the fact that:

- it is difficult to simultaneously use different models of problem solving within the same system for the same complex problem solving, it is practically impossible to combine different models to solve a problem for which there is no a priori algorithm;
- it is practically impossible to use technical solutions implemented in one system in other systems, i.e. the possibility of the component approach implementation in the problem solvers design is very limited. As a consequence, there is a large number of duplications of similar solutions in different systems;
- in fact, there are no complex methods and tools for problem solvers construction that would provide the ability to design, implement and debug solvers of various types.

The consequences of these problems are:

- high complexity of each solver development, increase of their development time, and as a result - the increase of the corresponding intelligent systems development and support cost;

- high complexity of making changes to the already developed solvers, i.e. there is no possibility or it is very difficult to add new components to the already developed solver and to make changes to existing components during the system operation process;
- a high level of professional requirements to the developers of solvers;

B. The proposed approach

The development of a problem solvers model that meets the above requirements, as well as the corresponding method and tools for their construction and modification, is proposed to be implemented within the *OSTIS Technology* [9]. As a formal basis for representation of knowledge within this technology approach, a unified semantic network with set-theoretic interpretation is used. This representation model is called ***SC-code***. Elements of such a semantic network are called sc-nodes and sc-connectors (sc-arcs, sc-edges). The model of an entity, described by means of SC-code will be called the semantic model of the specified entity or simply a *sc-model*. Computer systems, based on OSTIS Technology are called ***ostis-systems***. Each ostis-system consists of a sc-models interpreting platform and the sc-model of this system, which does not depend on the platform of its interpretation.

Orientation to OSTIS Technology and SC-code is due to their following advantages and peculiarities:

- SC-code is oriented to the **sense** representation of knowledge, which allows to generalize the problems solving models and thus significantly reduce the variety of different models, which is largely due to the representation forms of certain knowledge types, rather than their meaning;
- SC-code, unlike other widely used knowledge representation languages, allows to represent in a unified form any kinds of knowledge, including logical statements and programs. This fact makes it possible to unify also the problem solution models on the basis of knowledge presented in this form and to ensure integration of various problem solving models on the basis of such formalism;
- the associativity and structural reconfigurability of the semantic memory (sc-memory), in which the SC-code constructions are stored, makes it possible to provide modifiable models for solving problems, presented on its basis.

In addition, the work involves a number of solutions developed within the technology:

- a graph procedural programming language SCP, programs of which are also written using SC-code and which will be used as the base programming language within the proposed approach;
- models of knowledge base structuring and models of various types of knowledge representation, built on the basis of SC-code, presented in [10];
- the method of consistent construction and modification of knowledge bases, presented in the work [11];
- means for automatic editing and verification of various knowledge types, presented in the same work;
- an implementation version of the computer systems sc-models interpreting platform, considered in the work [12].

At the base of the proposed approach to the previously formulated problems solution are the following principles:

- as a basis for problem solvers design, it is proposed to use a multi-agent approach that will allow to build parallel asynchronous systems, having a distributed architecture, as well as to increase the modifiability and performance of the solvers developed.
- the process of any problem solution is proposed to decompose into logically atomic actions, which will provide compatibility and modifiability of the solvers too.
- the solver is suggested to be considered as a hierarchical system consisting of several interconnected levels. This approach allows to provide the possibility of designing, debugging and verifying components at different levels independently of other levels.
- in order to ensure the reflexivity of the designed intelligent systems, it is proposed to record all information about the solver and the problems it solves with the means of SC-code in the same knowledge base as the subject knowledge of the system. In general, this information includes: (1) the specification of the solver's agents, including the full texts of agent programs (in SCP language); (2) specification of all information processes performed by agents in the semantic memory, including - constructions that ensure synchronization of parallel processes; (3) specification of all problems on solution of which the specified information processes are directed;
- when designing the solver as a hierarchical system, it is suggested to use at each level the component approach, which allows to significantly reduce the development time and improve the reliability of solvers by using well-debugged components. To implement this approach, it is proposed to develop within the IMS metasystem [13] a library of solvers components of various levels, as well as a method for solvers constructing and modifying, which takes into account the existence of such a library.
- it is proposed to build automation and information support tools for solvers developers using OSTIS Technology, that is, including using models, methods and tools offered in this work. Such an approach will allow to ensure high rates of development of these tools, as well as significantly improve the effectiveness of information support tools, allowing to build these tools as part of the intelligent metasystem IMS, that is, as a kind of intellectual subsystem.

Orientation to the multi-agent approach as a basis for the modifiable solvers constructing is due to the following main advantages of this approach [14]:

- autonomy (independence) of agents within such a system, which allows to localize the changes introduced into the

solver during its evolution, and reduce the corresponding labor costs;
- Decentralization of processing, i.e. the absence of a single monitoring center, which also allows to localize the changes introduced into the solver.

The most common and widely used definition of intelligent agent is given in [15].

Existing approaches to the construction of multi-agent systems are discussed in detail in the papers [16], [17], [18], [19], also a specialized journal Autonomous Agents and Multi-Agent Systems [20] is devoted to multi-agent systems.

In the general case, in order to construct a concrete multi-agent system, it is necessary to clarify the following components:

- a model of the agent itself, which is part of such a system, including the classification of such agents and a set of concepts that characterize each agent within the system. Currently, the most popular is the BDI (belief-desire-intention) model, in which it is intended to describe the «beliefs», «desires» and «intentions» of each agent of the system in appropriate languages.
- a model of the environment within agents are located, on the events in which they react and within which they can perform some transformations. A survey on the varieties of environments for multi-agent systems is given in [21].
- an agent communication model that specifies the agent interaction language (the structure and classification of messages) and the way messages are exchanged between agents. Currently, there are a number of standards describing agent interaction languages, for example, KQML [22] and ACL [23].
- a model of agents coordination, regulating the principles of their activities, including mechanisms for resolving possible conflicts. Currently, the most works in the field of multi-agent systems is aimed specifically at mechanisms for coordinating agents, including the allocation of a higher level agents (meta-agents) [24], various socio-psychological models [25], [26], ontology-based behavior [17] and others [27], [28].

The main disadvantages of most popular modern tools for building multi-agent systems [29], [30], [31], [32], [33], [2], [34], [35] include the following:

- the rigid orientation of most tools to the BDI model leads to significant overhead costs associated with the need to express a particular practical task in the BDI concept system. At the same time, the orientation toward the BDI model implicitly provokes the artificial separation of languages, describing the BDI components themselves and the agent's knowledge about the external environment. That leads to the lack of unification of the representation and, correspondingly, to additional overhead costs.
- most modern means of multi-agent systems construction are oriented to the representation of agent's knowledge using highly specialized languages, often not intended to represent knowledge in a broad sense. Here we mean both the agent's knowledge of himself (for example, in accordance with the BDI model) and knowledge about the external environment. In some approaches, an ontology is first constructed, which, however, often uses tools with low expressiveness that are not designed for ontologies building [32], [33]. Ultimately, this approach leads to a strong limitation of the capabilities of the developed multi-agent systems and their incompatibility.
- the absolute majority of modern tools assume that agents interact through messaging directly from the agent to the agent. This approach has a significant disadvantage due to the fact that in this case each agent of the system should have sufficient information about other agents in the system, which leads to additional resource costs. In addition, adding or removing one or more agents leads to the need to notify other agents about that change. This problem is solved by organizing agents' communication on the «blackboard» principle [36], which assumes that messages are placed in some common area for all agents, and each agent in general case may not know to which of the agents the message is addressed and from which agent the message was received. However, this approach does not exclude the problem associated with the need to develop a specialized agent interaction language that is not generally associated with a language that describes the agent's knowledge about the problems to be solved and about the environment.
- a lot of means of multi-agent systems construction are designed in such a way that the logical level of agents' interaction is rigidly tied to the physical level of the multi-agent system implementation. For example, when sending messages from an agent to an agent, the developer of a multi-agent system needs, in addition to the semantically significant information, to specify the ip-address of the computer on which the receiving agent is located, the encoding with which the message text was encoded and other technical information, which depends on the features of the concrete tools implementation.
- in most approaches, the environment with which agents interact is specified separately by the developer for each multi-agent system, which, on the one hand, expands the possibilities of corresponding means using, but on the other hand leads to significant overhead and incompatibility of such multi-agent systems. In addition, in some cases, the developer also has to take into account the specifics of the technical implementation of development tools in terms of their docking with the intended environment, which, for example, can be a local or global network.

Within this work, the listed disadvantages are supposed to be eliminated by using the following principles:

- communication of agents is suggested to be implemented on the basis of the «blackboard» principle, however, unlike the classical approach, in the role of messages there are specifications in the general semantic memory of the actions (processes) performed by agents and aimed at solving any problems, and the role of communication environment is played by this semantic memory itself. This approach allows to:

  - exclude the need to develop a specialized language for messaging;
  - ensure the «impersonality» of communication, i.e. each agent generally does not know which other agents are in the system, who has formulated the request and to whom this or that request is addressed. Thus, adding or removing agents to the system does not lead to changes in other agents, which ensures the modifiability of the entire system;
  - agents, including the end user, get the opportunity to formulate tasks in the *declarative way*, i.e. do not declare for each problem the way to solve it. Thus, the agent does not need to know in advance how the system will solve a particular problem, it is enough only to specify the final result;

  It should be noted that this approach allows, if necessary, to organize messaging between agents directly, and, thus, can be the basis for modeling multi-agent systems that implement other ways of interaction between agents.
- in the role of the external environment for agents is the same semantic memory, in which problems are formulated and through which agents interact. This approach ensures unification of the environment for different agent systems, which in turn ensures their compatibility.
- the specification of each agent is described by means of SC-code in the same semantic memory, which allows:
  - minimize the number of specialized means required to specify agents, including language and tools;
  - on the one hand - to minimize the necessary specification of the agent in the general case, which includes the condition of its initiation and the program that describes the algorithm of the agent, on the other hand - to provide the possibility of unlimited expansion of such specification for each specific case, including the possibility of implementing the BDI model and others;
- synchronization of the agents activities is supposed to be carried out at the level of the processes performed by them, aimed at solving certain problems in the semantic memory. Thus, each agent is treated as an abstract processor, which is able to solve the tasks of a particular class. With this approach, it is necessary to solve the problem of ensuring the interaction of parallel asynchronous processes in the common semantic memory, for the solution of which the solutions used in traditional linear memory can be adapted.
- each information process at any time has associative access to the necessary fragments of the knowledge base stored in the semantic memory, except of fragments blocked by other processes in accordance with the synchronization mechanism discussed below. Thus, on the one hand, the need to store information about the external environment by each agent is excluded, on the other hand, each agent, like in classical multi-agent systems, has only a part of all the information necessary to solve the problem.

  It is important to note that in the general case it is impossible to predict a priori which knowledge, models and methods of problem solving will be needed for the system to solve a specific problem. In this regard, it is necessary to ensure, on the one hand, the ability to access all the necessary fragments of the knowledge base (in the limit, to the entire knowledge base), on the other hand, to be able to localize the area of the problem solution search, for example, within a single subject domain [11]. Each agent has a set of key elements (usually concepts) that it uses as starting points for associative search within the knowledge base. A set of such elements for each agent is specified at the stages of a multi-agent system design in accordance with the method considered below. Reducing the number of key elements of the agent makes it more universal, but it reduces the effectiveness of its work due to the need to perform additional search operations.

Next, consider the model of knowledge processing and the model of the problem solver itself in accordance with the listed principles.

II. General model of knowledge processing in ostis-systems

The model of processing the knowledge stored in semantic memory can be conditionally divided into two components (figure 3):

- the model of information processes performed in such a memory, including the classification of such processes, the mechanisms for their execution regulating, the means for various conflicts solving, including associated with the parallel execution of such processes, means of specifying the state of information processes (executing, delayed, planned, etc.);
- the model of the problem solver, which is treated as an abstract processor that performs the specified information processes, and, accordingly, the model of which is constructed taking into account the model of information processes in the semantic memory. The solver consists of a platform-independent part (solver program) that includes a model of the operational semantics of the SCP language (scp-interpreter model) and a platform-dependent part that includes the implementation of the scp-interpreter. In addition, some components of the solver program can, if necessary, be implemented within the sc-model interpretation platform, for example, to improve the solver performance.

This approach to the knowledge processing organization makes it possible to ensure the independence of the ostis-system sc-model (including the solver program) from the interpretation platform of such models. Thus, the development of

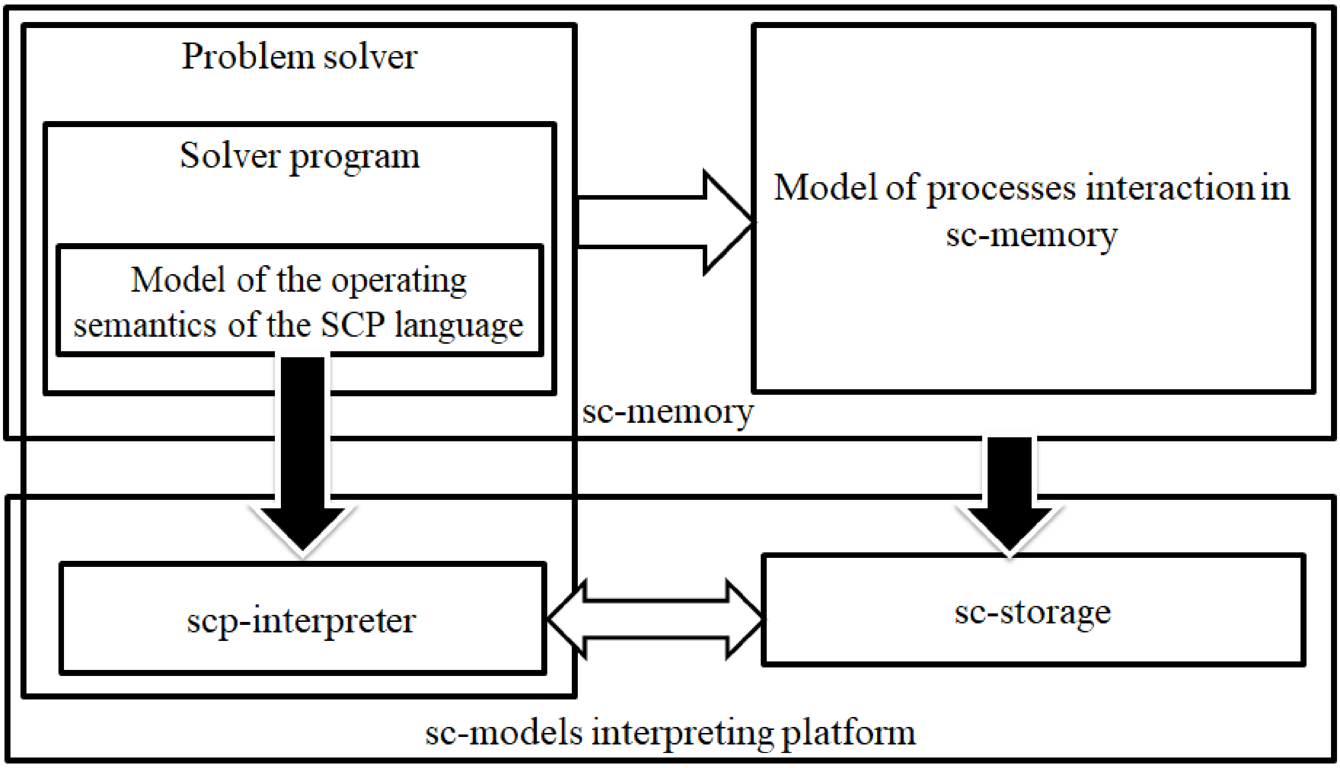


Figure 3. General scheme of knowledge processing organization in ostis-systems

a solver for a particular ostis-system is reduced to developing a program for this solver, i.e. the collective of the solver's agents and the programs corresponding to them. If the agent programs are implemented in the SCP language, such a solver program can be transferred from one implementation of the sc-models interpretation platform to another, including the hardware platform, without any changes.

In its turn, the formal model of the operational semantics of the SCP language is, in fact, a technical task for the implementation of the sc-model interpretation platform, both in software version and in hardware. The complete specification of the denotational and operational semantics of the SCP language is described in the corresponding sections of the IMS metasystem [13].

The development of a certain entity sc-model (including a program, an agent and a solver) supposes a formal refinement of the concepts system that is used to describe this entity in the knowledge base of the ostis-system. To achieve this goal, in accordance with the principles of OSTIS Technology, it is necessary to develop the sc-models of one or several interconnected *subject domains* and the corresponding *ontologies* [10].

The implementation of the approach proposed in this work requires the construction of several subject domains (SD) sc-models connected with each other, as shown in the figure 4.

Next, the most important concepts that are *researched concepts* [10] within the specified subject domains will be discussed in more detail.

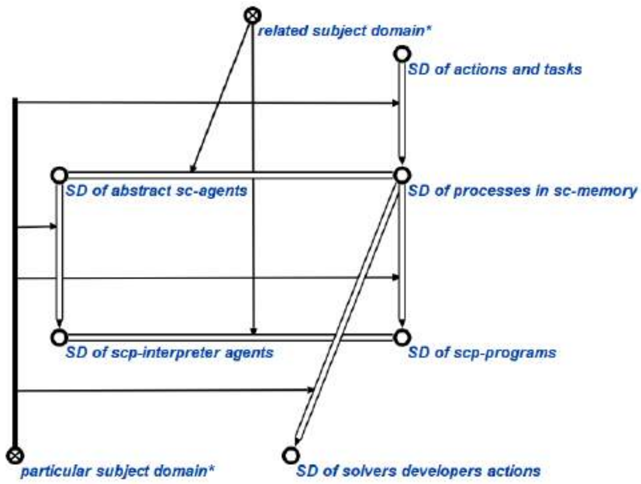


Figure 4. Hierarchy of subject domains

A. Means of specification of problems being solved

As it was said before, the principle of communication of sc-agents within the proposed approach is based on the specification by sc-agents of all the actions performed by them in semantic memory, i.e. in the formal description in the knowledge base of all problems being solved.

In the proposed approach, the problem is treated as a specification of some action, which in general can include such information as: the subject and the object of the performed action; the timing of the problem, its priority, dependent problems, etc .; detailed specification of the specified action execution process by its decomposition into sub-actions (procedural formulation of the problem); specification of the action goal (declarative formulation of the problem), etc.

The figure 5 shows an example of an information task in SCg language.

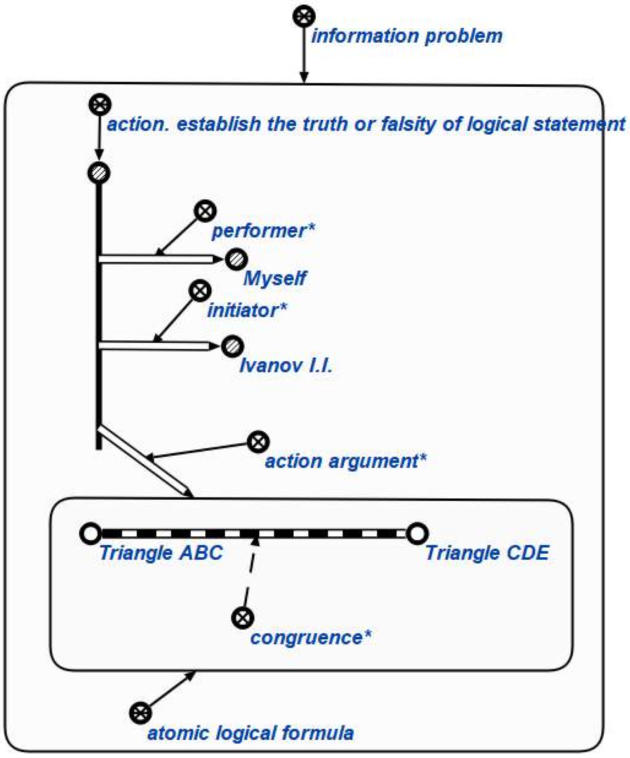


Figure 5. The problem of establishing the truth or falsity of a statement

The model of the activity performed by agents while solving problems in semantic memory is defined as follows:

$$M_A = \{A_C, A_{CM}, A_R, A_{CS}\}, \quad (1)$$

where A_C – set of classes of actions performed by different subjects;

A_{CM} – set of classes of actions performed by agents in semantic memory, $A_{CM} \subset A_C$;

A_R – set of relations specifying actions belonging to classes from A_C;

A_{CS} – a set of specification classes of actions belonging to classes from A_C, such as a task, a protocol for an action executing, etc.;

In more detail, the means for specifying the actions performed in semantic memory are discussed in [37].

One of the basic principles underlying the proposed approach to the construction of solvers is the principle of dividing the process of any problem solving in an intelligent system by *logically atomic actions*.

We will assume that each *action* belonging to some particular *class of logically atomic actions* has two necessary properties:

- The execution of an action does not depend on whether the specified action is part of the decomposition of a some general action. When this action is carried out, it should also not take into account the fact that the action precedes or follows any other action;
- the specified action should be a logically complete act of transformation, for example, in semantic memory. Such an action is essentially a transaction, i.e. the result of this transformation is the new state of the system being converted, and the action to be performed must either be executed completely, or not executed at all, partial execution is not allowed.

At the same time, logical atomicity does not prevent decomposition of the executed action into partial ones, each of which in turn will also belong to some class of logically atomic actions.

B. The concept of sc-agent

The only kind of entities that perform transformations in *sc-memory* is *sc-agents*. We will call an *sc-agent* some *subject* which is able to execute *actions in sc-memory*, belonging to some particular *class of logically atomic actions*.

Formally, the sc-agent model is defined as follows:

$$M_S = \{S_C, S_R\}, \quad (2)$$

where S_C – a set of classes (types) of sc agents; S_R – set of relations defined on the set of sc-agents.

Logical atomicity of sc-agent actions assumes that each sc-agent responds to the corresponding class of events occurring in the sc-memory, and performs a certain transformation of the sc-text (SC-code text) located in the semantic neighborhood of the event being processed. In this case, each sc-agent in general does not have information about what other sc-agents are currently present in the system and interacts with other sc-agents solely through the formation of certain constructions (usually the action specifications) in the common sc-memory. Such a message can be, for example, a question addressed to other sc-agents in the system (not known in advance what exactly) or the answer to the question formed by other sc-agents (again, it is not known exactly what). Thus, each sc-agent controls only the knowledge base fragment in the context of the problem solved by the agent at any given time, the state of the rest of the knowledge base is unpredictable for the sc-agent in the general case.

To ensure the availability of such a multi-agent system, it is necessary that each sc-agent in its composition specifies in the sc-memory all the results of its actions. It is assumed that after solving a certain problem:

- all the intermediate constructs generated in the solution process and having no meaning outside this process must be deleted;

- all processes (actions) in sc-memory, aimed at solving the same problem, should be terminated, except when it is supposed to receive several independent responses to the same questions.

It is important to note that the end-user of the ostis-system in terms of knowledge processing also acts as an sc-agent, forming the messages in the sc-memory by performing the elementary actions provided by the user interface. In the same way the ostis-system interacts with other systems and the environment in general. All information gets in and out the ostis-system exclusively through the appropriate sc-agents of the interface.

Let's list some advantages of the offered approach to the organization of knowledge processing in sc-memory:

- because of processing is performed by agents that exchange messages only via common memory, adding a new agent or excluding (deactivating) one or more existing agents usually does not result in changes to other agents, since agents do not exchange messages directly;
- the agents are initiated in a decentralized manner and, most often, independently of each other. Thus, even a significant increase of the agents number within the same system does not lead to its productivity reduce;
- agent specifications and, as will be shown below, their programs can be written in the same language as the processed knowledge, which significantly reduces the list of specialized means designed to develop such agents and their groups, and simplifies system development by using more universal components;

Since it is supposed that copies of the same *sc-agent* or functionally equivalent *sc-agents* can work in different ostis-systems, being physically different sc-agents, it is advisable to consider properties and classification of not sc-agents, but classes of functionally equivalent sc-agents, which we will call *abstract sc-agents*. Thus, an *abstract sc-agent* is a class of functionally equivalent *sc-agents*, different instances (i.e. elements) of which can be implemented in different ways.

Each *abstract sc-agent* has a corresponding specification, which specifies the key sc-elements of the specified abstract sc-agent, as well as a description of the initiating condition for this *sc-agent*, i.e. class of those *situations* in *sc-memory* that initiate the activity of this *sc-agent*. In addition, for each abstract sc-agent, the variant of its implementation is specified. From the implementation point of view, two classes of *abstract sc-agents* can be distinguished:

- *non-atomic abstract sc-agent*, which is decomposed into a group of simpler *abstract sc-agents*, each of which in turn can be both *atomic abstract sc-agent*, and *non-atomic abstract sc-agent*. However, in some version of *abstract sc-agent decomposition**, the child *non-atomic abstract sc-agent* can become an *atomic abstract sc-agent*, and be implemented accordingly.
- *atomic abstract sc-agent* is an abstract sc-agent, for which the platform of its implementation is specified, i.e. there is a corresponding connection of the *sc-agent program** relation.

In turn, *atomic abstract sc-agents* are subdivided into *platform-independent abstract sc-agents* and *platform-dependent abstract sc-agents*.

The *platform-independent abstract sc-agents* are *atomic abstract sc-agents* implemented in *SCP language*.

When describing *platform-independent abstract sc-agents*, platform independence is understood from the point of view of OSTIS Technology, i.e. implementation in the SCP language, because *atomic sc-agents* implemented in the specified language can be easily transferred from one sc-models interpretation platform to another.

The *platform-dependent abstract sc-agents* are *atomic abstract sc-agents*, implemented below the level of sc-models, i.e. not in the *SCP language*, but in some other language of the program description.

An example of an atomic abstract sc-agent specification including the agent program, its key sc-elements, and the initiation condition is given in the figure 6.

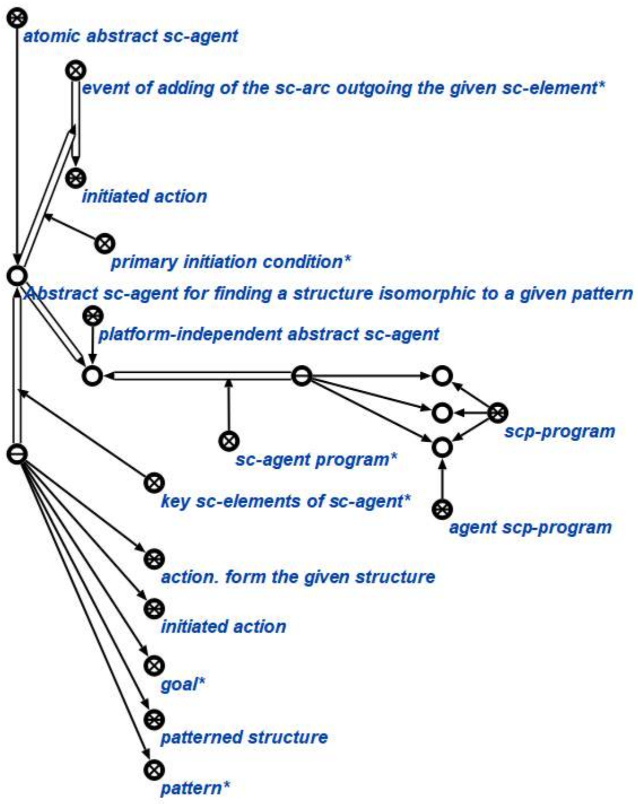


Figure 6. Atomic abstract sc-agent of search for structures isomorphic to a given pattern

The construction of non-atomic sc-agents allows to ensure the hierarchy of the designed multi-agent system and the ability to consider it at different levels of detail, which, in turn, provides the convenience of such a system designing and debugging through the ability to design and debug components of varying complexity independently. In addition, the allocation of non-atomic sc-agents is the basis for the formation of a hierarchical library of problem solvers reusable components, which will include components of different complexity levels, including even entire solvers.

C. Basic model of knowledge processing

The *SCP language*, developed within the OSTIS Technology, is proposed as the base development language for the programs describing the activity of sc-agents within *sc-memory*. *SCP language* is focused on processing of unified semantic networks represented with *SC-code*.

The basic *sc-text* processing model includes:

- model of *scp-programs* subject domain, which includes all the texts of *scp-programs*, and in which the classification of these programs operators and the means of their specification are researched.
- model of subject domain of the scp-programs interpretation agents (also called *Abstract scp-machine*), which is part of the sc-model interpretation platform. The term *abstract* in this case, as in the case of the *abstract sc-agent*, shows that a semantic model of the scp-interpreter is being developed, including the specification of each agent in its composition, which can later be implemented within the any sc-model interpretation platform, including hardware.

The main features of the SCP language include the following:

- texts of scp-programs are written using SC-code, as well as processed information;
- each *scp-program* is a generalized structure in sc-memory, each time a *scp-program* is called, an independent *scp-process* is created on its basis.

The main advantages of the SCP language, due to these features:

- simultaneously several independent processes can be executed in the common memory, and processes corresponding to the same *scp-program* can be executed on different servers, in case of distributed implementation of the sc-model interpretation platform.
- *SCP language* allows concurrent asynchronous subroutine calls (creating subprocesses within scp-processes), as well as concurrently execute of scp-operators within a single *scp-process*;
- since *scp-programs* are written using *SC-code*, the transfer of the *sc-agent* implemented with the SCP language from one system to another supposes a simple transfer of the knowledge base fragment, without any additional operations, which depends on the sc-model interpretation platform;
- the fact that sc-agents' specifications and their programs can be written in the same language as the knowledge they are processing significantly reduces the list of specialized tools, designed to build and modify problem solvers and simplifies their development by using more universal components;
- the fact that a unique *scp-process* is created for the *scp-program* interpreting allows to optimize the execution plan as much as possible before its implementation and even directly during the execution without potential danger to break the general algorithm of the entire program. Moreover, this approach to programs designing and interpreting allows to talk about the possibility of creating self-reconfigurable programs;

III. Semantic model of problem solver

Using the concepts discussed above, we will say that the *sc-model of the integrated problem solver* is a non-atomic abstract sc-agent, which is the result of combining of all abstract sc-agents within a particular ostis-system into one. In other words, the *sc-model of the integrated problem solver* is the collective of all sc-agents within a given ostis-system, considered as a single whole.

Formally, the semantic model of the integrated problem solver is given as follows:

$$M_{IPS} = \{AG_{NA}, AG_A, AG_R\}, \qquad (3)$$

where AG_{NA} – a set of non-atomic abstract sc-agents within the solver;

AG_A – a set of atomic abstract sc-agents within the solver;

AG_R – a set of concepts specifying abstract sc-agents within the solver, including those describing the decomposition of non-atomic agents into atomic ones;

There are several basic levels of detail for the problem solver:

- level of the solver itself;
- level of non-atomic sc-agents within the solver, including particular solvers;
- level of atomic sc-agents;
- level of scp-programs or programs implemented at the level of *sc-model interpretation platforms*.

Such a level hierarchy, firstly, provides the possibility of step-by-step design of task solvers with a gradual increase of the detail level from the upper to the lower, and secondly, the possibility of designing, debugging and verifying components at different levels independently from other levels, which significantly simplifies the task of solvers construction and modification due to lower overhead costs.

In addition, the proposed approach to the construction of a solver model allows to provide modifiability of solvers and the possibility of consistent use of different problem solving models within a single solver.

IV. Semantic model of parallel processes interaction in the common semantic memory

Taking into account the models considered earlier, a formal model of information processes interaction in semantic memory is constructed, which is defined as follows:

$$M_{IPM} = \{M_A, M_S, M_{SYNC}, M_{SCP}\}, \qquad (4)$$

where M_A – model of the activity performed by agents in semantic memory;

M_S – model of sc-agent that performs transformations in semantic memory;

M_{SYNC} – model of processes execution synchronization in the semantic memory;

M_{SCP} – a model of a basic programming language, oriented to the processing of unified semantic networks, which, in turn, is defined as:

$$M_{SCP} = \{M_P, M_I\}, \quad (5)$$

where M_P – model of basic programming language program; M_I – model of basic programming language programs interpreter;

To synchronize the execution of *processes in sc-memory*, the lock mechanism is used. The *lock** relation connects the signs of the *processes in the sc-memory* to the signs of the situational *structures* that contain sc-elements that are locked for the duration of this process or for some part of this period. Each such *structure* belongs to any of the *lock types*.

In the current version, three *lock types* are distinguished to synchronize the processes execution in the sc-memory:

- *full lock*;
- *lock on any change*;
- *lock on delete*.

In turn, from the point of view of synchronization tools, three classes of *sc-agents* can be distinguished:

- textit sc-agents of scp-programs interpreting, which are implemented at the sc-models interpretation platform level and one of tasks of which is to provide the described synchronization mechanism. In turn, the principles of synchronization of agents of this class are more trivial than in the case of *program sc agents*, and are described separately.
- textitprogram sc-agents, providing the main functionality of the system, that is, its ability to solve certain tasks and working in accordance with the considering mechanism.
- *sc-metaagents*, task of which is to coordinate the activity of *program sc-agents*, in particular, to solve the deadlocks problem.

For more details on the locks mechanism in semantic memory, see [38], [39].

Thus, each sc-agent can generally correspond to several concurrently running processes in the sc-memory, the interaction of which is regulated by the described locks mechanism.

An example of a description of locks in semantic memory that correspond to several processes in this memory is shown in the figure 7.

V. Method for problem solvers constructing and modifying

As mentioned above, all platform-independent components of problem solvers can be represented using SC-code. In this case, we are talking about the specifications of sc-agents, and the full texts of scp-programs that describe the algorithms of these agents.

Thus, the construction of the ostis-system problem solver is reduced to the development of a special kind of knowledge base fragment of such a system. In this regard, when constructing and modifying solvers, all existing automation tools for the knowledge bases construction and modifying using OSTIS Technology can be used, considered, in particular, in the work [11].

Due to that, when constructing a method and means for solvers developing, it is necessary to clarify only some aspects of the development that are specific to the problem solvers based on the model considered earlier. It is important to note that, according to the model presented earlier, the task solver is an *abstract sc-agent*, and therefore the development of the solver is reduced to the such an agent development.

To develop problem solvers on the basis of the solver model considered, a method is proposed that assumes the use of formal ontology of such solvers developers activity for solvers development and is oriented to the use of reusable solver components at each level of the structural hierarchy of the solver being developed, which makes it possible to reduce the complexity of their development.

The proposed method includes several stages (figure 8):

The formally proposed method for solvers constructing and modifying is defined as follows:

$$M_{PA} = \{PA_C, PA_R\}, \quad (6)$$

where PA_C – a set of classes of actions performed by the solver developers;

PA_R – set of relations specifying these actions, including relations, specifying the order of actions implementation and decomposition of some actions into sub-actions.

The main advantage of the proposed method is its orientation to the ontology of the solver developers activity, which, on the one hand, will allow to automate this activity, and on the other hand, will allow to present the specifications of this activity within the knowledge base of the IMS metasystem and thus provide information support to the solvers developers.

VI. Tools for problem solvers constructing and modifying

To implement the proposed method, tools were developed to automate the process of problem solvers construction and modifying (TAPCM). The tools include a system for automating the process of solvers constructing and modifying, which in turn is conditionally divided into an automation system for the agents development and automation system for scp-programs development, as well as the library of solvers reusable components as part of the IMS metasystem for ostis-systems design support. Schematically, the architecture of the tools is shown in the figure 9:

In its turn, the library includes:

- the set of problem solvers components;
- means for specifying the problem solvers components;
- search tools for problem solver components based on their specification;

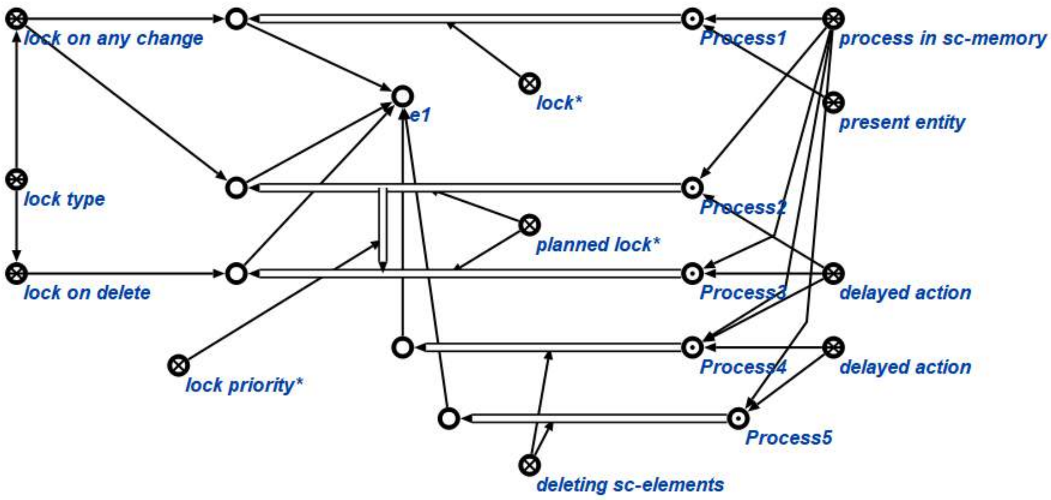


Figure 7. Example of the specification of locks in semantic memory

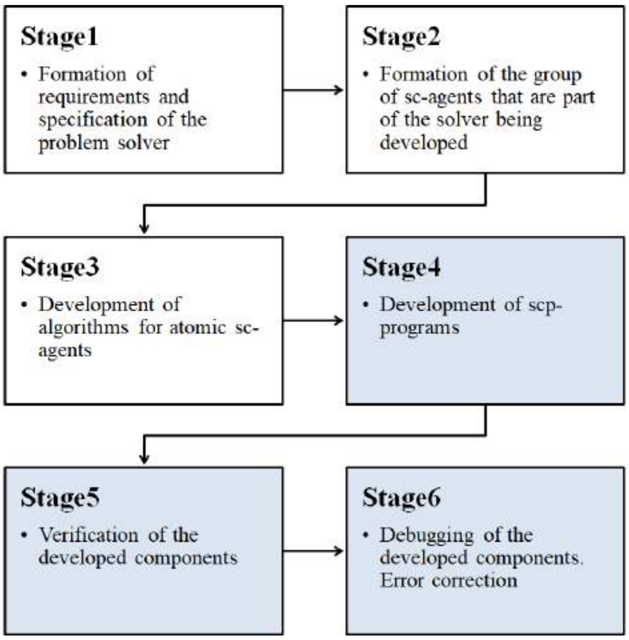


Figure 8. Stages of the method for problem solvers constructing and modifying

General structure of *Library of problem solvers reusable components* in the language SCn:

Library of problem solvers reusable components
= *problem solvers reusable component*
<= *subdividing**:
{
• *Library of reusable problem solvers*
• *Library of reusable atomic abstract sc-agents*
• *Library of reusable programs for sc-texts processing*
}

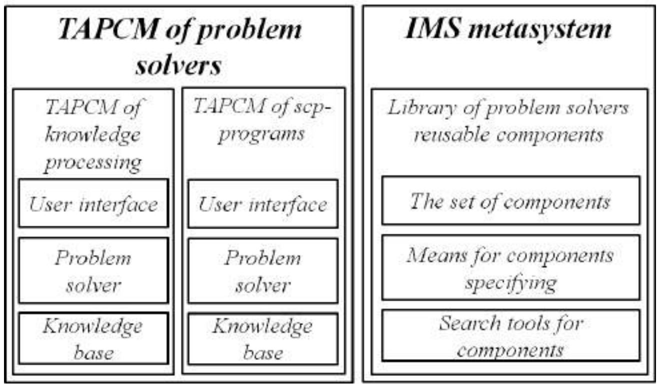


Figure 9. The architecture of tools for problem solvers constructing and modifying

In turn, the *Library of reusable abstract sc-agents* has the following structure:

Library of reusable abstract sc-agents
= *reusable abstract sc-agent*
<= *subdividing**:
{
• *Library of information search sc-agents*
• *Library of sc-agents of integrable knowledge immersion into the knowledge base*
• *Library of sc-agents for aligning the ontology of integrable knowledge with the basic ontology of the knowledge base current state*
• *Library of sc-agents for planning solutions to explicitly formulated tasks*
• *Library of logical inference sc-agents*
• *Library of sc-models of high-level programming languages and their interpreters*
• *Library of sc-agents of knowledge base verification*

- *Library of sc-agents of knowledge base editing*
- *Library of sc-agents for automation of activity of knowledge-base developers*

}

The solvers debugging tools developed provide the possibility of debugging on two levels of the solver:

- debugging at the sc-agents level;
- debugging at the scp-program level;

In the case of debugging at the sc-agents level, the act of each agent executing is considered to be indivisible and can not be interrupted. At the same time, debugging of atomic sc-agents and non-atomic agents can be performed. Initiation of an agent, including one that is part of a non-atomic one, is carried out by creating appropriate structures in the sc-memory, so debugging can be performed at different levels of agent detail, up to atomic levels.

Debugging at the level of sc-agents involves the ability to set or release the locks, enable or disable any agents, etc. Taking into account that the model of agent interaction proposed in this paper uses a universal variant of interaction of agents through common memory, the considered system of agent design support can serve as a basis for modeling systems of agents using other principles of communication [40], [41], for example, direct exchange of messages between agents.

Debugging at the level of scp-programs is carried out in a manner similar to existing modern approaches to procedural programs debugging and assumes the possibility of setting breakpoints, step-by-step program execution, etc.

The main feature of the considered tools for solver constructing and modifying is their implementation on the basis of OSTIS Technology, that is, including the use of previously considered solver model for the tools constructing. This feature allows to ensure the modifiability of the tools themselves, i.e. ease of their functionality expansion, including by using components from the permanently extending library of solvers reusable component.

VII. Conclusion

The proposed models, method and tools have been successfully applied to the development of problem solvers for intelligent reference systems in various academic disciplines, such as geometry, graph theory, history, chemistry, as well as in the development of a prototype automation system for batch production enterprises [42]. On certain examples, the proposed approach showed such advantages as the universality of the developed agents and the ease of modification of solvers constructed on the basis of the proposed model.

Acknowledgment

This work was supported by BRFFR (BRFFR-RFFR №Ф16Р-102, BRFFR-SFFRU №Ф16К-068).

Список литературы

[1] A. Kolesnikov, I. Kirikov, and S. Listopad, Gibridnye intellektual'nye sistemy s samoorganizatsiei: koordinatsiya, soglasovannost', spor [Hybrid intelligent systems with self-organization: coordination, consistency, dispute]. M.: IPI RAN, 2014, (in Russian).

[2] O. Castillo, P. Melin, and J. Kacprzyk, Recent Advances on Hybrid Intelligent Systems. Springer, 2014.

[3] S. Nikolenko and T. A.L., Samoobuchayushchiesya sistemy [Self-learning systems]. M.: MTsNMO, 2009, (in Russian).

[4] C. Kaner, J. Falk, and H. Nguyen, Testing Computer Software, 2nd ed. Wiley, 1999.

[5] E. Efimov, Reshateli intellektual'nykh zadach [Intelligent problem solvers]. M.: Nauka, 1982, (in Russian).

[6] A. Podkolzin, Komp'yuternoe modelirovanie logicheskikh protsessov. Arkhitektura reshatelya i yazyki reshatelya zadach [Computer modeling of logical processes. Solver architecture and problem solver languages]. M.: Fizmatlit, 2008, (in Russian).

[7] (2017, Jun.) Wolfram alpha. [Online]. Available: http://www.wolframalpha.com/

[8] A. Vladimirov, O. Varlamov, A. Nosov, and T. Potapova, "Programmnyi kompleks "UDAV": prakticheskaya realizatsiya aktivnogo obuchaemogo logicheskogo vvoda s lineinoi vychislitel'noi slozhnost'yu na osnove mivarnoi seti pravil [the program complex "UDAV": practical implementation of the active learning logical input with linear computational complexity on the basis of a miwar rule network]," Trudy nauchno-issledovatel'skogo instituta radio [Proceedings of the radio research institute], no. 1, pp. 108–116, 2010, (in Russian).

[9] V. Golenkov and N. Gulyakina, "Proekt otkrytoi semanticheskoi tekhnologii komponentnogo proektirovaniya intellektual'nykh sistem. Chast' 2: Unifitsirovannye modeli proektirovaniya [project of open semantic technology of component design of intelligent systems. part 2: Unified design models]," Ontologiya proektirovaniya [Ontology of design], no. 4, pp. 34–53, 2014, (in Russian).

[10] I. Davydenko, N. Grakova, E. Sergienko, and A. Fedotova, "Sredstva strukturizatsii semanticheskikh modelei baz znanii [Means of semantic models of knowledge bases structuring]," in Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Minsk.: BSUIR, 2016, pp. 93–106, (in Russian).

[11] I. T. Davydenko, "Ontology-based knowledge base design," in Open semantic technologies for intelligent systems (OSTIS-2017): materials of VII International.sc.-tech.conf. Minsk: BSUIR, 2017, pp. 57–72.

[12] D. Koronchik, "Realizatsiya khranilishcha unifitsirovannykh semanticheskikh setei [implementation of the unified semantic networks storage]," in Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Mn.: BSUIR, 2013, pp. 125–129, (in Russian).

[13] (2017, Jun.) IMS metasystem. [Online]. Available: http://ims.ostis.net/

[14] M. Wooldridge, An Introduction to MultiAgent Systems - Second Edition. Wiley, 2009.

[15] S. Russel and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Prentice Hall, 2010.

[16] V. Gorodetskii, M. Grushinskii, and A. Khabalov, "Mnogoagentnye sistemy (obzor) [Multi-agent systems (survey)]," Novosti iskusstvennogo intellekta [Artificial intelligence news], no. 2, pp. 64–116, 1998, (in Russian).

[17] V. Gorodetskii, V. Samoilov, and D. Trotskii, "Bazovaya ontologiya kollektivnogo povedeniya avtonomnykh agentov i ee rasshireniya [Basic ontology of autonomous agents collective behavior and its extension]," Izvestiya RAN. Teoriya i sistemy upravleniya [Proceedings of the RAS. Theory and control systems], no. 5, pp. 102–121, 2015, (in Russian).

[18] V. Tarasov, Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam [From multi-agent systems to intelligent organizations]. M.: Editorial URSS, 2002, (in Russian).

[19] V. F. Răzvan, Autonomous artificial intelligent agents. Romania: Center for Cognitive and Neural Studies (Coneural), 2003.
[20] (2017, Jun.) Autonomous agents and multi-agent systems. [Online]. Available: http://www.springer.com/computer/ai/journal/10458
[21] D. Weyns, A. Omicini, and J. Odell, "Environment as a first class abstraction in multiagent systems," ser. 1, vol. 14, February 2007, pp. 5–30.
[22] T. Finin, R. Fritzson, D. McKay, and R. McEntire, "Kqml as an agent communication language," in Proceedings of the third international conference on Information and knowledge management - CIKM '94, 1994.
[23] (2017, Jun.) FIPA ACL Message structure specification. [Online]. Available: http://www.fipa.org/specs/fipa00061/SC00061G.html
[24] R. L. Hartung and A. Hakansson, "Using meta-agents to reason with multiple ontologies," in Agent and Multi-Agent Systems: Technologies and Applications, 2008, pp. 261–270.
[25] W. W. Vasconcelos, M. J. Kollingbaum, and T. J. Norman, "Normative conflict resolution in multi-agent systems," ser. 2, vol. 19, June 2009, pp. 124–152.
[26] T. Rumbell, J. Barnden, S. Denham, and T. Wennekers, "Emotions in autonomous agents: comparative analysis of mechanisms and functions," ser. 1, vol. 25, June 2012, pp. 1–45.
[27] A. Omicini and F. Zambonelli, "Coordination for internet application development," Autonomous Agents and Multi-Agent Systems, vol. 2, no. 2, pp. 251–269, June 1999, issue 2.
[28] M. V. Nagendra Prasad and V. R. Lesser, "Learning situation-specific coordination in cooperative multi-agent systems," Autonomous Agents and Multi-Agent Systems, vol. 2, no. 2, pp. 173–207, 1999, issue 2.
[29] (2017, Jun.) GOAL agent programming language. [Online]. Available: https://goalapl.atlassian.net/wiki/spaces/GOAL/overview
[30] (2017, Jun.) GAMA Platform. [Online]. Available: http://gama-platform.org/
[31] (2017, Jun.) EVE - a web-based agent platform. [Online]. Available: http://eve.almende.com/index.html
[32] R. Evertsz, M. Fletcher, R. Jones, J. Jarvis, J. Brusey, and S. Dance, Implementing Industrial Multi-agent Systems Using JACK. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 18–48.
[33] (2017, Jun.) JAVA Agent DEvelopment framework. [Online]. Available: http://jade.tilab.com/
[34] O. Boissier, R. H. Bordini, J. F. Hübner, A. Ricci, and A. Santi, "Multi-agent oriented programming with jacamo," Sci. Comput. Program., vol. 78, no. 6, pp. 747–761, Jun. 2013.
[35] M. W. Rafael H. Bordini, Jomi Fred Hübner, Programming Multi-Agent Systems in AgentSpeak using Jason. Wiley, 2007.
[36] V. Jagannathan, K. Dodhiawala, and L. Baum, Blackboard Architectures and Applications. N.Y.: Academic Press, 1989.
[37] D. Shunkevich, A. Gubarevich, M. Svyatkina, and O. Morosin, "Formal'noe semanticheskoe opisanie tselenapravlennoi deyatel'nosti razlichnogo vida sub"ektov [Formal semantic description of the purposeful activity of various types of subjects]," in Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Mn.: BSUIR, 2016, pp. 125–136, (in Russian).
[38] D. Shunkevich, "Vzaimodeistvie asinkhronnykh parallel'nykh protsessov obrabotki znanii v obshchei semanticheskoi pamyati [Interaction of asynchronous parallel processes of knowledge processing in common semantic memory]," in Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Mn.: BSUIR, 2016, pp. 137–144, (in Russian).
[39] D. V. Shunkevich, "Ontology-based design of knowledge processing machines," in Open semantic technologies for intelligent systems (OSTIS-2017): materials of VII International.sc.-tech.conf. Minsk: BSUIR, 2017, pp. 73–94.
[40] C. M. Macal and M. J. North, "Tutorial on agent-based modeling and simulation," in Proceedings of the 2005 Winter Simulation Conference. WSC'05, 2005, pp. 2–15.
[41] M. Marietto, N. David, J. S. Sichman, and H. Coelho, "Requirements analysis of agent-based simulation platforms: State of the art and new prospects," in Third International Workshop of Multi-Agent-Based Simulation. Bologna, 2002, pp. 125–141.
[42] V. Taberko, D. Ivanyuk, V. Golenkov, K. Rusetskii, D. Shunkevich, I. Davydenko, V. Zakharov, V. Ivashenko, and D. Koronchik, "Proektirovanie predpriyatii retsepturnogo proizvodstva na osnove ontologii [Designing enterprises of batch production on the basis of ontologies]," Ontologiya proektirovaniya [Ontology of design], no. 2, pp. 123–144, 2017, (in Russian).

АГЕНТНО-ОРИЕНТИРОВАННЫЕ МОДЕЛИ, МЕТОД И СРЕДСТВА РАЗРАБОТКИ СОВМЕСТИМЫХ РЕШАТЕЛЕЙ ЗАДАЧ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Шункевич Д.В. (БГУИР)

Статья посвящена разработке агентно-ориентированных моделей, метода и средств разработки совместимых решателей задач интеллектуальных систем, способных решать комплексные задачи. Рассматриваются требования, предъявляемые к таким решателям, модель решателя задач, удовлетворяющего предъявленным требованиям, а также метод и средства разработки и модификации таких решателей.

Основной проблемой, решаемой в работе, является проблема низкой согласованности принципов, лежащих в основе реализации различных моделей решения задач. Как следствие, затруднена возможность одновременного использования различных моделей решения задач в рамках одной системы при решении одной и той же задачи, практически невозможно повторно использовать технические решения, реализованные в некоторой системе, фактически отсутствуют комплексные методы и средства построения решателей задач.

Как основу для построения решателей задач предлагается использовать многоагентный подход. Процесс решения любой задачи предлагается декомпозировать на логически атомарные действия, что позволит обеспечить совместимость и модифицируемость решателей. Сам решатель предлагается рассматривать как иерархическую систему, состоящую из нескольких взаимосвязанных уровней, что позволяет обеспечить возможность проектирования, отладки и верификации компонентов на разных уровнях.

Integration of artificial neural networks and knowledge bases

Golovko Vladimir Adamovich
Kroshchanka Alexsander Alexsandrovich
Brest State Technical University
Brest, Belarus
gva@bstu.by
kroschenko@gmail.com

Golenkov Vladimir Vasilyevich
Ivashenko Valerian Petrovich
Kovalev Mikhail Vladimirovich
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
golen@bsuir.by
ivashenko@bsuir.by
michail.kovalev7@gmail.com

Taberko Valery Vasilyevich
Ivaniuk Dzmitry Sergeevich
JSC "Savushkin product"
Brest, Belarus
id@pda.savushkin.by

***Abstract*—This article reviews the questions and directions of integration of artificial neural networks with knowledge bases. There are two main directions of integration:**

1) the inputs and outputs of artificial neural network to use integration of knowledge bases and artificial neural networks for solutions of application problems;

2) by artificial neural network representation on the basis of ontological structures and its interpretation by means of knowledge processing in the knowledge base providing an intelligent environment for the development, training and integration of different artificial neural networks compatible with knowledge bases.

The knowledge bases that are integrated with artificial neural networks are built on the basis of homogeneous semantic networks and multiagent approach to represent and process knowledge.

***Keywords*—ANN, knowledge base, integration, frameworks**

INTRODUCTION

At the moment, the idea that various areas of artificial intelligence should not develop in isolation has great acknowledgement. Fortunately, using synergetic approach, we can rely on not only the solutions of particular problems, but also on the achievement of ambitious goals such as the development of strong artificial intelligence systems.

The popularity of problem solving methods based on machine learning is inspired by the development of contemporary theoretical models of artificial neural networks and high-performance hardware platforms for their implementation. The variety of architectures, methods, directions and applications of artificial neural networks has accumulated and continues to be updated.

The complexity of contemporary intelligent systems that use artificial neural network models, as well as big data processing, require means of monitoring, understanding and explanation of the mechanisms of their work to verbalize and optimize their activity.

Therefore, it becomes actual to develop integration approaches for artificial neural networks and knowledge bases based on ontologies. Such integrated systems are able to combine the possibility of semantic interpretation of data that processed by an artificial neural network (ANN) including input and output data specification and the representation of ANN solvable applied problems with structure representation and the description of characteristics and states of artificial neural network making easier an understanding of its operation.

This article will describe the following questions:

- Is an artificial neural network a kind of knowledge and what kind if yes?
- What classes of problems are convenient to solve with the help of artificial neural networks, and which ones are not?
  - What classes of problems are convenient to solve by integrating artificial neural networks and knowledge bases?
  - What are the benefits of embedding an artificial neural network and its input and output data into knowledge base?
- How artificial neural network interacts with knowledge base?
- How to present artificial neural network in knowledge base?
- What is the structure of the domain of artificial neural networks?
- What agents are needed to process artificial neural networks in the knowledge base?

I. ARTIFICIAL NEURAL NETWORKS AND KNOWLEDGE

The appearance of artificial neural networks is associated with the works of W. McCulloch and W. Pitts [2], in which, however, preferred the term *neurons net*. The term *artificial neural network* has been fixed since the works of K. Fukushima [3]. Especial interest is the use of deep artificial neural networks for the formation of abstractions. These networks have an important advantage that distinguishes its from superficial models – the ability to form a complex hierarchy of characteristics (Figure 1 [4]). For the first time, an effective algorithm for training deep artificial neural networks was proposed by J. Hinton in 2006 [5]. Since that moment, a lot of work has been devoted to training this type of networks (for example, [6], [7]). Also, fundamentally new approaches were proposed([8], [9]).

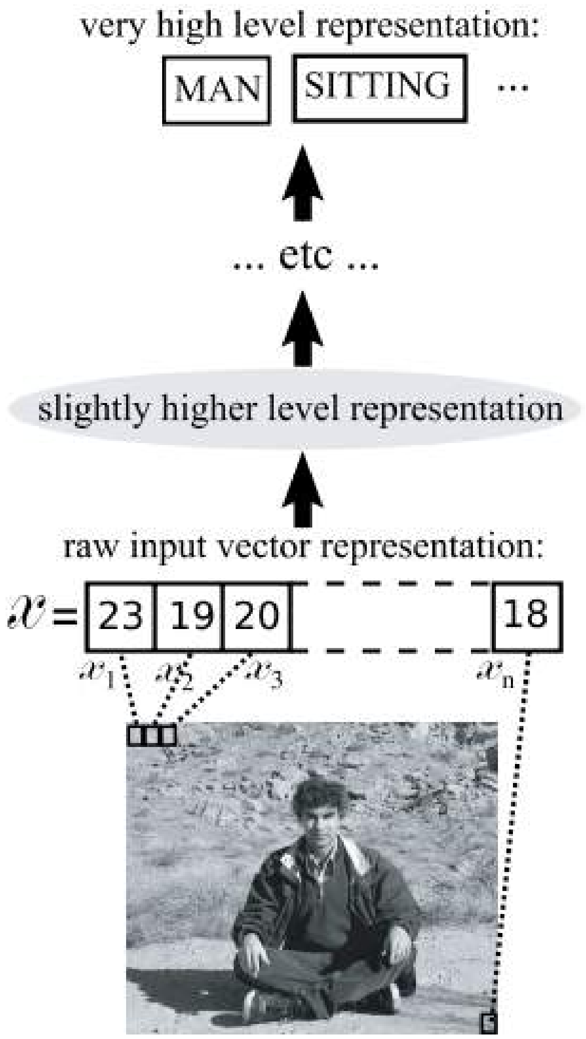


Figure 1. Hierarchy of features.

In this article, we consider artificial neural networks based on the concept of a formal neuron [10]. The formal neuron is specified by the composition of two functions:

1) function of synaptic transformation;

2) neuron activation function.

Artificial neural network that based on formal neurons is specified by four components:

1) set of vertices of artificial neural network (V).

2) set of edges of artificial neural network (E ($E \subseteq V \times V$)).

3) set of formal neurons, their properties and parameters of artificial neural network (N).

4) set of mapping between the union of vertex sets and edges of artificial neural network and the set of formal neurons, their properties and parameters of artificial neural network S

$$S \subseteq 2^{(V \cup E) \times N} \quad (1)$$

.

Artificial neural networks can be correlated with the concept of knowledge in accordance with different approaches to understanding knowledge. Correlation based on the [11] characteristics:

1) connectivity(artificial neural networks are based on a connected oriented graph, but not always strongly connected);

2) complex structure(there are artificial neural networks with a complex structure, consisting of layers set, vertices of artificial neural network and back connection);

3) interpretability(there are software implementations of artificial neural networks, where they are interpreted);

4) activity(software implementations of artificial neural networks have activity);

5) semantic metric (semantic metric in the general form and in the case when vertices of artificial neural networks is absent).

In accordance with IDEF-5 [12] knowledge is presented as relations, properties, kinds and attributes. In accordance with this, artificial neural networks have only numerical attributes and, possibly, sometimes kinds and relations, given with the help of predicates. However, in the general case, the predicates functions are not defined and are not known.

From the structural approach point of view, the artificial neural network is a mathematical structure(model) that corresponds to the principles of connectionism, which corresponds to the information process(becoming) that takes place in the technical(cybernetic) system expressed as a mapping or function (the dependence of the output data on the input).

From a pragmatic-mathematical point of view, the data processed by artificial neural network are elements of certain set, often interpreted as a sign set (image, attributes of some entity), and set itself as a feature(vector) space. There are binary, nominal, ordinal and quantitative features. Future is a numeric, quantitative attribute of an entity, an entity that can be expressed, specified by a mapping or function.

II. Classification of tasks that can be solved by artificial neural networks

Artificial neural networks are used to solve tasks. The task can be specified by the relation between the description of the initial situation and the description of a set of target situations.

Classify the tasks can be in accordance with different characteristics.

Since set of formal neurons are open, outline a class of tasks that are convenient to solve using artificial neural networks is problematic. Along with this, the accumulated experience shows that artificial neural networks are successfully used to solve tasks of the following types (see Table 1).

Table I
CLASSIFICATION OF ARTIFICIAL NEURAL NETWORK TASKS.

conceptually-pragmatic	**mathematical**
	approximation
processes management	- optimization
processes prediction	- extrapolation
images generation	
images matching detection of anomalous phenomena	
logical inference	logical inference
images translation	
- images clustering	- vector quantization
- images compression	- decrease in dimension
- linear division of images	- increase in dimension
- images association	
- images classification	
- images recognition	

Nevertheless, based on general methodological principles, it is possible to specify the signs of classes of tasks that are convenient to solve using artificial neural networks:

1) difficult-formalizable tasks, the solution of which has natural(vector) parallelism or data parallelism;

2) tasks, the solution of which is stable to the presence of NOT-factors in data and knowledge;

Also, there are list of possible signs of classes of tasks that are unprofitable to solve only with the help of artificial neural networks:

1) a complex conceptual description in task;

2) the presence of NOT-factors and the instability of solution to its;

3) the dominance of complexly described successive processes in the solution of the task;

4) need for introspective analysis and explanation of the obtained results in complex artificial neural networks.

Among the classes of tasks, it is possible to distinguish the classes of tasks that are conditionally specialized and the classes of tasks conditionally being abstract-fundamental [13].

III. USING OF INTEGRATION OF ARTIFICIAL NEURAL NETWORKS AND KNOWLEDGE BASES FOR SOLVING APPLIED TASKS

Intelligent systems with knowledge bases for solving applied problems with the help of artificial neural network algorithms can be used both for internal tasks, such as training neural network or optimizing its operation, and for solving a target problem, by processing input and output parameters of artificial neural network. Such intelligent systems can use artificial neural network methods on a par with other methods available in the system for solving only one, determined by system from subtasks during solving applied task.

Further, we will describe examples of use integration of artificial neural networks and knowledge bases for solving applied tasks.

A. Implementation of knowledge base integration into marking recognition system at JSC "Savushkin product"

One of the derived tasks at JSC "Savushkin product" is goods marking recognition task [14]. There is artificial neural networks [15], [16] based project to analyze image from camera, located above the production line (Fig. 2).

Figure 2. Production line with the camera unit

The unit recognize information from the bottles cover as string with the date of manufacture and the expiration date (Fig. 3).

Figure 3. Yogurt bottles with marked covers

This project is used by the engineer of Electrical Control & Instrumentation (EC&I engineer). The simplest way of integration into the bottle line system is to stop the line and inform about this the EC&I engineer using emergency sound and light alarm in case of detection n (e.g. 5) bottles with errors in recognition from the bottles cover. He should determine the cause of the problem and resolve it. In this case the module recognize and send the result as string to the knowledge base, where it should be processed according to this rules:

1) if the resulting string does not match the reference and it is only the one occurrence thus we should just discard this bottle (if there is a rejection device);

2) if the resulting string is empty and it is repeated for n occurrence consecutive bottles that is the ink has run out and we need to stop the line and notify the EC&I engineer to refill the printer;

3) if the resulting string is not empty but does not match the reference and it is repeated for n occurrence consecutive bottles that is nozzles are jammed and we need stop the line and notify the EC&I engineer to clean the printer unit.

Such rules can be a lot. In addition, we must be able to view and edit them (possibly even by EC&I engineer).

Since consumer packaging is constantly changing (and marking too), the next question arises: who and how will tune this system to a new kind of marking? Suppose today Arabic numerals are printed, and tomorrow Marketing department will decide to use Roman numerals. What needs to be done is to create a training set, train a new artificial neural network, launch the modified project. Such job can not be done by an EC&I engineer even if he wants (there are not enough knowledge in this area). But if there is some agent between them in the form of an intelligent system, then everything can look like this: the intelligent system receives information from artificial neural network that it is not possible to recognize the marking string, and from the user (engineer of the instrumentation and automation) information that everything in order and it was the launch a new type of labeling, then it can independently perform the tasks listed earlier to adapt the recognition module to this new type of labeling.

Proceeding from the above classification of tasks, this task can be classified as task of the detection of anomalous phenomena using the tasks of the images recognition. Also this task has signs of tasks described above, which are unprofitable to solve only with the help of artificial neural networks, because task has a dominance of difficult to describe successive processes in the solution. Using knowledge base integration allows to use artificial neural network algorithms to solve this task despite the presence of unprofitable factors for this.

B. Application of artificial neural network to support and update knowledge base

Today researchers proposed architectures, that consist from networks of different types. One of the ways of integration of artificial neural networks and knowledge bases lies in the use of such hybrid architectures, that form system for processing streams of knowledge to solve complex tasks.

An example of this structure is a hybrid network that uses convolutional neural networks (CNN) and recurrent LSTM networks (long short-term memory) [17].

This architecture (also known as LRCN – Long-term Recurrent Convolutional Network) is depicted in Fig. 4 [17]. Here are present in fact two variants such architecture, each of them allows to solve own task. First variant assumes averaging of results LSTM-network and forming compact description for depicted action (for example, HighJump). This variant can be consider as recognition of action, depicted on series of images. For second variant we need sequentially apply LSTM-network to form text description for image. Both variants uses two types of artificial neural networks. First type is convolutional neural network (CNN), which is used to process original image and form features. Second type is LSTM-network, which gets features and generates text description.

According to classification from section II, this is patterns recognition task (for first variant of model) and patterns generation task (for second variant of model).

The final text description can be analyzed, formulated into rules and integrated in the semantic network. Semantic network accumulates the knowledge in the form of specific objects and relations between them, and is able to perform a logical conclusion. In addition to image analysis, a similar architecture is used to generate text descriptions of the video stream.

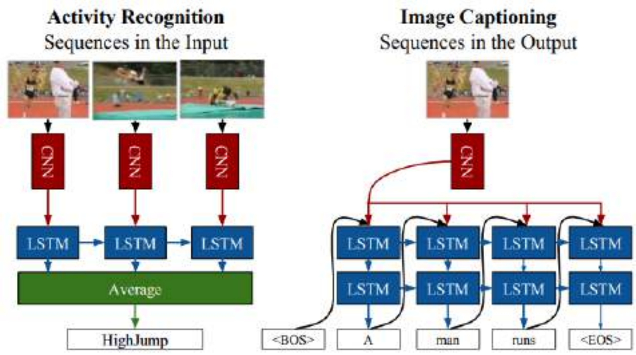


Figure 4. Long-term Recurrent Convolutional Network

Another use case is based on the word2vec approach [18]. This method uses for a semantic analysis of the text. The method is based on the use of simple shallow artificial neural network, which provides the formation of context based on one central word (skip-grams model – Fig. 5 [18]), or the formation of a central word based on the context (CBOW – Continuous Bag of Words model). The context is a set of words, which surrounding the central word and taken within a certain window. The trained network performs mapping of the data (words) from one-hot form to lower dimensional space, which then used to estimate the semantic similarity of words. The resulting embedding can be used for prediction semantic relations. For example, **king for queen** is also that **father for ?**.

Here is an example of the visualization obtained by us for a training set of 100000 English Wikipedia documents and vocabulary size of 50000 words. In this experiment used a simplified architecture of a skip-grams, including 50000 input neurons corresponding to the central word, 300 hidden and 50000 output neurons corresponding to the context word. Training set consist of pairs in form **(central_word, context_word)**, which fed to artificial neural network in mini-batches for 128 pairs in each. After training we applied t-SNE for dimensionality reduction of data. The resulting two dimensional map of semantic similarity depicted in Fig. 6.

Word2vec can be used to form knowledge base in certain subject domain (for example, [19]) and to extract semantic relations in common [20].

According to classification from section II, this is dimensionality reduction task.

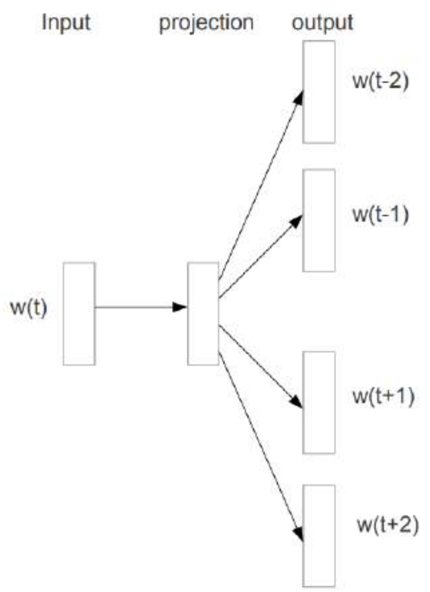


Figure 5. Skip-grams model

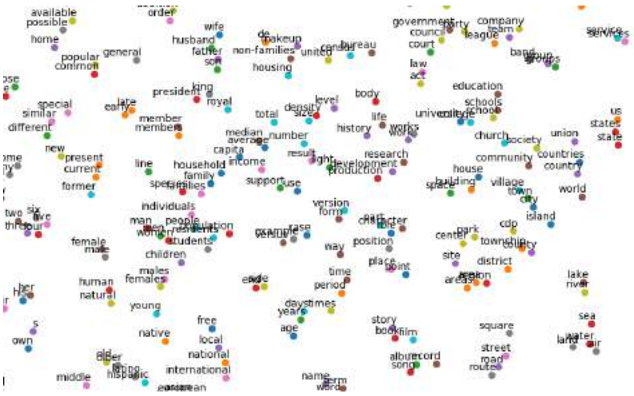


Figure 6. Map of semantic similarity

C. Using knowledge base for artificial neural network training

Interesting variation for integrating artificial neural networks and knowledge bases is to use knowledge base for artificial neural network training. With this approach, knowledge base comprise description of subject domain ontologies with in which task is solved, and with which you can describe artificial neural network inputs and outputs values. It also means the presence of knowledge-processing machine that can interpret the stored rules of **validation**(checking whether submerged value corresponds to certain conditions), checking for **non-consistency**(checking that new knowledge does not violate logical rules stored in knowledge base) and **adjustment**(replacing some knowledge with others) knowledge. This integration can be used to resolve the following tasks:

1) Validation and adjustment of training, test and examination samples. All elements of the samples are put to knowledge base and checked for consistency. Also in system can be stored a set of rules for validation and adjustment of features values of input elements.

2) Check consistency and correctness of artificial neural network output during training phase and implement appropriate adjustment of training. Training of artificial neural network is adjusted depending on intermediate results, which, with help of the reference to knowledge base, are checked for validity, consistency and plausibility. Also, result can be adjusted, for which knowledge base must have appropriate rules. Depending on the result of call to knowledge base, the weights calculation algorithm can be changed, which will allow training algorithms to be expanded by directly taking into account semantics.

3) Validation and adjustment of the results of trained artificial neural network. Artificial neural network outout can also be checked for validity, consistency and plausibility, and also adjusted with the help of rules used during training , and with the help of another specialized set of rules.

For resolving these tasks, knowledge base of a certain intelligent system can be used, as well as knowledge base, which has been specially developed for training particular artificial neural network.

Let us consider an example of classification task. In general, the statement of classification task is next: there is a group of objects. Each object has **n** features, it is necessary to assign each object to one of the **p** specified classes. Accordingly, artificial neural network is trained on some sample of objects with known features, and depending on type of training(with or without a teacher), with a known belonging to the class.

The training process can be improved by describing object classes and its features in knowledge base. Then, on training stage, for example, without a teacher, artificial neural network classified object to class **j**. Since knowledge base stores a description of this class, it will be possible to check consistency of object features with description of class that was detected. And already knowing this, it will be possible to do conclusion about need and method of synaptic connection weights adjustment of artificial neural network, thus expanding training algorithm.

Let's take the conditional task of defining the functional text style. There is text and author. The text extracts such conventional features as average length of sentences, average length of paragraphs, presence of direct speech, frequency of using verbs and nouns. It is indispensable to determine which of the five traditional styles – official, scientific, publicistic, newspaper and belles-lettres – related to text. Input features are very conventional and are used only as an illustration of proposed method. Figure 7 shows a schematic representation of artificial neural network inputs and outputs that solves this task.

Each element of training sample set is store in knowledge base. Semantic neighborhood of immersed elements are expanded by already existing knowledge in system. Figure 8 shows a fragment of knowledge base that describes submerged sample of texts to determine its style. Its semantic neighbor-

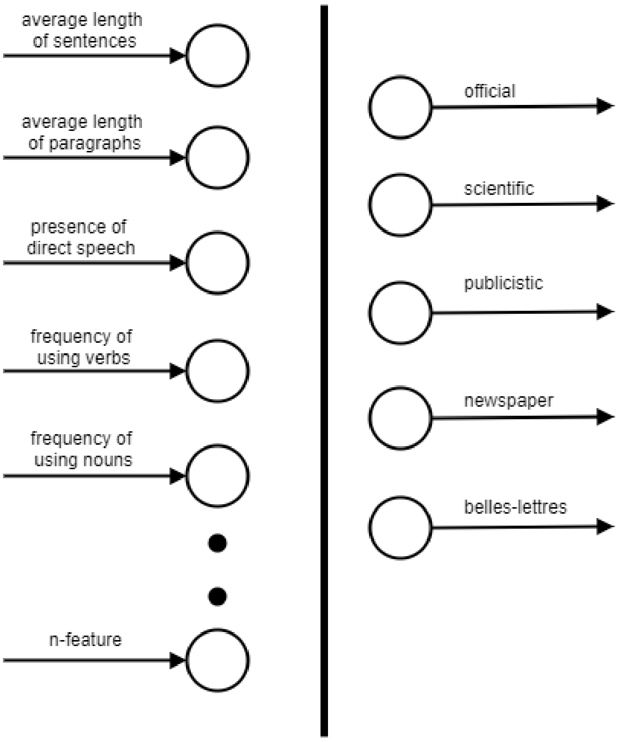


Figure 7. Schematic representation of artificial neural network inputs and outputs for solving text style determining task

hood is expanded by already existing in the system knowledge about text author and styles that he uses in its texts.

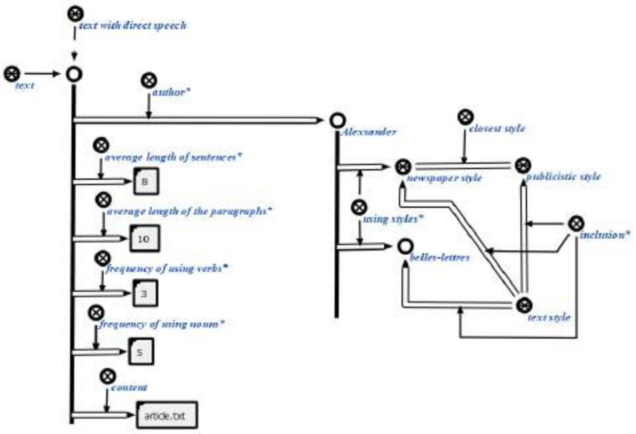


Figure 8. Knowledge base fragment that describes submerged sample of texts to determine its style

This information can be used to validate and adjust artificial neural network outputs during training. For example, during training, artificial neural network defined text style as publicistic. Knowledge base has validation rule, in which only style that author has previously used can be defined. Also, system has rule for adjustment the result, in which if text style that author has not previously used is defined, then it can be replaced with a similar one that author used. From knowledge base fragment that presented on Figure 8, system will be able to apply described rules of validation and adjustment, and replace the publicistic style to similar newspaper style. And depending on whether validation and adjustment rules were applied, as well as results of applying these rules, the algorithm of synaptic connection weight calculation in artificial neural network can change.

Described approach will help to increase degree of semantics influence during training, to discard or adjust deliberately unsuccessful sample elements, as well as to check artificial neural network outputs on plausibility.

IV. Intelligent environment of artificial neural network processing

A. *Review of existing frameworks*

The main drawback of all developed special frameworks is the requirement for knowledge about model structures, which should be used for each specific task. Despite the success of recent years, which reflects in the emergence of high-level frameworks with support of using graphics card (GPU) as hardware platform to calculations, and frameworks that are able to parallelize computations on different devices as a single computer or a whole cluster of computers [21], these developments are still only quantitative, but on the quality of such systems and their usability and comprehensibility for the end-user still need to perfect.

Despite the aforementioned drawback of existing frameworks, the threshold for occurrence in field of artificial neural networks, now more than ever low. Today, a large number of different libraries implement solutions based on intelligent algorithms. However, it is necessary to have the knowledge and skills to modify and improve the standard solution.

Here is a short overview of currently used open-source frameworks.

TensorFlow is one of the most popular libraries. Was developed by Google (2015). It allows to run training of models on several CPU and/or GPU devices. Available for different platforms, and supports different programming languages (C++, R, Python).

The main Tensorflow features includes

1) Multi-GPU support
2) Training across distributed resources (for example, on cloud)
3) Visualizing Tensorflow graph using TensorBoard – specific tool, which supports training process visualization and training data visualization
4) Model checkpointing – users of Tensorflow may stop training process and restart it from certain checkpoint.

Caffe/Caffe2 one of the first deep training libraries. It is written in C++, has a Python interface. Mainly oriented on training and the use of multi-layers and convolutional networks. It's created many pre trained networks for Caffe. In 2017 was proposed a new version Caffe2 by Facebook, which offers greater flexibility in building high-performance deep models. Can be used to develop architectures for run on mobile devices. It's well documented.

The Caffe features include the following:

- Trainable model is described in a special file with the extension "prototxt". In this file is stores the architecture of model and its main parameters.
- In the file with the same extension stores the parameters for the model training (number of training epochs, momentum, the training rate, weight decay, etc.)
- The training data are saved in files of special format (hdf5, lmdb) or in a text file with list of individual items (useful for images in separate files).
- The trained network is saved in a file with the extension "caffemodel". In the future it may be used directly to solve the basic tasks of machine training (classification, regression, etc.), and as an part of more complex models (for example, to solve the problem of detection of objects in images).
- To use the model, you need to have deploy-file that has the extension "prototxt", which, in fact, coincides completely with the file that contains the description of the model, except for settings that have a direct relation to the training (format of data, minimized function, etc.).

Theano was one of the first libraries that implement algorithms for deep training. It has some problems with scalability and computing on a cluster of graphics cards. Overall enjoyed sustained popularity with the experts in the field of deep training. It's often used as underlying framework for more high-level libraries that provide API wrappers (for example, Keras).

Keras provides a simplified interface to work with Theano, Tensorflow or CNTK. Very lightweight and easy to train and use. It's well documented. Allows you to create and train artificial neural network in a few lines of code. Written in Python.

Torch is framework based on Lua. It is possible to use libraries of C/C++ and CUDA. A very simple library, process of model building as simple as possible. There is a more modern implementation for Python, called PyTorch.

Additionally we need to refer on other frameworks, which are active use in last years (Fig. 9). These are Microsoft CNTK, MXNet and based on it Gluon, ONNX, released in September 2017 and presents open format to represent deep training models [22] and more others.

Number of frameworks is progressively to grow, but in fact each of them are completely repeat others and special differences include only additionally models and support of new hardware features.

B. The development of intelligent environment of artificial neural network processing

Intelligent environment is a environment with ability of automatic parameters selection of model (and, ideally, whole model) depending on tasks with minimal user involvement. This system allows people who are not specialists in the field of machine training and artificial neural networks, to successfully apply the latest developments in their activity.

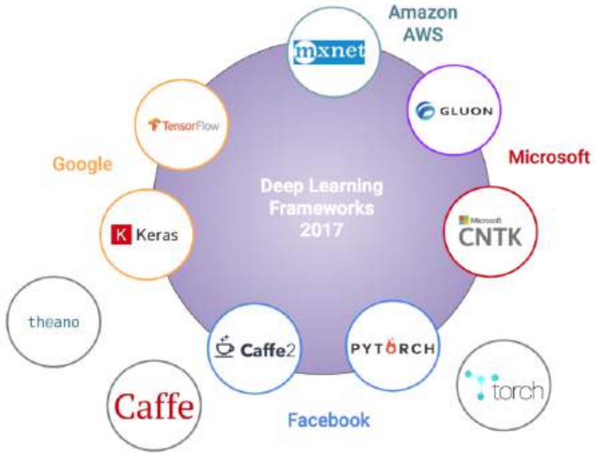


Figure 9. Open-source frameworks in 2017 [22]

Obviously the relevance of such technologies is high. However, to develop similar software systems need to determine the main criteria for the selection of a particular model. It is known that the choice of artificial neural network to solve the task — is creative and empirical process, as it involves getting results through the selection and evaluation of the effectiveness of different architectures of networks.

But already developed approaches for the automatic generation of artificial neural network models [23], which show the advantage in comparison to manual selection of parameters.

V. Subject domain of artificial neural networks and its denotational and operational semantics

The main part of the system that developed using OSTIS([14], [24]) technology is the ontological model (knowledge base sc-model), which is built on the basis of the ontological approach. This approach includes building of ontologies as systems of absolute and relative concepts that describe a particular subject domain. In the OSTIS technology, the concept of ontology is defined as a specification of the subject domain [1], [25], its typology is specified.

A. Representing neural networks in knowledge base

The proposed approach is based on the use of knowledge bases corresponding to the model of the unified semantic representation of knowledge. This model uses homogeneous semantic networks which are semantic networks based on the basic set-theoretic semantic interpretation. This interpretation built on the (situtational) (non-)member-of relation whose links are denoted by sc-arc's. This relation links elements to a set and are relation of single basic type only [1]. The languages that included in the model of the unified semantic representation of knowledge are called sc-languages, the texts of which consist of sc-elements, and formed from its situative sets, structures and ontological representations, ontological models are called sc-sets, sc-structures and sc-models. The languages semantics of the unified model of knowledges semantic representation corresponds to the model of situative sets. Situative sets represent a more flexible and adequate apparatus for representing knowledge than classical sets, allowing to consider NOT-factors of knowledge and to adapt to the represented problem area while preserving the ontological model and its semantics. This is achieved not only by considering temporal properties, which makes it possible to interpret the system of situative sets as the development of systems of L-fuzzy sets, but also due to the fact that within the model of unified semantic representation of knowledge the process of using the apparatus of situative sets can be dynamic that support by dynamic of alphabetic labels. Therefore, it seems expedient to use situative sets and its advantages to represent artificial neural networks using the model of unified semantic representation of knowledge. It should also be noted that semantics of the texts of sc-languages in the model of unified semantic representation of knowledge is a model one, but it is possible to describe it in denotational form, in this case we can talk about the denotational semantics of sc-languages on basis of situative sets. Thus, when representing artificial neural network, each node (vertex) that is not receptor can be treated as a denotation of the situative set (sc-set) of all its vertices from which the signal comes to this node (Fig. 10.).

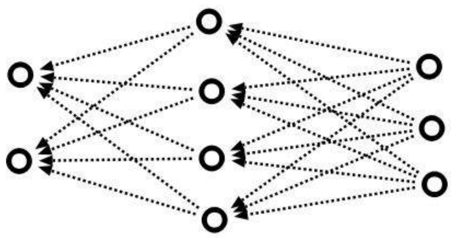


Figure 10. Representation of multilayer artificial neural network

Each receptor vertex is denotation of situative set of attributes (components of a features description) of the entities it expresses. Artificial neural network is represented as a situative set of sc-elements, denoting vertices, relations, its parameters, functions, properties, links and correspondence between them (Fig. 11, Fig. 12).

The synaptic connection weights and outputs of neural elements are represented as a situative measured value of a certain quantity that obtained by measuring the corresponding elements of artificial neural network (Fig. 13).

To specify the functions of the synaptic transformation and activation (Fig. 14), synaptic connection weights(Fig. 13) and outputs signals of neural elements, the key elements are used: *synaptic transformation function*, *activation function**, *synapse*, *neuron node* denoting situative relations.

Mathematical relationships can be established between these values, expressed through mathematical operations and relations, such as sum and product (Fig. 15).

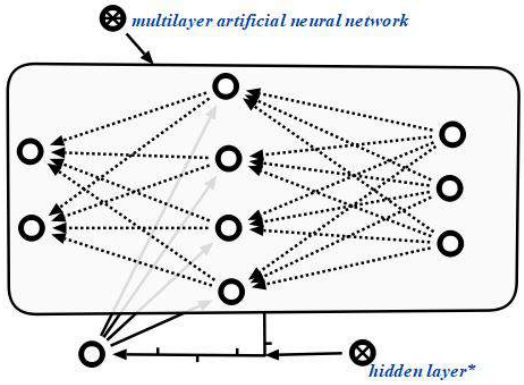


Figure 11. Representation of multilayer artificial neural network with a hidden layer

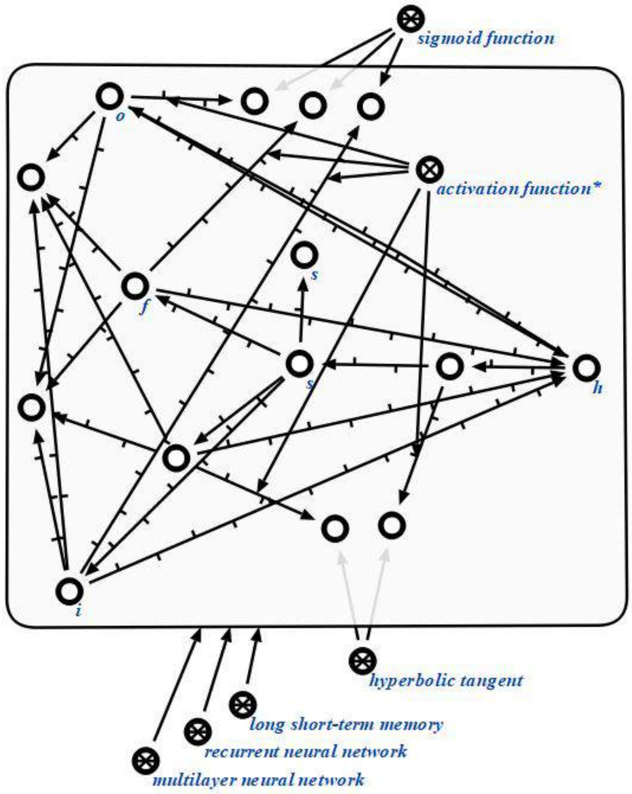


Figure 12. Long short-term memory representation

These mathematical relations can be given in general form by a sentence or program (Fig. 16).

Representation of artificial neural networks with complex structure makes it possible to represent its separate parts, such as hidden layers in multilayer artificial neural networks (Fig. 11), counter-connections in artificial neural networks of counter propagation (Fig. 17), reverse connections in recurrent artificial neural networks (Fig. 12) and etc.

As has already been said, each artificial neural network specifies a mapping between its inputs and outputs (Fig. 18, Fig. 19), a description of this fact is represented using the key

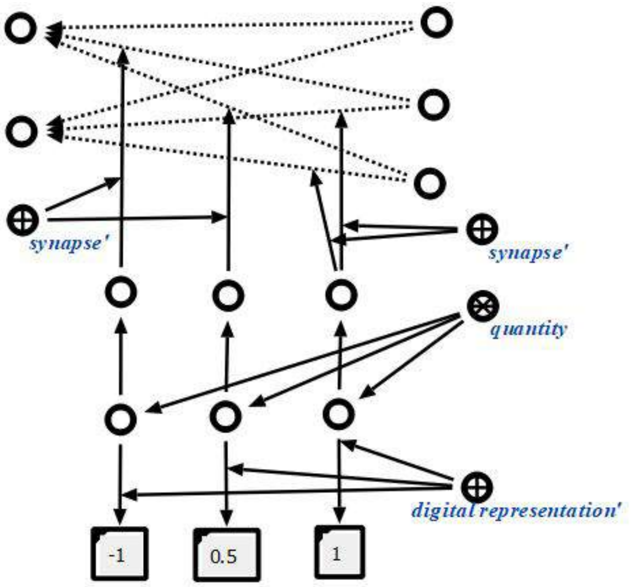


Figure 13. Representation of synaptic connection weights of artificial neural network

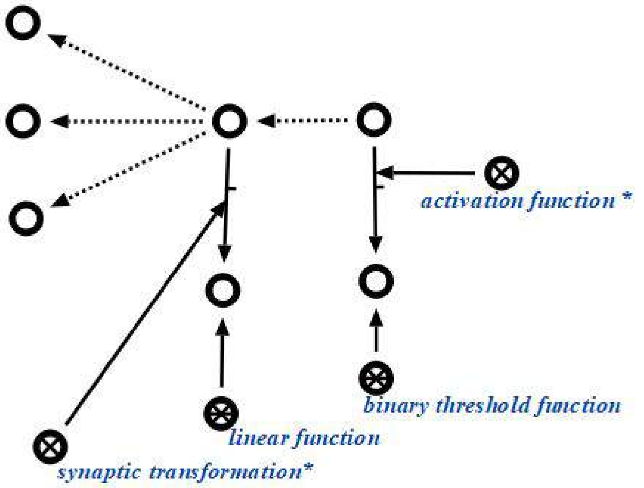


Figure 14. Representation of functions of synaptic transformation and activation of artificial neural network nodes

elements *mapping*, *domain of mapping*, *domain´*, *mapping of domain´*, *range of mapping*, *range´*, *mapping of range´*.

To clarify the values of input neural elements of artificial neural network, the key elements are *domain´*, *domain*, *relation´*. The listed key elements, along with others, are used for the specification of features(Fig. 20), whose values are the values of the input and output neural elements (receptors and effectors) of artificial neural network (Fig. 19).

B. The structure of subject domain

Subject domain is a structure that is defined on a set of sc-elements. On subject domains the same relations are defined as on sc-structures, which can be connected by inclusion relations (situative subset), isomorphic embedding, etc. These relations

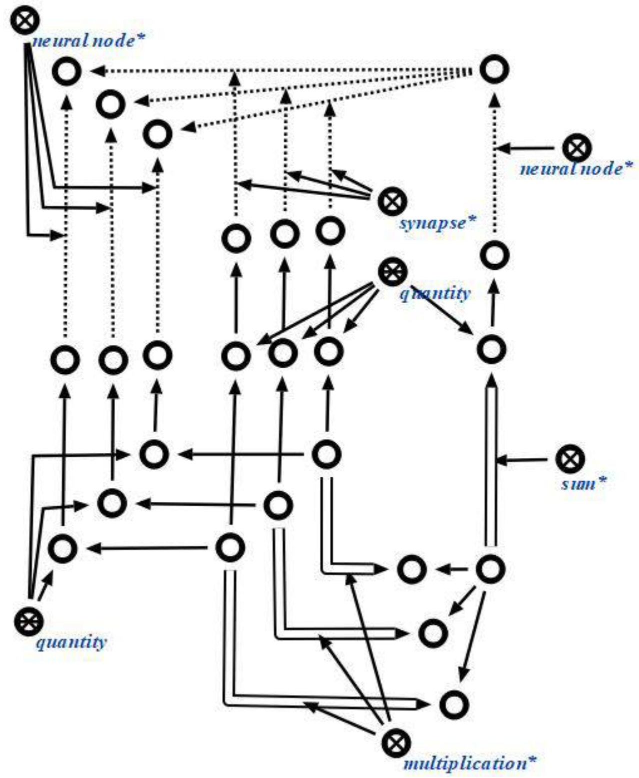


Figure 15. Representation of the output values of elements of artificial neural network and its mathematical relations

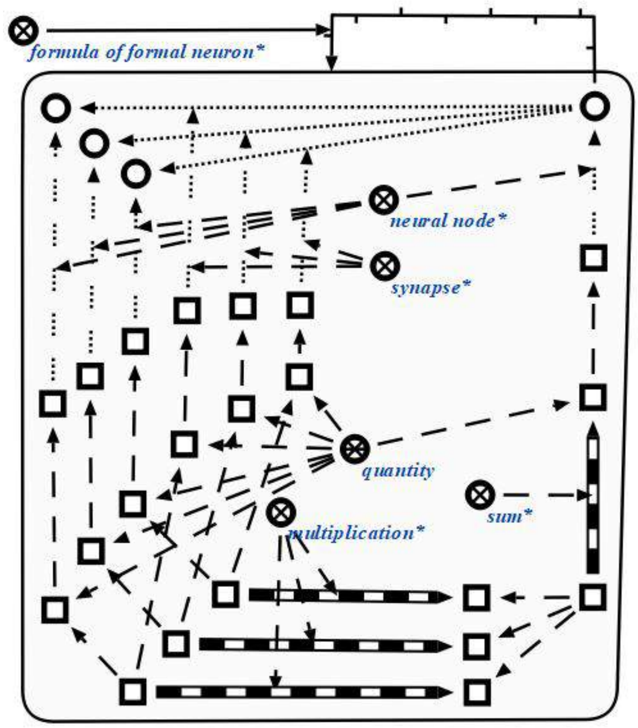


Figure 16. Representation of the program, a formal description of the dependence of the output values of the elements of artificial neural network

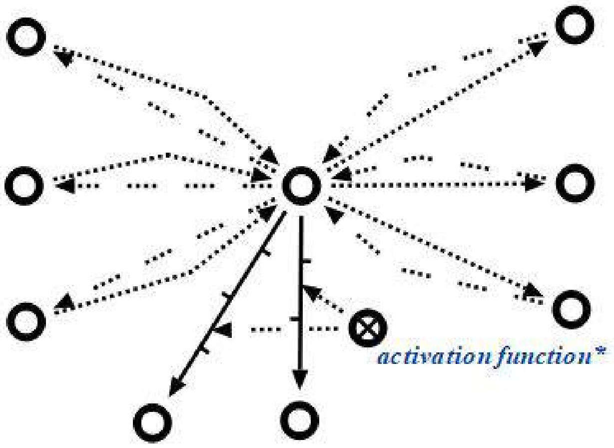


Figure 17. Representation of multilayer artificial neural network with a hidden layer

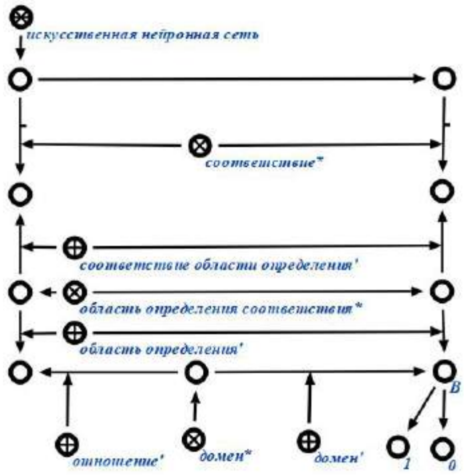


Figure 19. Representation of input artificial neural network element

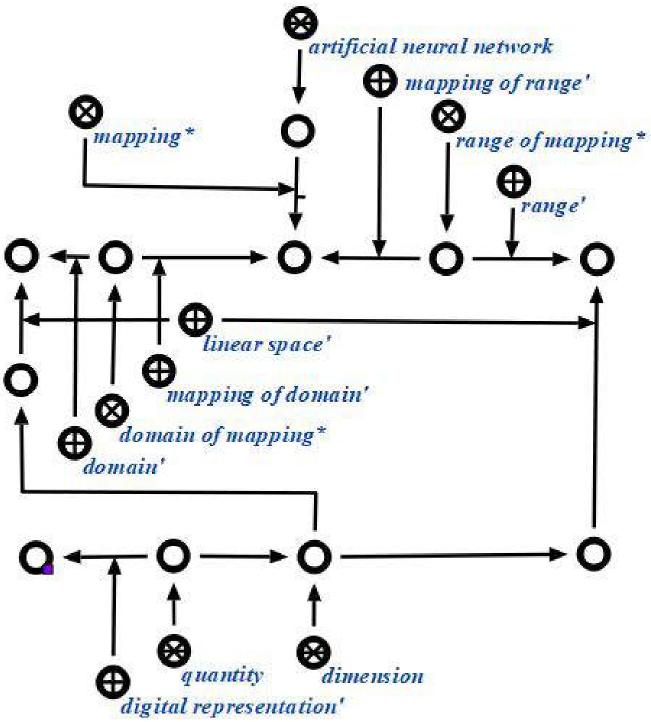


Figure 18. Representation of inputs and outputs of artificial neural network

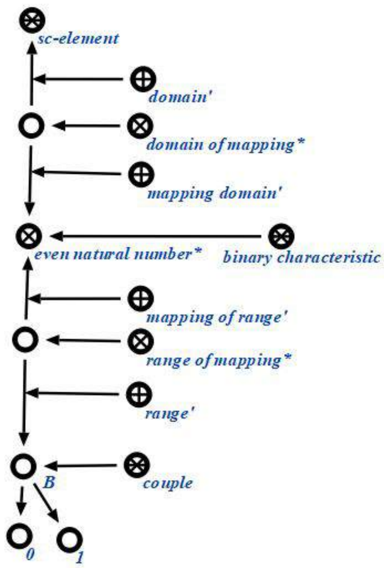


Figure 20. Representation of the binary feature

are analogous to the relations of the knowledge specification model [29]. It is necessary to distinguish these relations from situative relations of temporary inclusion, hypothetical inclusion, temporary isomorphic embedding and hypothetical isomorphic embedding, which are caused by the presence of such NOT-factors as incompleteness and uncertainty.

Subject domain **A** is a partial subject domain of subject domain **B** if it is included in **B** and its maximum investigated set is included in the maximal investigated set **B**.

The most important part in the development of the OSTIS knowledge base is the formation of a hierarchy of subject domains (based on the concept of a partial subject domain) that determines the structure of knowledge base and its specification [25]. Consideration of knowledge base from the position of the relationship with subject domain allows us to consider the objects under study at different levels of detailing, which are reflected in different types of ontologies that describe a certain direction of description the objects characteristic within the described subject domain. Such ontologies include: structural specification of subject domain, logical ontology, logical system of concepts and its definitions, set-theory ontology.

Artificial neural networks:

- Finite-valued artificial neural networks
  - •• Binarized artificial neural networks

- ••• Binary artificial neural networks
- ••• Bipolar artificial neural networks
- •• Ternary artificial neural networks
- Complex-numerical neural networks
  - •• Real-numerical artificial neural networks
  - ••• Rationally-numerical artificial neural networks
  - •••• Integer artificial neural networks
- Rational neural networks
- Dimension-increasing artificial neural networks
- Dimension-decreasing artificial neural networks
- Discontinuous artificial neural networks
- Continuous artificial neural networks
- Differentiable artificial neural networks
- Nondifferentiable artificial neural networks
- Homogeneous artificial neural networks
- Heterogeneous artificial neural networks
- Artificial neural networks without contextual neurons
- Artificial neural networks with contextual neurons
- Artificial neural networks with contextual connections
- Artificial neural networks without contextual connections
- Artificial neural networks with hidden neurons
- Artificial neural networks without hidden neurons
- Multilayer artificial neural networks
- Single-layer artificial neural networks
- Stochastic artificial neural networks
- Deterministic artificial neural networks
- Relaxation artificial neural networks
  - •• Relaxation networks with chaotic behavior
  - •• Relaxation networks with stable behavior

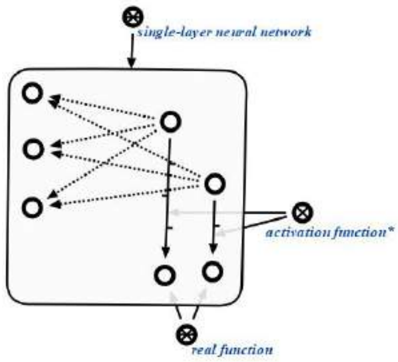


Figure 21. Single-layer artificial neural network

In addition to the distinguished classes of artificial neural networks (Figure 21, Figure 22.) and corresponding subject domains, it is possible to single out tasks, classes of artificial neural networks and corresponding subject domains associated with the training of artificial neural networks. There are two main classes:

- Training of artificial neural networks based on the target values of artificial neural network output elements.

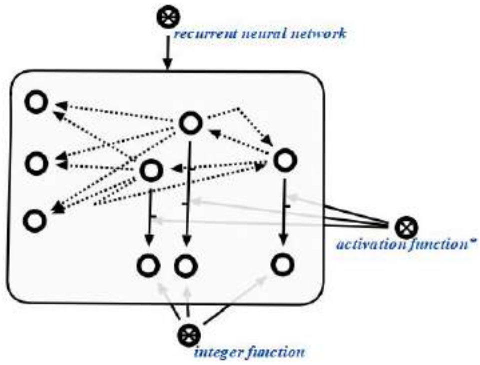


Figure 22. Recurrent artificial neural network

- Training of artificial neural networks without using the target values of artificial neural network output elements.

The first class includes subclasses:

- Training of artificial neural networks based on delta rule.
- Training of artificial neural networks based on gradient methods.
  - •• Training of artificial neural networks using the method of back propagation of an error.
  - •• Training of artificial neural networks by the conjugate gradient method.
  - •• training of artificial neural networks using the variable metric method.

The private classes of second class are:

- training of artificial neural networks without using target values based on the Hebb rule [26].
- training of artificial neural networks without using reference values based on the WTA rule [27].

In addition, the intelligent system may require knowledge of the following subject domains:

- Subject domain of the images and the features description of entities.
- Subject domain of set theory.
- Subject domain of logic.
- Subject domain of the natural numbers arithmetic(Peano).
- Subject domain of the integer number field.
- Subject domain of rational numbers field.
- Subject domain of dual rational numbers field.
- Subject domain of polynomial expressions.
- Subject domain of algebraic numbers field.
- Subject domain of dual algebraic numbers field.
- Subject domain of symbolic differentiation of polynomial expressions.
- Subject domain of functions and formal series.
- Subject domain of numerical sequences.
- Subject domain of common topology.
- Subject domain of graphs.
- Subject domain of measure theory.
- Subject domain of differentiable functions.
- Subject domain of linear spaces.
- Subject domain of tensor algebra.

- Subject domain of differentiable programming space.
- Subject domain of probability theory.
- Subject domain of cascades.
- Subject domain of cascades and dynamic systems.
- Subject domain of discrete optimization.
- Subject domain of optimization tasks.
- Subject domain of training artificial neural networks.

C. Agents of intelligent environment of artificial neural network processing

Operational semantics of knowledge bases that constructed in correspondence with model of unified semantic representation of knowledge is expressed in commands of knowledge processing sc-machine [28]. Each sc-machine corresponds to a formal information processing model, which language is some sc-language. Also, each sc-machine has initial information structure and set of operations that it implements, which can be programmed in procedures form.

In accordance with multiagent approach, each sc-machine can be implemented as a collection of agents (sc-agents). Operational semantics of artificial neural network in knowledge base is reduced to operational semantics of sc-agents that implements it and operational semantics of which, in turn, reduces to operational semantics of programs(commands) of its operations.

All operations of sc-agents are performed asynchronously, i. e. its are implemented in such a way that its joint work is reduced to its consistent work. All sc-agents interact via shared memory, passing each other data in form of semantic network constructs (sc-language texts that consist of sc-elements).

The condition for starting operation(initiating sc agent) is some event in the shared (graph) memory. Such events are changes in temporary non-belonging to temporal (actual) belonging of element to situative set, which is treated as set of commands for initiating sc-agents. Each command is data that will be processed by the agent. Such data can be a single sc-element (Figure 23) and its semantic neighborhood, available in shared memory, or some structure (sc-structure), denoted by such a sc-element (Figure 24).

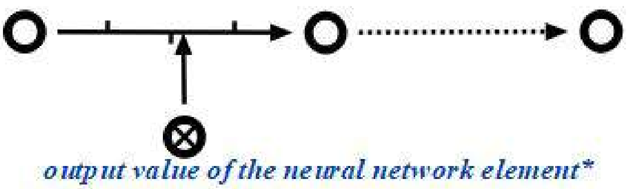


Figure 23. Data exchange between sc-agents

After operation start and complete, the temporary accessory is replaced with a temporary (actual) non-belonging (Figure 25), but new temporary (actual) accessories may appear that will initiate work of other sc-agents. Thus, work of artificial neural network and processes of processing knowledge in knowledge base are reduced to certain order of changes in temporal belonging to non-belonging and vice versa. Agents that provide semantic logging [30] artificial neural networks work form meta-descriptions as structures of special kind that provide ability to search and analyze order of artificial neural networks operations, which reduces to the interaction of sc-agents in shared graph memory.

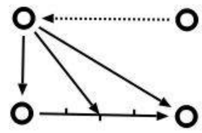


Figure 24. Transmission of data as a structure

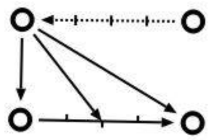


Figure 25. Processed data as a structure

For implementation of artificial neural networks, following agents classes was distinguished:

1) Agents for interpretation and processing of artificial neural networks inputs and outputs:
   - Agents of receptoral signals recording
   - Agents of receptoral signals validation and adjustment
   - Agents of effectoral signals reading
   - Agents of effectoral signals validation and adjustment
   - Agents of interpretation artificial neural networks nodes
   - Agents of interpretation artificial neural networks layers
   - Agents of semantic logging of artificial neural networks work
   - Agents of validation semantic logging of operation of artificial neural networks operations
2) Agents of training of artificial neural networks
   - Agents of artificial neural network training management
     - Agents of stochastic values generation
     - Agents of global training parameters changing
     - Agents of errors calculations
     - Agents of samples management
   - Agents of training processes semantic logging
3) Agents of integration of different artificial neural networks
   - Agents of cloning of artificial neural network
   - Agents of artificial neural network receptors search
   - Agents of artificial neural network effectors search

- Agents of synthesis of artificial neural network
  - Agents of synthesis of artificial neural network layers
- Agents of synthesis of multiple neurons agents that synaptically connected to neurons set
  - Agents of synthesis of multiple neurons agents that synaptically completely connected to neurons set
  - Agents of synthesis multiple neurons agents that synaptically incompletely connected to neurons set
- Agents of synthesis agents of backpropagation dual network
- Agents of synaptic connections deleting
- Agents of artificial neural network deleting
- Agents of translating ontological representation of artificial neural network to programming languages
- Agents of semantic logging of artificial neural networks integrations

CONCLUSION

The considered directions of integration of artificial neural networks with knowledge bases will help to solve more high-level tasks, making the solution of these tasks more structured and transparent. The implementation of the described intellectual system for the theory of artificial neural networks, as well as intelligent environment, helps to reduce the requirements to the skills of developers for methods of solving tasks using artificial neural networks. The possibility of artificial neural network introspection provided by the intelligent environment with the ability to memorize the state of artificial neural network during learning allows a deeper analysis of its work.

REFERENCES

[1] Horoshevskij B.F., Gavrilova T.A., Bazy znanij intellektual'nyh sistem(Intelligent Systems Knowledge Base), St.Petersburg, Piter, 2001, 384 p.

[2] McCulloch W., Pitts W. A Logical Calculus of Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, 1943, pp. 115–133.

[3] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, Biological Cybernetics, 1980, pp. 193-202.

[4] Bengio, Y. Learning deep architectures for AI. Foundations and Trends in Machine Learning, Foundations and Trends® in Machine Learning, 2009, Vol. 2: No. 1, pp 1-127

[5] Hinton G.E., Osindero S., Teh Y. A fast learning algorithm for deep belief nets, Neural Computation, 2006, pp. 1527–1554.

[6] Glorot, X., Bordes, A., Bengio, Y. Deep sparse rectifier neural networks, 14th International Conference on Artificial Intelligence and Statistics, 2011, vol. 15, pp. 315–323.

[7] Golovko V.A., Kroshchanka A. A., Rubanau U., Jankowski S., Learning Technique for Deep Belief Neural Networks, Communication in Computer and Information Science, 2014. vol. 440. . – pp. 136-146.

[8] Golovko, V.A. Deep learning: an overview and main paradigms // Optical Memory and Neural Networks (Information Optics), 2017, Volume 26, Issue 1, pp 1–17.

[9] Golovko, V. The nature of unsupervised learning in deep neural networks: A new understanding and novel approach / V. Golovko, A. Kroschanka // Optical memory and neural networks. – 2016. – No 3. – p.127–141.

[10] Rosenblatt F., Principles of neurodynamics; perceptrons and the theory of brain mechanisms, Washington Spartan Books, 1962, pp. 24-25.

[11] Pospelov D.A. Iskusstvennyj intellekt: V 3 knigah. Kniga 2. Modeli i metody: Spravochnik(Artificial Intelligence: In 3 books. Book. 2. Models and methods: Handbook), Moscow, Radio i svyaz', 1990, 304 p.

[12] Information Integration for Concurrent Engineering (IICE), IDEF5 Method Report, Texas, USA, Knowledge Based Systems, Inc., 1994.

[13] Graves A., Wayne G., Reynolds M., Harley T., Danihelka I., Grabska-Barwińska A., Colmenarejo S.G., Grefenstette E., Ramalho T., Hybrid computing using a neural network with dynamic external memory, Nature, 2016, pp. 471–476.

[14] Golenkov, V. V. Ontology-based Design of Batch Manufacturing Enterprises, Open Semantic Technologies for Intelligent Systems, 2017, pp. 265-280.

[15] Diederik P. Kingma, Jimmy Lei Ba., Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG], December 2014.

[16] Klambauer G., Unterthiner T., Mayr A., Hochreiter S., Self-normalizing neural networks, CoRR, arXiv:1706.02515, 2017

[17] Donahue J., Long-Term Recurrent Convolutional Networks for Visual Recognition and Description, in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, vol. 39, no. 4, pp. 677-691.

[18] Mikolov T., Chen K., Corrado G., Dean J., Efficient estimation of word representations in vector space, ICLR conference, arXiv preprint arXiv:1301.3781.

[19] Xiong S., Wang X., Duan P., Yu Z. and Dahou A. Deep Knowledge Representation based on Compositional Semantics for Chinese Geography , In Proceedings of the 9th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART, 2017, pp. 17–23.

[20] Pelevina M., Arefyev N., Biemann C., Panchenko A. Making Sense of Word Embeddings, 2017, arXiv preprint arXiv:1708.03390.

[21] Goldsborough P., A Tour of TensorFlow, ICLR conference, 2016, arXiv preprint arXiv:1610.01178.

[22] Battle of the Deep Learning frameworks – Part I: 2017, even more frameworks and interfaces. Available at: https://towardsdatascience.com/battle-of-the-deep-learning-frameworks-part-i-cff0e3841750 (accessed 2017, Dec)

[23] Liu H., Simonyan K., Vinyals O., Fernando C., Kavukcuoglu K., Hierarchical Representations for Efficient Architecture Search, ICLR conference, 2017, arXiv preprint arXiv:1711.00436.

[24] IMS. Available at: http://ims.ostis.net (accessed 2017, Nov)

[25] Davydenko I.T., Tekhnologiya komponentnogo proektirovaniya baz znanij na osnove unificirovannyh semanticheskih setej(The technology of component design of knowledge bases on the basis of unified semantic networks), Open Semantic Technologies for Intelligent Systems, 2013, pp. 185-190.

[26] Hebb D.O., The Organization of Behavior, New York, Wiley & Sons, 1949, 383p

[27] Grossberg S., Contour enhancement, short-term memory, and constancies in reverberating neural networks, Studies in Applied Mathematics, 52:213, 1973

[28] Shunkevich, D.V, Models and Means of Component Design of Knowledge-Processing Machines on the Basis of Semantic Networks. In: Open semantic technologies for designing intelligent systems (OSTIS2013): materials of the International Science and Technology Conf. Minsk, February 21-23, 2013) – Minsk: BSUIR, 2013, – P. 269–280

[29] Ivashenko, V.P., Modeli i algoritmy integracii znanij na osnove odnorodnyh semanticheskih setej(Models and algorithms of knowledge integration based on homogeneous semantic networks), Open Semantic Technologies for Intelligent Systems, 2015, pp 111-132.

[30] Ivashenko, V.P., Semanticheskoe protokolirovanie processov obrabotki znanij(Semantic logging of knowledge processing), Open Semantic Technologies for Intelligent Systems, 2017, pp. 110 – 111.

ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ С БАЗАМИ ЗНАНИЙ

Головко В.А., БГТУ
Голенков В.В., БГУИР
Ивашенко В.П., БГУИР
Таберко В.В., ОАО "Савушкин продукт"
Иванюк Д.С., ОАО "Савушкин продукт"
Крощенко А.А., БГТУ
Ковалёв М.В., БГУИР

Статья посвящена вопросам и направлениям интеграции искусственных нейронных сетей с базами знаний. Выделены два основных направления интеграции:

1) через входы и выходы искусственной нейронной сети, с целью использования такой интеграции баз знаний и искусственных нейронных сетей для решения прикладных задач;

2) через представление искусственных нейронных сетей с помощью онтологических структур и их интерпретацию средствами представления знаний в базе знаний, с целью создания интеллектуальной среды по разработке, обучению и интеграции различных искусственных нейронных сетей, совместимых с базами знаний.

Базы знаний, с которыми интегрируются искусственные нейронные сети, построены на основе однородных семантических сетей, а обработка знаний в них осуществляется с помощью многоагентного подхода.

The implementation of graphodynamic paradigm using the metagraph approach

Yuriy E. Gapanyuk
Computer Science and Control Systems Department
Bauman Moscow State Technical University
Moscow, Russia
gapyu@bmstu.ru

Yuriy T. Kaganov
Computer Science and Control Systems Department
Bauman Moscow State Technical University
Moscow, Russia
kaganov.y.t@bmstu.ru

Georgiy I. Revunkov
Computer Science and Control Systems Department
Bauman Moscow State Technical University
Moscow, Russia
revunkov@bmstu.ru

***Abstract*—This paper proposes an approach for implementation of graphodynamic paradigm using complex graphs. The metagraph model is used as complex graph model. The brief history of metagraph model development is discussed. It is shown that proposed version of the metagraph model provides the implementation of emergence principle using metavertices. Metavertices include vertices, edges and lower-level metavertices. Metaedges are used for process description. The metagraph and hypergraph models comparison is given. It is shown that the hypergraph model does not fully implement the emergence principle. The metagraph and hypernetwork models comparison are given. It is shown that the metagraph model is more flexible then hypernetwork model. Metagraph agents provide dynamical part of graphodynamic paradigm.**

***Keywords*—graphodynamic paradigm, metagraph, metavertex, metaedge, hypergraph, hypernetwork, metagraph agent**

I. Introduction

The graphodynamic paradigm was proposed by Professor Vladimir Golenkov with colleagues in monography [1]. Nowadays the ideas of graphodynamics are widely used in intelligent systems.

The graphodynamic paradigm assumes the following provisions:

- The graph-based model is used as a data model.
- The ways for graph-based data model transformation should be considered.

Currently, complex graph models are used increasingly instead of plain graph models.

The main idea of this paper is the combination of graphodynamic paradigm with complex graph model. We propose to use the metagraph model as a data model. The metagraph agents are used for model transformation.

II. The metagraph model

A. The brief history of metagraph model

At present time there is no single version of metagraph model. There are several "complex graphs with emergence" models that are similar in the basic provisions, but differ in details.

The original version of metagraph model (and term "metagraph") was proposed by A. Basu and R. Blanning in their monography [2].

The terms "metavertex" and "metaedge" were proposed in paper [3]. According to this model, metavertex is a set of vertices (which is isomorphic to the hyperedge of hypergraph). An edge connects two vertices while metaedge connects vertex and metavertex or two metavertices.

The model [4] (presented at OSTIS-2015) used term "metavertex" in the sense of model [3] for fuzzy knowledge-bases representation.

Our paper [5] also used term "metavertex" and "metaedge" but in a different sense compared to [3]. The definition of metavertex is recursive and metavertex may include vertices, edges and other metavertices.

In our model, the metavertex is used for complex data description while metaedge is used for process description. The set of metagraph agents are used for model transformation.

In the following sections we will describe our model in details.

B. The proposed version of metagraph model

The metagraph is described as follows: $MG = \langle V, MV, E, ME \rangle$, where MG – metagraph; V – set of metagraph vertices; MV – set of metagraph metavertices; E – set of metagraph edges; ME – set of metagraph metaedges.

A metagraph vertex is described by the set of attributes: $v_i = \{atr_k\}, v_i \in V$, where v_i – metagraph vertex and atr_k – attribute.

A metagraph edge is described by set of attributes, the source and destination vertices (or metavertices) and edge direction flag: $e_i = \langle v_S, v_E, eo, \{atr_k\} \rangle, e_i \in E, eo = true|false$, where e_i – metagraph edge; v_S – source vertex (metavertex) of the edge; v_E – destination vertex (metavertex) of the edge; eo – edge direction flag ($eo = true$ – directed edge, $eo = false$ – undirected edge); atr_k – attribute.

The metagraph fragment is defined as $MG_i = \{ev_j\}, ev_j \in (V \cup E \cup MV \cup ME)$, where MG_i – metagraph fragment; ev_j – an element that belongs to union of vertices, metavertices, edges and metaedges.

The metagraph metavertex: $mv_i = \langle \{atr_k\}, MG_j \rangle, mv_i \in MV$, where mv_i – metagraph metavertex belongs to set of metagraph metavertices MV; atr_k – attribute, MG_j – metagraph fragment.

The metagraph metaedge: $me_i = \langle v_S, v_E, eo, \{atr_k\}, MG_j \rangle, me_i \in ME, eo = true|false$, where me_i – metagraph metaedge belongs to set of metagraph metaedges ME; v_S – source vertex (metavertex) of the metaedge; v_E – destination vertex (metavertex) of the metaedge; eo – metaedge direction flag ($eo = true$ – directed metaedge, $eo = false$ – undirected metaedge); atr_k – attribute, MG_j – metagraph fragment.

C. The examples of proposed metagraph model

The example of metavertices representation is shown in figure 1.

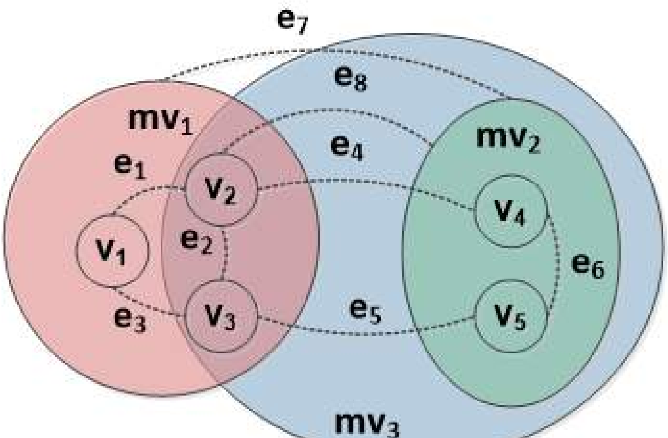


Figure 1. The example of metavertices representation.

This example contains three metavertices: mv_1, mv_2 and mv_3. Metavertex mv_1 contains vertices v_1, v_2, v_3 and connecting them edges e_1, e_2, e_3. Metavertex mv_2 contains vertices v_4, v_5 and connecting them edge e_6. Edges e_4, e_5 are examples of edges connecting vertices $v_2 - v_4$ and $v_3 - v_5$ respectively, and are contained in different metavertices mv_1 and mv_2. Edge e_7 is an example of an edge connecting metavertices mv_1 and mv_2. Edge e_8 is an example of an edge connecting vertex v_2 and metavertex mv_2. Metavertex mv_3 contains metavertex mv_2, vertices v_2, v_3 and edge e_2 from metavertex mv_1 and also edges e_4, e_5, e_8 showing the complex nature of the metagraph structure.

Thus a metavertex in addition to the attributes includes a fragment of the metagraph. The presence of private attributes and connections for a metavertex is distinguishing feature of a metagraph model. It makes the definition of metagraph holonic – a metavertex may include a number of lower level elements and in turn, may be included in a number of higher level elements.

From the general system theory point of view, a metavertex is a special case of the manifestation of the emergence principle, which means that a metavertex with its private attributes and connections become a whole that cannot be separated into its component parts.

The figure 1 helps us to show differences between metagraph model [3] and our model.

In sense of model [3] edges cannot be included in metavertex. In our model metavertex may include both vertices (metavertices) and edges.

Also in sense of model [3] edges e_7 (connecting two metavertices) and e_8 (connecting vertex and metavertex) are metaedges. In our model metaedge is used for process description. The example of metaedge is shown in figure 2.

The directed metaedge contains metavertices $mv_S, \ldots mv_i, \ldots mv_E$ and connecting them edges. The source vertex contains a nested metagraph fragment. During the transition to the destination vertex the nested metagraph fragment became more complex, new vertices, edges, and metavertices are added. Thus, metaedge allows binding the stages of nested metagraph fragment development to the steps of the process described with metaedge.

III. The comparison of metagraph model and other complex graph models

Currently, there are two well-known complex graph models exist: hypergraph model and hypernetwork model. In this section we will compare these models with the metagraph model.

A. The metagraph and hypergraph models comparison

Hypergraph definition according to [6]: $HG = \langle V, HE \rangle, v_i \in V, he_j \in HE$, where HG – hypergraph; V – set of hypergraph vertices; HE – set of non-empty subsets of V called hyperedges; v_i – hypergraph vertex; he_j – hypergraph hyperedge.

A hypergraph may be directed or undirected. A hyperedge in an undirected hypergraph only includes vertices whereas, in a directed hypergraph, a hyperedge defines the order of traversal of vertices. The example of an undirected hypergraph is shown in figure 3.

This example contains thee hyperedges: he_1, he_2, and he_3. Hyperedge he_1 contains vertices v_1, v_2, v_4, v_5. Hyperedge he_2 contains vertices v_2 and v_3. Hyperedge he_3 contains vertices v_4 and v_5. Hyperedges he_1 and he_2 have a common vertex v_2. All vertices of hyperedge he_3 are also vertices of hyperedge he_1.

Comparing metagraph and hypergraph models, it should be noted that the metagraph model is more expressive than the hypergraph model. Comparing figures 1 and 3 it is possible to note some similarities between the metagraph metavertex and the hypergraph hyperedge, but the metagraph offers more details and clarity because the metavertex explicitly defines metavertices, vertices and edges inclusion, whereas the hyperedge does not. The inclusion of hyperedge he_3 in hyperedge he_1 is only graphical and informal, because according to hypergraph definition a hyperedge inclusion operation is not explicitly defined.

Thus the metagraph is a holonic graph model whereas the hypergraph is a near flat graph model that does not fully implement the emergence principle.

B. The metagraph and hypernetwork models comparison

Currently, there are two versions of hypernetwork model exist.

The first version of the hypernetwork model was proposed by Professor Vladimir Popkov with colleagues in 1980s. Professor V. Popkov proposes several kinds of hypernetwork models with complex formalization and therefore only main ideas of hypernetworks will be discussed in this section. According to [7] given the hypergraphs $PS \equiv WS_0, WS_1, WS_2, \dots WS_K$. The hypergraph $PS \equiv WS_0$ is called primary network. The hypergraph WS_i is called secondary network of order i. Also given the sequence of mappings between networks of different orders: $WS_K \xrightarrow{\Phi_K} WS_{K-1} \xrightarrow{\Phi_K - 1} \dots WS_1 \xrightarrow{\Phi_1} PS$. Then the hierarchical abstract hypernetwork of order K may be defined as $AS^K = \langle PS, WS_1, \dots WS_K; \Phi_1, \dots \Phi_K \rangle$. The emergence in this model occurs because of the mappings Φ_i between the layers of hypergraphs.

The second version of the hypernetwork model was proposed by Professor Jeffrey Johnson in his monography [8]. The main idea of Professor J. Johnson variant of hypernetwork model is the idea of hypersimplex (the term is adopted from polyhedral combinatorics). According to [8] a hypersimplex is an ordered set of vertices with an explicit n-ary relation and hypernetwork is a set of hypersimplices. In hierarchical system, the hypersimplex combines k elements at level N (base) with one element at level N+1 (apex). Thus, hypersimplex establishes an emergence between two adjoining levels.

The example of hypernetwork that combines the ideas of two approaches is shown in figure 4.

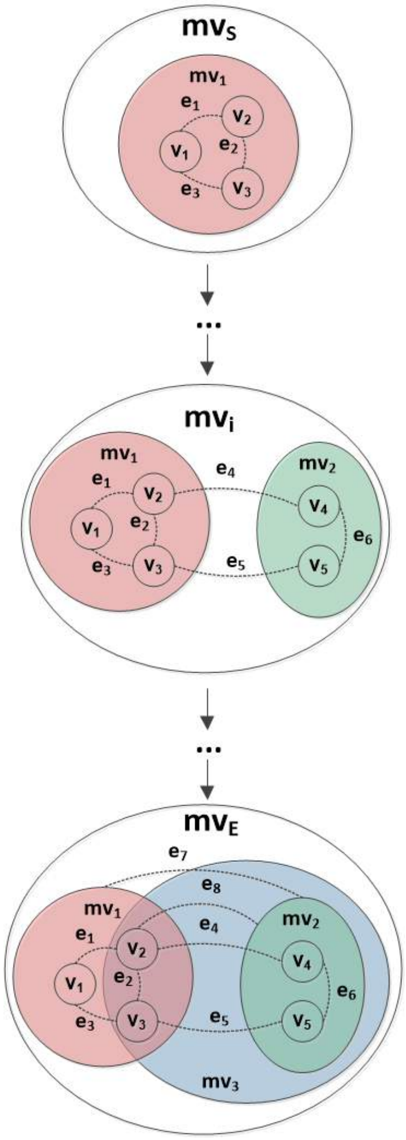


Figure 2. The example of metaedge representation.

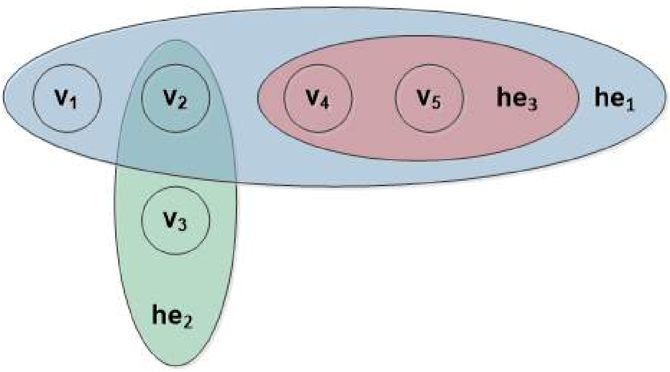


Figure 3. The example of the hypergraph.

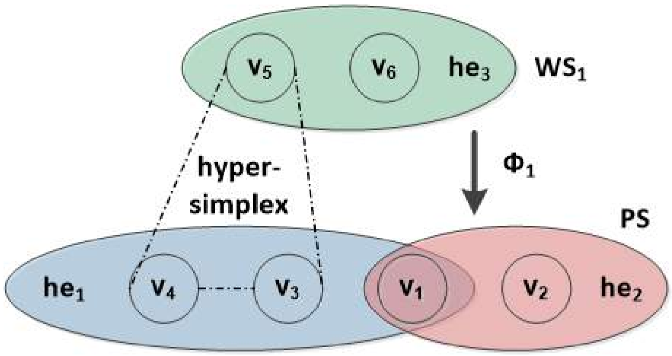


Figure 4. The example of hypernetwork.

The primary network PS is formed by the vertices of hyperedges he_1 and he_2. The first level WS_1 of secondary network is formed by the vertices of hyperedge he_3. Mapping Φ_1 is shown with an arrow. The hypersimplex is emphasized with the dash-dotted line. The hypersimplex is formed by the base (vertices v_3 and v_4 of PS) and apex (vertex v_5 of WS_1).

It should be noted that unlike the relatively simple hypergraph model the hypernetwork model is full model with emergence. Consider the differences between the hypernetwork and metagraph models.

According to the definition of a hypernetwork, it is a layered description of graphs. It is assumed that the hypergraphs may be divided into homogeneous layers and then mapped with mappings or combined with hypersimplices. Metagraph approach is more flexible. It allows combining arbitrary elements that may be layered or not using metavertices.

Comparing the hypernetwork and metagraph models we can make the following notes:

- Hypernetwork model may be considered as "horizontal" or layer-oriented. The emergence appears between adjoining levels using hypersimplices. The metagraph model may be considered as "vertical" or aspect-oriented. The emergence appears at any levels using metavertices.
- In hypernetwork model the elements are organized using hypergraphs inside layers and using mappings or hypersimplices between layers. In metagraph model metavertices are used for organizing elements both inside layers and between layers. Hypersimplex may be considered as a special case of metavertex.
- Metagraph model allows organizing the results of previous organizations. The fragments of flat graph may be organized into metavertices, metavertices may be organized in higher-level metavertices and so on. Metavertex organization is more flexible then hypersimplex organization because hypersimplex assumes base and apex usage and metavertex may include general form graph.
- Metavertex may represent a separate aspect of organization. The same fragment of a flat graph may be included in different metavertices whether these metavertices are used for modeling different aspects.

Thus, we can draw a conclusion that metagraph model is more flexible then hypernetwork model. However, it must be emphasized that from the historical point of view the hypernetwork model was the first complex graph with an emergence model and it helps to understand many aspects of complex graphs with an emergence.

IV. The metagraph model transformation using metagraph agents

The metagraph itself is not more than a complex data structure. To process and transform metagraph data the metagraph agents are used. There are two kinds of metagraph agents: the metagraph function agent (ag^F) and the metagraph rule agent (ag^R). Thus $ag^{MG} = ag^F | ag^R$.

The metagraph function agent serves as a function with input and output parameter in form of metagraph: $ag^F = \langle MG_{IN}, MG_{OUT}, AST \rangle$, where ag^F – metagraph function agent; MG_{IN} – input parameter metagraph; MG_{OUT} – output parameter metagraph; AST – abstract syntax tree of metagraph function agent in form of metagraph.

The metagraph rule agent uses rule-based approach: $ag^R = \langle MG, R, AG^{ST} \rangle, R = \{r_i\}, r_i : MG_j \rightarrow OP^{MG}$, where ag^R – metagraph rule agent; MG – working metagraph, a metagraph on the basis of which the rules of agent are performed; R – set of rules r_i; AG^{ST} – start condition (metagraph fragment for start rule check or start rule); MG_j – a metagraph fragment on the basis of which the rule is performed; OP^{MG} – set of actions performed on metagraph.

The antecedent of a rule is a condition over metagraph fragment, the consequent of rule is a set of actions performed on metagraph. Rules can be divided into open and closed. If the agent contains only open rules it is called open agent. If the agent contains only closed rules it is called closed agent.

The consequent of an open rule is not permitted to change metagraph fragment occurring in rule antecedent. In this case, the input and output metagraph fragments may be separated. The open rule is similar to the template that generates the output metagraph based on the input metagraph.

The consequent of closed rule is permitted to change metagraph fragment occurring in rule antecedent. The metagraph fragment changing in rule consequent cause to trigger the antecedents of other rules bound to the same metagraph fragment. But incorrectly designed closed rules system can cause an infinite loop of metagraph rule agent.

Thus metagraph rule agent can generate the output metagraph based on the input metagraph (using open rules) or can modify the single metagraph (using closed rules).

The distinguishing feature of metagraph agent is its homoiconicity which means that it can be a data structure for itself. This is due to the fact that according to definition metagraph agent may be represented as a set of metagraph fragments and this set can be combined in a single metagraph. Thus higher-level metagraph agent can change the structure of lower-level metagraph agents.

The example of metagraph rule agent is shown in figure 5.

The metagraph rule agent "metagraph rule agent 1" is represented as metagraph metavertex. According to definition, it is bound to the working metagraph MG_1, which is shown with edge e_4.

The metagraph rule agent description contains inner metavertices corresponds to agent rules (rule 1 ... rule N). Each rule metavertex contains antecedent and consequent inner vertices. In given example mv_2 metavertex bound with antecedent which is shown with edge e_2 and mv_3 metavertex bound with consequent which is shown with edge e_3. Antecedent conditions and consequent actions are defined in form of attributes bound to antecedent and consequent corresponding vertices.

The start condition is given in form of attribute "start=true". If the start condition is defined as a start metagraph fragment then the edge bound start metagraph fragment to agent metavertex (edge e_1 in given example) is annotated with

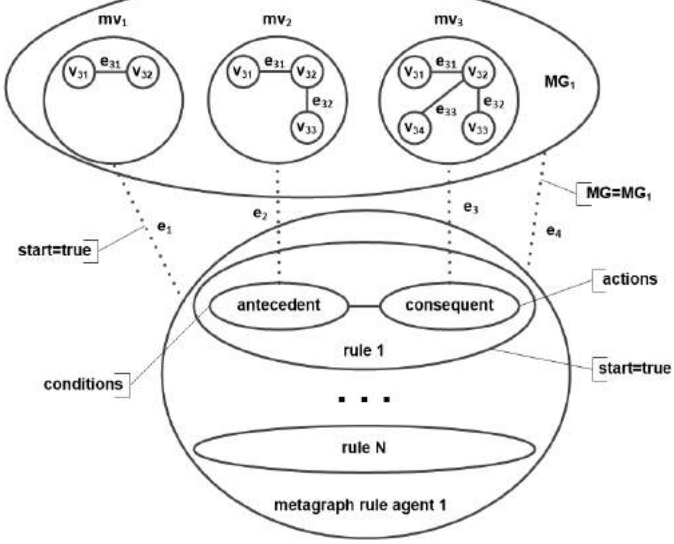


Figure 5. The example of metagraph rule agent.

attribute “start=true”. If the start condition is defined as a start rule then the rule metavertex is annotated with attribute “start=true” (rule 1 in given example), fig. 5 shows both cases corresponding to the start metagraph fragment and to the start rule.

Thus, metagraph agents provide “dynamical” part of graphodynamic paradigm.

V. Modelling the polypeptide chain synthesis for learning software

In this section, we will consider the example of metagraph approach usage for the learning software in the field of molecular biology.

Molecular biology is considered to be one of the most difficult to study topics of biological science. It’s hard to believe that the complexity of functioning of the biological cell invisible to the human eye exceeds the complexity of functioning of a large ERP-system, which can contain thousands of business processes. The difficulty of studying biological processes is also because in studying it is impossible to abstract from the physical and chemical features that accompany these processes. Therefore, the development of learning software that helps to understand even one complex process better is a valid task.

We will review the process of synthesis of a polypeptide chain which is also called “translation” in molecular biology. Translation is an essential part of the protein biosynthesis. This process is very valid from an educational point of view because protein biosynthesis is considered in almost all textbooks of molecular biology. The translation process is very complicated, and in this section, we review it in a simplified way.

The first main actor of the translation process is messenger RNA or mRNA, which may be represented as a chain of codons. The second main actor of the translation process is ribosome consisting of the large subunit and a small subunit. The small subunit is responsible for reading information from mRNA, and large subunit is responsible for generating fragments of the polypeptide chain.

According to [9] the translation process consists of three stages.

The first stage is initiation. At this stage the ribosome assembles around the target mRNA. The small subunit is attached at the start codon.

The second stage is elongation. The small subunit reads information from the current codon. Using this information the large subunit generates fragment of polypeptide chain. Then ribosome moves (translocates) to the next mRNA codon.

The third stage is termination. When the stop codon is reached, the ribosome releases the synthesized polypeptide chain. Under some conditions the ribosome may be disassembled.

From the graphodynamic paradigm point of view the translation process may be considered as a kind of graph automaton that reads codon information and generates polypeptide chain. We will use metagraph approach for translation process modelling. The representation is shown in figure 6.

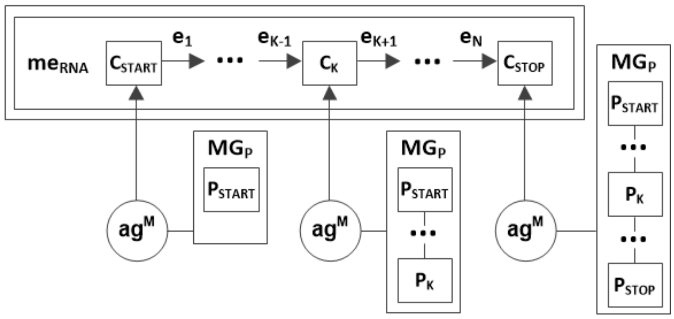


Figure 6. The representation of the translation process based on metagraph approach.

The mRNA is shown in figure 6 as metaedge $me_{RNA} = \langle C_{START}, C_{STOP}, eo = true, \{atr_k\}, MG_{RNA}\rangle$, where C_{START} – source metavertex (start codon); C_{STOP} – destination metavertex (stop codon); $eo = true$ – directed metaedge; atr_k – attribute, MG_{RNA} – metagraph fragment, containing inner codons of mRNA (C_K) linked with edges.

Codon (shown in figure 6 as elementary vertex) may also be represented as metavertex, containing inner vertices and edges according to the required level of detailing.

Ribosome may be represented as metagraph rule agent $ag^{RB} = \langle me_{RNA}, R, C_{START}\rangle, R = \{r_i\}, r_i : C_K \to P_K$, where me_{RNA} – mRNA metaedge representation used as working metagraph; R – set of rules r_i; C_{START} – start codon used as start agent condition; C_K – codon on the basis of which the rule is performed; P_K – the added fragment of polypeptide chain.

The antecedent of rule is approximately corresponds to the small subunit of ribosome modelling. The consequent of rule is approximately corresponds to the large subunit of ribosome modelling.

Agent ag^{RB} is open agent generating output metagraph MG_P based on input metaedge me_{RNA}. The input and output metagraph fragments don't contain common elements.

While processing codons of mRNA agent ag^{RB} sequentially adds fragments of polypeptide chain P_K to the output metagraph MG_P. Vertices P_K are linking using undirected edges.

The process represented in figure 6 is very higher-level. But metagraph approach allows representing linked processes with different levels of abstraction.

For example for each codon or peptide we can link metavertex with its inner representation. And we can modify this representation during translation process using metagraph agents.

Thus, metagraph approach allowed us to represent a model of polypeptide chain synthesis which can be the basis for the learning software. And this is a special case of graphodynamic paradigm.

VI. Conclusion

- The main idea of this paper is the combination of graphodynamic paradigm and complex graph model.
- As complex graph model, we propose to use metagraph model.
- The metagraph model includes vertices, edges, metavertices and metaedges.
- The proposed version of the metagraph model provides the implementation of emergence principle using metavertices. Metavertices include vertices, edges and lower-level metavertices.
- For process description metaedges are used.
- The hypergraph model does not fully implement the emergence principle.
- The hypernetwork model fully implements the emergence principle using hypersimplices. The metagraph model is more flexible then hypernetwork model.
- For metagraph model processing the metagraph function agents and the metagraph rule agents are used. Thus metagraph agents provide "dynamical" part of graphodynamic paradigm.

References

[1] Knowledge Representation and Processing in Graph-Dynamic Associative Computers. V.V. Golenkov, O.E. Eliseeva, V.P. Ivashenko and others. Edited by V.V. Golenkov. Minsk, BSUIR, 2001. 412 p.

[2] A. Basu, R. Blanning. Metagraphs and their applications. New York, Springer, 2007. 173 p.

[3] S.V. Astanin, N.V. Dragnish, N.K. Zhukovskaya. Vlozhennye metagrafy kak modeli slozhnykh ob"ektov [Nested metagraphs as models of complex objects]. Inzhenernyj vestnik Dona, 2012, vol. 23, no. 4-2 (23), p. 76.

[4] L.S. Globa, M.Y. Ternovoy, O.S. Shtogrina. Metagrafy kak osnova dlya predstavleniya i ispol'zovaniya baz nechetkih znaniy [Metagraph based representation and processing of fuzzy knowledgebases] Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2015, pp. 237-240.

[5] E.N. Samokhvalov, G.I. Revunkov, Yu.E. Gapanyuk Ispolzovaniye metagraphov dlya opisaniya semantiki i pragmatiki informatsionnykh sistem [Metagraphs for information systems semantics and pragmatics definition]. Vestnik MGTU im. N.E. Baumana, seriya "Priborostroeniye" [Herald of the Bauman Moscow State Technical University, "Instrument Engineering"], 2015, no. 1, pp. 83-99.

[6] Vitaly I. Voloshin. Introduction to Graph and Hypergraph Theory. New York, Nova Science Publishers, 2009. 303 p.

[7] A.T. Akhmediyarova, J.R. Kuandykova, B.S. Kubekov, I.T. Utepbergenov, V.K. Popkov. Objective of Modeling and Computation of City Electric Transportation Networks Properties. In: International Conference on Information Science and Management Engineering (Icisme 2015), Destech Publications, Phuket, 2015, pp. 106–111.

[8] J. Johnson. Hypernetworks in the Science of Complex Systems. London, Imperial College Press, 2013. 349 p.

[9] I. Samish. Computational Protein Design. New York, Springer Science+Business Media, 2017. 450 p.

РЕАЛИЗАЦИЯ ГРАФОДИНАМИЧЕСКОЙ ПАРАДИГМЫ С ИСПОЛЬЗОВАНИЕМ МЕТАГРАФОВОГО ПОДХОДА

Гапанюк Ю.Е., Каганов Ю.Т., Ревунков Г.И.
Московский государственный технический университет имени Н. Э. Баумана, факультет "Информатика и системы управления".

В статье рассматривается реализация графодинамической парадигмы на основе сложных сетей. В качестве модели сложной сети используется метаграфовая модель. Рассматривается краткая история развития метаграфовой модели. Показано, что предложенная версия метагарфовой модели обеспечивает реализацию принципа эмерджентности с использованием метавершин. Метавершины могут включать вершины, ребра и метвершины нижнего уровня. Метаребра используются для описания процессов. Проведено сравнение моделей метаграфа и гиперграфа. Показано, что гиперграфовая модель не в полной мере реализует принцип эмерджентности. Проведено сравнение моделей метаграфа и гиперсети. Показано, что метаграфовая модель является более гибкой по сравнению с гиперсетевой моделью. Метаграфовые агенты реализуют динамическую часть графодинамической парадигмы.

Knowledge-based ontology concept for numerical data clustering

Peter Grabusts
Rezekne Academy of Technologies
Rezekne, Latvia
peteris.grabusts@rta.lv

***Abstract*—Classical clustering algorithms are sufficiently well studied, they are used for grouping numerical data in similar structures - clusters. Similar objects are placed in the same cluster, different objects in another cluster. All of the classic clustering algorithms have common parameters, and successful selection of which also determines the clustering result. The most important parameters characterizing clustering are: clustering algorithm, metrics, initial number of clusters, criteria for clustering accuracy. In recent years, there has been a tendency towards the possibility of obtaining rules from clusters. Classical clustering algorithms do not apply semantic knowledge. It creates difficulties in interpreting the results of clustering. Presently, the use of ontology opportunities is developing very rapidly, that allows to gain knowledge about a certain data model. The paper analyzes the concept of ontology and prototype development for numerical data clusterization, which includes the most significant indicators characterizing clusterization. The aim of the work is to develop a concept for analyzing clustering data with the help of ontologies. As a result of the work, a study has been conducted on the use of ontologies in this type of tasks.**
***Keywords*—clustering, cluster analysis, ontology**

I. Introduction

Nowadays there is a large amount of data in various fields of science, business, economics and other spheres and there is a need to analyze them for better management of a particular industry. Often, the needs of business stimulate to develop new intelligent methods for data analysis that are oriented towards practical application. The goal of cluster analysis as one of the basic tasks of intellectual data analysis - to search for independent groups (clusters) and their characteristics in analytic data [1], [2], [3]. Solving this problem allows to understand the data in a better way since clustering can be used in any application area where data analysis is required.

Author's research interests have been oriented to clustering analysis: clustering algorithms, fuzzy clustering, rule extraction from clustered data etc [4], [5]. The next step in the research would be the implementation of ontologies in cluster analysis [6].

In order to evaluate the efficiency aspects of clustering, the aim was set - to analyze and summarize the possibilities of clustering algorithms with the purpose of creating an ontology prototype for numerical data clusterization. Research tasks are subjected to the stated aim:

- To review clustering algorithms.
- Carry out the evaluation of the eligibility of the metrics selection.
- Characterize the impact of changes in the number of clusters.
- Evaluate the reliability of the results of clustering (clusters validity).
- Evaluate the possibility to get rules from clusters.
- Develop the ontology concept for numerical data clustering.

According to the previously obtained results of clustering study, the author will make an attempt to create ontology based prototype of clustering concepts using similarity measures, cluster numbers, cluster validity and others characteristic features.

II. An outline of classical clustering approach

The cluster analysis is based on the hypothesis of compactness. It is assumed, that the elements of the training set in the feature space are compact. The main task is to describe these formations formally. Clustering differs from the classification by the fact that there is no need to select a separate changeable group for analysis in the clustering process. From this point of view, clustering is treated as "non-teacher training" and is used at the initial stage of the research.

The cluster analysis is characterized by two features that distinguish it from other methods [2]:

- The result depends on the nature of the objects or their attributes, i.e. they can be uniqualy certain objects or objects with a fuzzy description.
- The result depends on the nature of possible relationship between the cluster and the objects in the clusters, i.e., the possibility of belonging the object to several clusters and the determination of the ownership of the object (strong or fuzzy belonging) should be taken into account.

Taking into account the important role of clustering in the analysis of data, the concept of object belonging was generalized to the function of classes that determines the class objects belonging to a particular class. Two types of classes characterizing functions are distinguished:

- A discrete function that accepts one of the two possible values – belongs/does not belong to the class (classical clusterization).
- A function that accepts values from the interval [0,1]. The closer function values are to 1, the «more» the object belongs to a particular class (fuzzy clustering).

Clustering algorithms are mainly intended for the processing of multidimensional data samples, when the data is given in the table form «object-attribute». They allow to group objects into certain groups, in which objects are related to each other according to a particular rule. It does not matter how the following groups are called - taxons, clusters or classes, the main thing is that they accurately reflect the properties of these objects. After the clustering, data for further analysis are used by other intellectual data analysis methods, in order to find out the nature of the acquired regularities and the possibilities of future use.

Clustering is commonly used in the data processing process as the first step of analysis. It identifies similar data groups that can later be used to explore data interrelation. The cluster analysis process formally consists of the following steps (see Fig. 1):

- Selection of necessary data for analysis.
- Determination of characterizing class sizes and boundaries for class data (clusters).
- Data grouping in clusters.
- Determining the class hierarchy and analyzing the results.

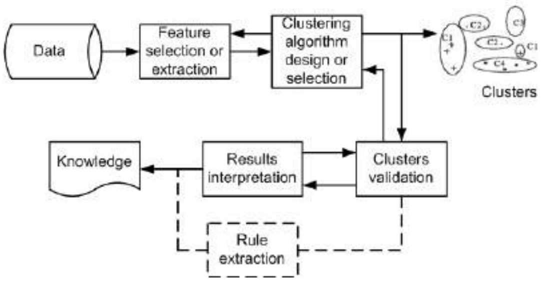


Figure 1. Clustering procedure

All clustering algorithms have common characteristics, the selection of which is characterized by a clustering efficiency. The most important clustering parameters are as follows: metric (cluster element distance to the cluster center), the number of clusters k, clustering validity assessment, opportunity to get rules [7], [8], [9].

III. Clustering characteristics

Classes of clustering algorithms are shown in the Fig. 2:

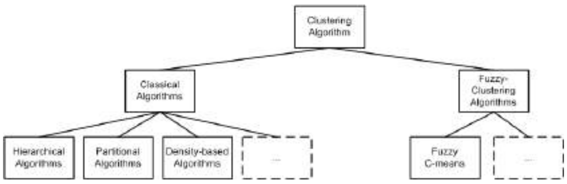


Figure 2. Hierarchical view of the clustering algorithm class

Metrics. The main purpose of metrics learning in a specific problem is to learn an appropriate distance/similarity function. A metrics or distance function is a function which defines a distance between elements of a set [10], [11]. A set with a metric is called a metric space. In many data retrieval and data mining applications, such as clustering, measuring similarity between objects has become an important part. In general, the task is to define a function Sim(X,Y), where X and Y are two objects or sets of a certain class, and the value of the function represents the degree of "similarity" between the two.

Euclidean distance is the most common use of distance – it computes the root of square differences between coordinates of a pair of objects.

Manhattan distance or city block distance represents distance between points in a city road grid. It computes the absolute differences between coordinates of a pair of objects.

Minkowski distance is the generalized metric distance.

Cosine distance is the angular difference between two vectors.

The distance measure can also be derived from the correlation coefficient, such as the Pearson correlation coefficient. Correlation coefficient is standardized angular separation by centering the coordinates to its mean value. It measures similarity rather than distance or dissimilarity.

The summary of the metrics is shown in the Fig. 3.

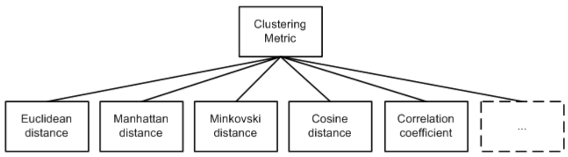


Figure 3. Hierarchical view of the clustering metrics class

Traditionally Euclidean distance is used in clustering algorithms, the choice of other metric in definite cases may be disputable. It depends on the task, the amount of data and on the complexity of the task.

Cluster numbers. An essential issue in implementing of clustering algorithms is the determination of the number of clusters and initial centers. In the simplest tasks it is assumed that the number of clusters is known apriori and it is suggested to take the first m points of the training set as the initial values of the m cluster centers.

Clustering validity. Cluster validity is a method to find a set of clusters that best fits natural partitions (number of clusters) without any class information. There are three fundamental criteria to investigate the cluster validity: external criteria, internal criteria, and relative criteria. In this case only external cluster validity index was analyzed [2].

Given a data set X and a clustering structure C derived from the application of a certain clustering algorithm on X, external criteria compare the obtained clustering structure C to a prespecified structure, which reflects a priori information on the clustering structure of X. For example, an external criterion can be used to examine the match between the cluster labels with the category labels based on apriori information.

The most popular clustering quality external criteria are shown in the Fig. 4.

For example, Rand index suggests an objective criterion for comparing two arbitrary clusterings based on how pairs of data

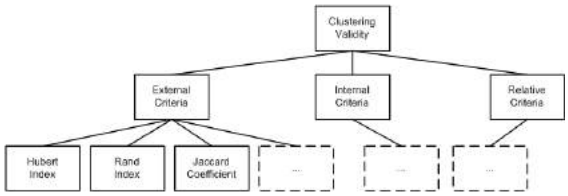


Figure 4. Hierarchical view of the clustering validity class

points are clustered. Given two clusterings, for any two data points there are two cases:

- The first case is that the two points are placed together in a cluster in each of two clusterings or they are assigned to different clusters in both clusterings.
- The second case is that the two points are placed together in a cluster in one clustering and they are assigned to different clusters in the other.

Rule extract. The possibility of direct transforming clustering information into a symbolic knowledge form is through rule extraction. Such assumptions are given as IF ... THEN ... rules [12]. The benefits of rule-making are as follows:

- The opportunity to check the acquired rules for different input data options is provided.
- Deficiencies in the training data can be identified, thus clustering can be improved by introducing or removing additional clusters.
- Determination of previously unknown regularities in the Data Mining industry.
- The rule base can be created from the acquired rules, which could be used in future for similar types of applications.

Several algorithms of artificial neural networks in the training process use clustering, which results in the creation of hidden units, which are in fact cluster centers [9], [13]. The nature of each hidden unit enables a simple translation into a single rule:

IF Feature1 is TRUE AND IF Feature2 is TRUE ... AND IF FeatureN is TRUE THEN ClassX.

Thus, the knowledge base is collected for the most important indicators characterizing clustering.

IV. Ontology based approach

In this paper author presents a formal clustering ontology framework concept, which can provide the background for numerical data clustering. Using the ontology, numerical clustering can become a knowledge-driven process.

As it was mentioned in the previous chapters clustering is using at the data level instead of the knowledge level, which helps from precisely identifying targets and understanding the clustering results. Existing clustering methods consider various constraints and they only consider limited knowledge concerning the domain and the users. In such a way, to include domain knowledge in the clustering methods and clustering process becomes an important topic in clustering data research and analysis.

There are many and different definitions of ontologies, but the following definition has recently been accepted as generally recognized: „An ontology is a formal explicit specification of a shared conceptualization" [14]. Ontologies are often equated with taxonomic hierarchies of classes. It can be said that the purpose of ontology is to accumulate knowledge in a general and formal way.

Ontologies can be classified in different forms. One of the most popular types of classification is offered by Guarino, who classified types of ontologies according to their level of dependence on a particular task or point of view [15].

It should be noted that ontologies are widely used in document clustering and Semantic Web, but undeservedly forgotten in numerical data clustering. Thus, an ontology is an explicit representation of knowledge. It is a formal, explicit specification of shared conceptualizations, representing the concepts and their relations that are relevant for a given domain of discourse [14].

The concept of the numerical data clustering ontology to be developed consists of the following classes:

Clustering-Task. It is an abstract class. This is connected to the clustering algorithm class. Depending on the purpose of the clustering and the domain, the clustering algorithm, the number of clusters and sample data are selected.

Clustering-Algorithm. This class represents a list of available clustering algorithms and their features (see Fig.2).

Clustering-Metric. This class represents a list of available distance metrics for clustering algorithms (see Fig.3).

Clustering-Numbers. This class represents a list of available numbers of clusters.

Clustering-Validity. This class represents a list of cluster validity methods (see Fig.4).

Clustering-Rule. This class represents a list of rule extraction methods from clusters (if it is possible).

Based on such class analysis the following approach is offered for ontology-based clustering, as shown in the Fig. 5.

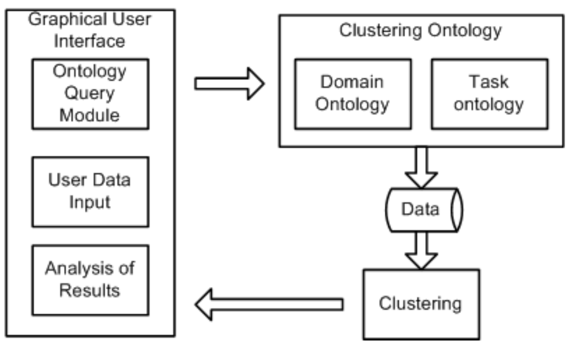


Figure 5. The framework concept of ontology-based numerical data clustering

Clustering ontology prototype should work according to the following scheme: numerical data selection, choice of clustering algorithm, determining the number of clusters, performance of clustering, validation of clustering, acquisition of rules (if possible) (see. Fig. 6.).

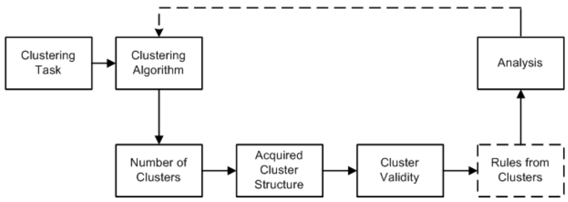


Figure 6. Working scheme of clustering prototype

Developing framework Protege OWL tool is used for construct this concept [16].

Protege is an ontology and knowledge base editor. Protege is a tool that enables the construction of domain ontologies, customized data entry forms to enter data. Protege allows the definition of classes, class hierarchies, variables and the relationships between classes and the properties of these relationships.

Protege is a special tool, which is thought to create and edit ontology, but OWL (Web Ontology Language) is a language through which it is possible to define the ontology. OWL ontology may include descriptions of classes, their characteristics and their instances. OWL formal semantics describes how, using these data get information which was not openly described in ontology, but which follows from the data semantics. Protege is a free open-source platform, which contains special tool kit which makes it possible to construct domain models and knowledge-based applications based on ontologies. In Protege environment a number of knowledge-modeling structures and actions that support ontology creation, visualization and editing of different display formats are implemented.

The development of ontologies with Protege begins with the definition and description of the classes hierarchy, after that the instances are assigned of these classes and different type of relationships (properties in Protege) in order to put more meaningful information within the ontology [17].

To demonstrate the development of ontology, four classes are taken: Clustering-Algorithm, Clustering-Metric, Clustering-Validity and Clustering-Numbers (see Fig. 7, Fig. 8 and Fig. 10).

Since the clustering algorithm refers to the partition algorithm class, K-Means member was included in the partition algorithms class. The K-means algorithm can use the metric Euclidean-distance or Manhattan-distance; then in Clustering-Metric class the members Euclidean-distance or Manhattan-distance are included.

The following properties were defined for K-Means:

K-Means – use -> Euclidean-distance or

K-Means – use -> Manhattan-distance

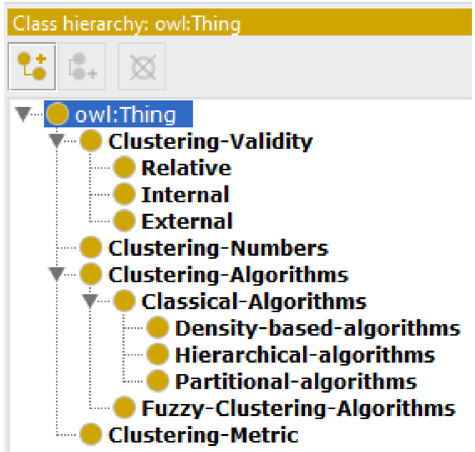


Figure 7. Clustering domain subclasses in the “Class hierachy” tab of Protege

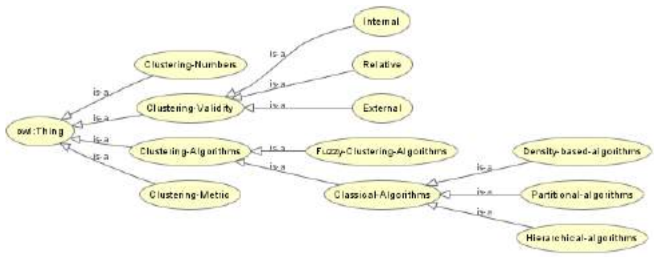


Figure 8. Clustering domain subclasses in the “OWLViz” tab of Protege

To simplify, we assume that three clusters are being studied in clustering and validation testing is carried out using the Randa index. Accordingly, we obtain:

K-Means – use -> 3C

Rand – use -> 3C

In turn, the Clustering-Metric object Euclidean-distance is assigned a property

Euclidean-distance – isUsedBy -> K-means

3C – isUsedBy -> K-means

3C – isUsedBy -> Rand

The result is shown in the Fig. 11.

Recently, the Protege developer has offered a new product-WebProtege, where the users can process their OWL data. WebProtege is an ontology development environment for the Web that makes it easy to create, upload, modify, and share ontologies for collaborative viewing and editing. WebProtege is lightweight ontology editor and knowledge acquisition tool for the Web [16].

Currently, there is no possibility to visualize the ontology as it was done with OntoGraf. Now it is possible to work with Classes, Properties, Individuals (see Figure 9).

An example of a demonstration shows that with a help of Protege an effective ontology description can be created,

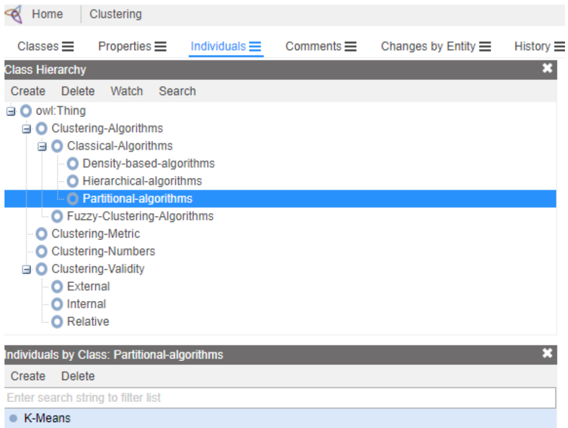


Figure 9. Clustering class hierarchy in WebProtege

but it is a sufficiently laborious process. The author plans to continue the work on the further development of numerical data clustering ontology.

V. Conclusion

There are no directly formalized criteria in the cluster analysis, therefore different clustering parameters are chosen in a subjective evaluation. This refers to the choice of the clustering algorithm, the choice of the number of clusters in each particular case, the determination of cluster validation criteria. Equally important is the acquisition of knowledge from clusters in the form of rules. All this causes a problem in the interpretation of clustering results. In recent decades, cluster analysis has transformed from a single data analysis section in a separate direction that is closely linked to knowledge support systems. Partly it happened due to the introduction of concepts of ontology in the description of clustering characteristics. The use of clustering ontologies in document and semantic web applications is developing very rapidly, but the clustering of numerical data is undeservedly neglected. The author is striving to formulate and create an ontology-based prototype for numerical data clustering. This concept contains several concept classes: clustering algorithms, cluster numbers, cluster validity, and other characteristics features. Future studies will focus on specification of these classes and developing a real model appropriately to the data clustering purposes.

References

[1] B. S. Everitt, *Cluster analysis*, John Viley and Sons, London, 1993.

[2] R.Xu and D.C.Wunch, *Clustering*, John Wiley and Sons, 2010, pp. 263-278.

[3] X. Rui and D. Wunsch II, Survey of clustering algorithms, *Neural Networks*, IEEE Transactions, 16(3):645–678, May 2005.

[4] F. Hoppner, F. Klawonn, R. Kruse and T. Runkler, *Fuzzy Cluster Analysis*, John Whiley and Sons, New York, 1999,289 p.

[5] M. Crawen and J. Shavlik, Using sampling and queries to extract rules from trained neural networks, *Machine Learning: Proceedings of the Eleventh International Conference*, San Francisco, CA, 1994.

[6] D. Gašević, D. Djurić and V. Devedžić, *Model driven architecture and ontology development*, Springer-Verlag, 2006.

[7] G. Gan, C. Ma and J. Wu, *Data clustering: Theory, algorithms and applications*, ASA-SIAM series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA, 2007.

[8] L. Kaufman and P. J. Rousseeuw, *Finding groups in data. An introduction to cluster analysis*, John Wiley and Sons, 2005.

[9] R. Andrews and S. Gewa, *RULEX and CEBP networks as the basis for a rule refinement system*, in J. Hallam et al, editor, Hybrid Problems, Hybrid Solutions. IOS Press, 1995.

[10] P. Vitanyi, *Universal similarity*, ITW2005, Rotorua, New Zealand, 2005.

[11] M. Li, X. Chen, B. Ma and P. Vitanyi, The similarity metric, *IEEE Transactions on Information Theory*, vol.50, No. 12, pp.3250-3264, 2004.

[12] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2010, 1132 p.

[13] D. R. Hush and B. G. Horne, Progress in Supervised Neural Networks. What's new since Lippmann?, *IEEE Signal Processing Magazine*, vol.10, No 1., p.8-39, 1993.

[14] T. R. Gruber, A translation approach to portable ontologies, *Knowledge Acquisition*, 5(2), 199-220, 1993.

[15] N. Guarino, Formal Ontology in Information Systems. *1st International Conference on Formal Ontology in Information Systems*, FOIS, Trento, Italy, IOS Press, 3-15, 1998.

[16] Protege project homepage. Available at: http://protege.stanford.edu/ (Accessed 2017 Nov).

[17] P. Grabusts, The Concept of Ontology for Numerical Data Clustering, *Proceedings of the 9th International Conference „Environment. Technology. Resources"*, Rezekne, Latvia, June,20-22, 2013, Vol.2., P.11-16.

КОНЦЕПЦИЯ ОНТОЛОГИИ ОСНОВАННАЯ НА ЗНАНИЯХ ДЛЯ КЛАСТЕРИЗАЦИИ ЧИСЛОВЫХ ДАННЫХ

Грабуст П.С.

Резекненская Академия Технологий, Латвия

Классические алгоритмы кластеризации достаточно хорошо изучены, они используются для группировки числовых данных в аналогичных структурах - кластерах. Аналогичные объекты помещаются в один кластер, разные объекты в другой кластер. Все классические алгоритмы кластеризации имеют общие параметры, и их успешный выбор также определяет результат кластеризации. Наиболее важными параметрами, характеризующими кластеризацию, являются: алгоритм кластеризации, метрика, начальное количество кластеров, критерии точности кластеризации. В последние годы наблюдается тенденция к возможности получения законов из кластеров. В классических алгоритмах кластеризации семантические знания не используются. Это создает трудности при интерпретации результатов кластеризации. В настоящее время использование возможностей онтологии очень быстро развивается, что позволяет получить знания об определенной модели данных. В данной работе проанализирована концепция онтологии и разработан прототип для кластеризации числовых данных, что включает в себя наиболее важные показатели, характеризующие кластеризацию. Целью работы является разработка концепции анализа данных кластеризации с помощью онтологий. В результате работы было проведено исследование о возможностях использования онтологий в задачах такого типа.

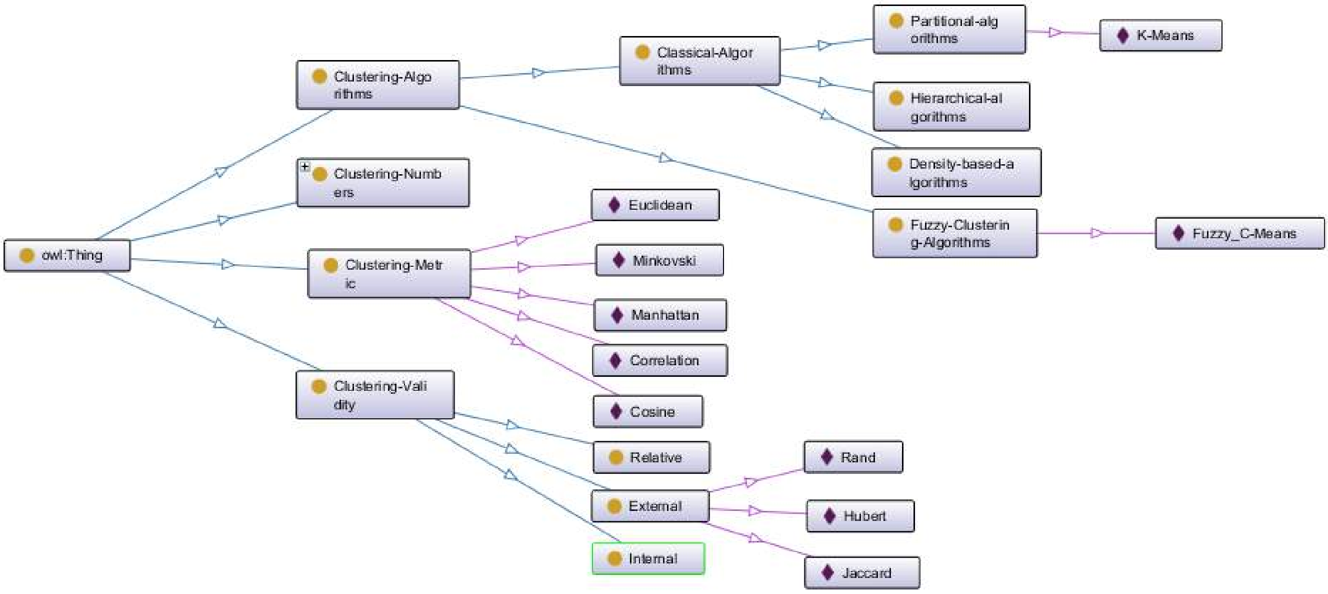


Figure 10. A visualization of the Clustering domain subclasses in OntoGraf tab

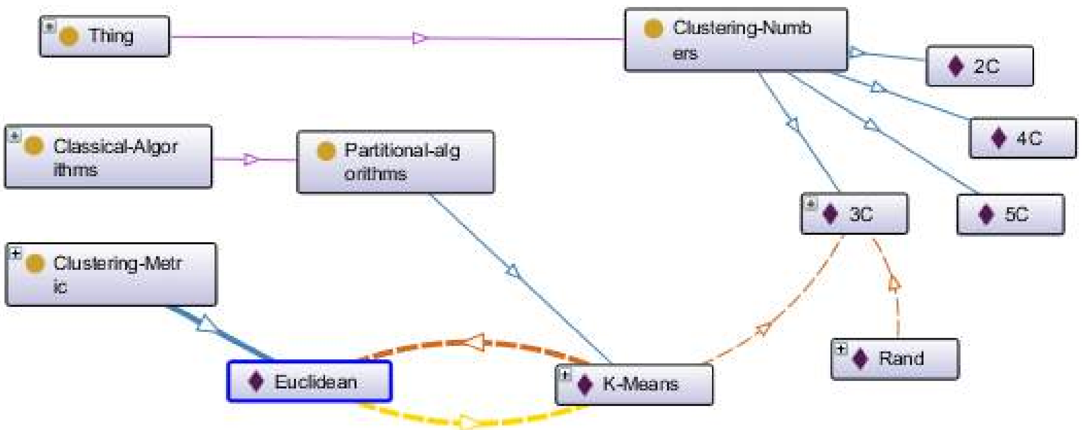


Figure 11. Visualization of K-means properties in the Clustering-Metric class

Processing of Wiki Resource Semantics on Base of Ontological Analysis

Julia Rogushina
Institute of Software Systems of National Academy of Sciences of Ukraine
Kyiv, Ukraine
ladamandraka2010@gmail.com

***Abstract*—The work objects the development of ontology-based means for processing of semantically matched Wiki resources. Such means allows more relevant representation of domain-specific information objects corresponded with Wiki pages. Proposed approach to semantic processing of Wiki resources is based on methods of ontological analyses of complex information objects. Ontological analyses provides matching of Wiki resource with ontology that allows to use various powerful mechanisms of knowledge management. Main results: methodological recommendations for development of hierarchy of categories and semantic properties of Wiki resource; their use for efficient search and navigation; formal representation of semantic elements of Wiki and their matching with relevant domain ontology. The results of proposed research work are used for design of the electronic version of the Big Ukrainian Encyclopedia (e-BUE) on technological platform of the Semantic MediaWiki. Processing of the semantic properties and categories of Wiki provides the e-BUE transformation into the distributed knowledge base with extended functionality**

***Keywords*—semantic Wiki, ontology, semantic properties**

I. Introduction

Now the use of the Web distributed knowledge determs the effectiveness of information systems. Therefore, the semantization of the Web information resources (IRs) that provides the possibility of their automated processing at the knowledge level is a topical issue of modern information technology. This is caused by great amount of such information as well as the complexity of its structure.

The creation of modern encyclopedic information sources, which includes the electronic version of the Big Ukrainian Encyclopedia (e-BUE), should be focused on intelligent technologies of information processing: the possibility of automated acquisition of relevant knowledge from the IR by other Web applications will make this resource really useful.

II. Formulation of the problem

The transformation of the Wiki resource into the distributed knowledge base requires the definition of the basic principles for page classification and the development of the system of categories and semantic properties that allow formalizing and integrating the knowledge contained on individual Wiki pages. Due to the fact that built-in semantic tools of Semantic MediaWiki do not support the analysis and visualization of this knowledge system, it is advisable to use external knowledge management tools based on ontological analysis. Therefore we propose to develop the rules for comparing the semantic markup of Wiki pages with elements of ontologies and the methods of their use.

III. Semantic MediaWiki

The Wiki technologies [1] provide distributed information processing on the Web and allows users to freely edit the content of the pages. Today, there is a large number of semantic extensions of this technology. One of them is Semantic MediaWiki (SMW) which turns MediaWiki into semantic resource by allowing automatic integration of information from various Wiki pages and supporting of complex semantic queries that process semantic properties and categories of information objects [2].

SMW has a fairly high expressive power, reliable implementation and user-friendly interface, and IRs on it‘s base are dynamically updated and rapidly expanding by the entire community of users, have a well-defined and easy to understand structure and provide information processing at the semantic level. Therefore SMW can be used as a technological platform for group knowledge management. It should be noted that now Semantic MediaWiki is a base for increasing number of sites, portals and encyclopedias. Such interest to semantization of Wiki technology causes the development of tools for intelligent processing of Wiki IRs.

Built-in tools of SMW allows to build ontologies represented in the OWL language but these tools are not flexible enough to generate the ontology relevant to needs of external application. Therefore we propose to develop advanced methods for creating and replenishing of ontologies based on the Wiki-resources, as well as comparing of ontologies created and improved on base of different variants of the Wiki-resources.

A. Categories in SMW

Categorization of Wiki pages is a convenient tool for classification of knowledge represented by Wiki resource. Wiki allows to use an arbitrary number of classifiers for every page. The main purposes of classifiers for Wiki resource: 1. to combine a group of pages with common properties or semantics;2. to retrieve the pages by classifiers.

Each page of the Wiki resource can be assigned to one or more categories of any classifier (depending on the classifier specifics). There is very important for realization of various encyclopedic IR on base of Wiki because the encyclopedia implies the co-operation and coexistence of different points of view without competitive extrusion where this is not required.

So the system of categories in any area is not obliged to be an attribute of only one classification tree, but may be a system of tree-shaped structures that are superimposed one on another.

That is, the same category or Wiki page can be classified into categories from different classifications, and categorizing of the page within some classification system is not important for it‘s categorized within another classification systems. For example, a Wiki page that describes organization can be classified by the type of institution (category "the National Academy of Sciences"), by its location (category "organization of Kiev"), by the number of employees (category "big organization").

Wiki-based encyclopedias can use either standard classifiers, such as UDC, or systems of categories proposed by IR developers. For example, the e-BUE uses the classification of pages by scientific directions and the set of categories-tags (common and specific for sciences).

B. Semantic properties in SMW

Semantic properties provide data binding to Wiki pages. Properties, in contrast to the Wiki categories, have not only a name but also a value. Each property has a type, name, and value; in addition, it has separate Wiki page in a special namespace that allows to specify a property type, define it’s position in the property hierarchy, and also documents its usage.

Semantic properties in SMW have values of various types – either common datatypes (Boolean, Number, Text etc.) or specific for Wiki (Page, Geographic coordinate, URL etc.). Properties and types are the basic way of entering semantic data.

IV. Types of information objects as a search component in Wiki resources

Information object (IO) is an information model of an object of a certain subject domain that defines the structure, attributes, limitation of integrity and, possibly, the behavior of this object [3]. In case of Wiki IRs IO is the content of some Wiki page together with it’s semantic matching. For example, every entry of e-BUE is a description of some IO.

Categories and semantic properties can be used as an instrument of representation of information objects: categories of encyclopaedias and guides based on Wiki group pages that have similar structure and the same semantic properties. Encyclopaedia developers can create a prototype pages with a unified content placement which simplify the information perception by users. The group of Wiki pages based on the same prototype page is named the type of IO.

For e-BUE, the type of IO is a group of e-BUE entries that has common semantic properties.

The presence of various IO types greatly improves the quality of the search and navigation into e-BUE, but developers of this IO system have to do it systematically. They have to define correctly (according to domain specifics) name spaces, hierarchy and other relations between categories, range of values and range of definition of semantic properties etc. It is advisable to use two approaches for creation of IO types – top-down and bottom-up.

Top-down approach – from top-level categories of IO to their subcategories that more clearly characterize the content. For example, at first the category “Person” is defined with the set of it‘s semantic properties, subsections and categories. On the next step the category “Scientist” is defined as a subcategory of “Person”. This subcategory inherits all properties, subsections and categories of the supercategory but can obtain additional –properties (such as “academic degree” or “publications”).

The bottom-up approach – from Wiki pages to IO types. In process of creating of new articles or reworking of existing ones expert tries to choose some similar pages and then defines the type of IO for them. For example, under the e-BUE article "Austro-Prussian War of 1866" expert can create IO type "War", and under the article "Abstract" – IO type "Document". Then these types can also be used for classification of other articles.

V. Search and navigation in e-BUE

One of the most important factors in the use of modern distributed knowledge bases is the organization of search and navigation, which allows the user to receive the relevant information [4]. In this case, both the access time and the clarity and convenience of the user interface are important.

Search tools can be subdivided into the following main groups: 1. by keywords (for e-BUE – by the article’s title or by the initial letters of the title); 2. by domain topics (for e-BUE – by knowledge directions); 3. by type of information object (for e-BUE – the search for personalities and concepts and by other IO types); 4. by IO semantics (for e-BUE – by semantic queries with categories and values of the semantic properties of Wiki-pages). The first one provides the fastest access to information, the second is based on commonly used classifiers (they allow to formalize the classification, but complicate the search for ordinary users), and the third one takes into account the semantics of the common user’s information needs and allows to obtain admittance to simular articles.

The fourth type of search is a semantic search. It results the information object with a complex structure that is used for search procedures. It can be considered as a special case of object recognition problem. It‘s implementationin e-BUE requires the installation of semantic plug-ins designed to handle semantic properties of articles.

VI. Components of semantic Wiki-resource

The formal model of any Wiki-resource consists of the following elements:

- the set of Wiki-pages P that includes:
  - the set of pages created by users;
  - the set of pages describing categories;
  - the set of pages describing templates;
  - the set of other special pages;

- the one-element set L that describes the links from one Wiki page of this resource to another Wiki page of this resource (although Wiki-resources provide links to other types of pages, they are not included in this model).

The formal model of semantically enriched Wiki-resources is more complex and includes a number of elements related to semantic properties: the set P is supplemented by the set of the Wiki pages that describe semantic properties, and the set L is supplemented by the set of links deal with semantic properties with type “Page”.

VII. Ontologies as a means of distributed knowledge representation

Ontologies are widely used to represent knowledge in the web-oriented information systems because they provide reuse of knowledge in different applications. It causes the great interest to methods and tools of ontological analysis [5].
One of the important areas of such research is the integration of ontology with other information resources of the Web, namely, with semantic Wiki resources. Such properties of the Web as heterogeneity and dynamism create a number of problems associated with the replenishment, use and evaluation of ontologies, and cause the need for more dynamic semantic Wiki resources that can be used to update these ontologies.

Processing of large amounts of knowledge (for example, into e-BUE with tens of thousands of pages) requires to automate this work on base of the approaches, methods and tools currently used for knowledge management. Creating of consistent, complete and understandable system of the IO types is extremely difficult even for domain experts that haven’t the skills and experience of knowledge engineer. Development of system with a large number and complex structure of IO types is implemented much more efficient with the help of methods and tools of ontological analysis: main elements of semantic markup are considered as classes, instances, and attitudes of the ontology of the corresponding domain.

In order to use such ontological tools, it is necessary to develop a method for transforming ontology elements into Semantic MediaWiki (categories and semantic properties). The next stage deals with refining the initial domain ontology by analysis of semantically-tagged Wiki-resources that requires a method for transforming Semantic MediaWiki constructs into ontological representation in OWL. Tools of ontological analysis enable to evaluate properties of this ontology and it’s pertinentness to user’s conceptions about domain. Reiteration of these actions should ensure the formation of an adequate domain ontology that can be used in different applications as a knowledge base.

VIII. Formal model of ontology

In the general case, the formal model of domain ontology is an ordered triple O = <T, R, F>, where T - the set of domain concepts; R – the set of relations between them; F – the set of interpretation functions for of concepts and relations.

This formal model can be specified in different ways depending on the purpose and scope of the ontology purpose and domain [6].

Now the Web-oriented information systems use most often various dialects of the OWL (Web Ontology Language) language for ontology representation [7]. OWL is one of components of Semantic Web [8]. Ontology in OWL is a sequence of axioms and facts, as well as references to other ontologies. OWL enables to obtain by logical inference of facts that are not represented directly in ontology but specified by its formal semantics. There are a lot of approaches to ontology formalization [9]. For processing semantics of the Wiki resources we should use formal model of ontology that consists of the following elements:

- X – the set of ontology concepts that joinsthe set of classes and the set of class instances;

- R - the set of relations between elements of ontology that joins: - hierarchical relation between ontology classes (with such properties as antisymmetry and transitivity);- the set of object properties that define relationships between class instances; the set of data properties that define relationships between class instances and values of these properties;

- F – the set of properties of ontology classes, class instances and their properties that can be used for logical output (for example, equivalence, difference);

- T – the set of data types (for example, string, integer).

The choice of such ontology model causes by the following reasons. Firstly, it quite enough corresponds to the intuitive representation of the ontology laid in the user interface of the broadly used tool for ontological analysis Protege [10]. Secondly, this model integrates quite easily with various ontology-based applications. Thirdly, this model allows to compare the ontological representation of domain knowledge with semantic constructions of Semantic MediaWiki.

Now ontologies are widely used for semantic markup of various information resources [11]. The problem of matching ontologies and semantic Wiki resources arises in several cases. First, development of semantic Wiki resources requires to create the set of categories and semantic properties. But Semantic MediaWiki’s built-in tools do not allow either to visualize this information or to assess its integrity and consistency. Therefore, there is advisable to construct the ontology for domain of Wiki resource and then use this ontology as the basis for semantic markup. Secondly, Wiki resources are more dynamic in comparison with ontologies because many users can participate in their development and improvement, and therefore they can be useful for improving the corresponding domain ontologies by actual information [12].

IX. Matching of ontologies and semantic Wikis

Some correspondences between the elements of domain ontology and the Semantic MediaWiki pages are one-to-one and can be detected automatically, some of which require additional refinements by the user but can also be partially automated. Matching is one-to-one to such elements:
- Link to another Wiki page with The "Page" object property;

- Semantic property of type "page" with Object property;
- Semantic property of any other type with Value of data. Some Wiki elements can be tansformed into ontological elements:
- Category into Class;
- Category Hierarchy into Class Hierarchy;
- Wiki page into Individual of class;
- Template into Class. But domain expert has to define what element of Wiki resource would be built for class of ontology – Category or Template: either Template or Category would be developed only for frequently used classes with unified structure, and Category only – for all other ones. For example, e-BUE contains templates and categories for Personality, Author BUE and Organization, and categories for Instrument, Group of persons and Region.

If OWL ontology is already built then it's easy to use it in Semantic MediaWiki. But the reverse process can not be completely automated. Moreover, automatically generating ontology for Semantic MediaWiki will lose the information from OWL ontology that deals with characteristics of classes and properties that have no analogues in the Wiki (in particular, about the equivalence of classes and properties, their noncrossing, their region of value, and definition). At the same time, some part of the Semantic MediaWiki content can not be directly transformed into ontology. For example, the fact that the pages use the same template suggests that these pages describe IOs of the same type. But representation of this fact in ontology requires to create a specific class and link it to the page element. In addition, it is not possible to associate with the ontology class the specific fragment of Wiki page.

X. Conclusions

The proposed approach to the use of semantic properties of Wiki pages to expand the functionality of the information resource can enhance the search and navigation efficiency in such resources by more complete and relevant satisfaction of information needs of users. Matching the elements of Wiki resource with elements of ontology allows, on the one hand, to take advantage of the ontological presentation for knowledge management of information from Wiki pages (for example, to check their consistency, perform logical output or perform complex queries) of various encyclopedias and Wikis. On the other hand, it allows to transform the encyclopedia into a distributed knowledge base that can be used not only by people but also by other intelligent applications. The availability of semantic markup on the basis of the domain ontology greatly extends both the scope of the use of such information resource and the effectiveness of working with them. Use of modern standards of knowledge representation ensures the interoperability of Wiki content.

References

[1] Wagner C. Wiki: A technology for conversational knowledge management and group collaboration *The Communications of the Association for Information Systems*, 2004, V. 13(1), pp. 264-289.

[2] Krotzsch M., Vrandecic D., Volkel M. Semantic MediaWiki. Available at: http://c.unik.no/images/6/6d/SemanticMW.pdf (accessed 2017, Nov).

[3] Rogushina J.V. The Use of Ontological Knowledge for Semantic Search of Complex Information Objects *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, 2017, pp.127-132.

[4] Rogushina J. Use of the Ontological Model for Personification of the Semantic Search *International Journal of Mathematical Sciences and Computing(IJMSC)*, V.2, N.1, 2016, pp.1-15.

[5] Gruber T.R. The role of common ontology in achieving sharable, reusable knowledge bases. Available at: http://www.cin.ufpe.br/ mtcfa/files/ 10.1.1.35.1743.pdf (accessed 2017, Oct).

[6] Guarino N. Formal Ontology in Information Systems. Formal Ontology in Information Systems. *Proceedings of FOIS'98*, 1998, pp.3-15.

[7] OWL 2 Web Ontology Language Document Overview. W3C. 2009 Available at: http://www.w3.org/TR/owl2-overview/ (accessed 2017, Oct).

[8] Davies J., Fensel D., van Harmelen F. *Towards the Semantic Web: Ontology-driven knowledge management*, John Wiley & Sons Ltd, England, 2002.

[9] Uschold M., Grüninger M. Ontologies: Principles, Methods and Applications. *Knowledge Engineering Review*, 11(2), 1996, pp.93–155.

[10] Tudorache T., Noy N.F., Musen M.A. Collaborative Protege: Enabling Community-based Authoring of Ontologies. *International Semantic Web Conference*, 2008.

[11] Vargas-Vera M., Motta E., Domingue J., Lanzoni M., Stutt A., Ciravegna F. MnM: Ontology driven tool for semantic markup. *Proc. Workshop Semantic Authoring, Annotation & Knowledge Markup (SAAKM 2002)*, 2002, pp.43-47.

[12] Rogushina, J.V. Development of distributed intelligent systems on base of ontological analysis and semantic wiki technologies. *Ontology of designing. 2017; 7(4): 453-472.*, 2017, pp.453-472.

ОБРАБОТКА СЕМАНТИКИ WIKI-РЕСУРСОВ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА

Рогушина Ю.В.

Институт программных систем НАН Украины, Киев, Украина

Работа посвящена разработке онтологических средств обработки семантически размеченных Wiki-ресурсов, которые позволяют более релевантно отобразить специфичные для предметной области информационные объекты, представленные на отдельных Wiki-страницам.

Предлагаемый подход к семантической обработке ресурсов Wiki базируется на методах онтологического анализа сложных информационных объектов. Онтологический анализ обеспечивает сопоставление Wiki-ресурса с онтологией, что позволяет применять к ним различные мощные механизмы менеджмента знаниями.

Основные результаты: методологические рекомендации для разработки иерархии категорий и семантических свойств Wiki-ресурса; использование семантической разметки для эффективного поиска и навигации; формальное представление семантических элементов Wiki и их сопоставление с релевантной онтологией предметной области.

Результаты предлагаемых исследований используются для разработки электронной версии Большой украинской энциклопедии (е-БУЕ) на технологической платформе Semantic MediaWiki. Обработка семантических свойств и категорий Wiki обеспечивает преобразование е-БУЕ в распределенную базу знаний с расширенными функциональными возможностями.

Ontology Model for Telecom Operator Big Data Representation

Globa L.S., Novogrudska R.L.
The National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
lgloba@its.kpi.ua
rinan@ukr.net

***Abstract*—This paper presents an approach to telecom operator Big Data description using ontology model. Elements of ontology model suggested the description of different data attribute, as well as set semantic correlation of such data. Using them it is possible to organize searching engine and perform logic inference. Special parameters for telecom Big Data are described. The estimation of telecom operator quality based on fuzzy logic is proposed and correlation between it and big data description based on ontology is proposed.**

***Keywords*—telecom operator, Big Data, parameters, ontology, fuzzy knowledge base.**

I. Introduction

At the present development stage of information and communication technologies, the term Big Data means a number of approaches, tools and methods of processing of structured and unstructured vast amounts of data and considerable diversity.

Big Data is a term used to identify data sets. It can't be coped with such data sets using existing methodologies and software because of their large size, rate of arrival, analysis and complexity. Such researchers as M. Hilbernt, S. Strinivasa and others developed methods and software tools for data transmission and information granules mining from Big Data (objects collections formation, that are usually formed for numeric attributes. They are placed side by side because of their similarity, functional or physical commonality), but the appearance of new data formats requires constant expansion and improvement of existing methods and data analysis tools [1].

With that, accumulating data from network nodes rapidly increase every year. This causes necessity in powerful constantly increasing computing resources to increase processing speed and data access. In this regard, necessity to improve large amounts of data processing algorithms becomes more relevance.

According to research were conducted by a number of leading companies in the world [2] telecom operators are faced with the urgent need the complex account of different characteristics of provided services (technical, economic, experience of using services by end-user) due to the rapidly growing range of services and the transition to digital space. They want a clear understanding and process management, that occurring between the operator and its subscribers. This whole range of parameters is too large and complex data for the collecting, processing and analysis with the use of current computing infrastructure and is characterized by:

– significant amount of data (from terabytes to petabytes);

– necessity for high processing speed in real time to reduce the volume of storage;

– heterogeneity (can be structured, unstructured, semi-structured);

– necessity to fulfil a validity requirement (may be disrupted due to the variety of data sources and processing methods, violations of safety requirements);

– importance (using of forecasted methods and analysis methods allows to predict the direction of company's development).

At the same time for assigned tasks data analysis is often performed based on data, which is obtained as a result of the economic operator's activity, or based on sociological interrogations, or based on technical parameters of the operator's infrastructure functioning (for example, the decision to invest in one or another part of the system does not consider the influence and analysis of all possible factors and consequences). For example, if we consider the problem of users' satisfaction degree by services, which are provided by telecom operator, it is quite obvious that the frequent technical failures affect satisfaction degree, price operator's policy and the service performance jointly impact on the final services performance evaluation by subscriber.

Modern facilities of Big Data analysis require a transition from unstructured to structured data, thus forming, relatively speaking, "volume data compression volume to their meaning" and generating data processing strategy for Big Data as "data - information - knowledge - prediction" (Fig. 1.), in this case the entered processing steps are understood as [2]:

– data – streams of raw facts such as business transactions;

– information – clusters of facts that are meaningful to human beings such as making decisions;

– knowledge – data/information organized to convey understanding, experience, accumulated learning, and expertise.

In Fig. 1. it is conditionally shown, that huge amount of data accumulates in time. The posibility of making prediction based on previosly gatherd data (thgougth its trnsformation into information and knowledge) is urgent task nowdays.

Telecommunication companies are investing a lot of money in analytical tools and services development. With such data

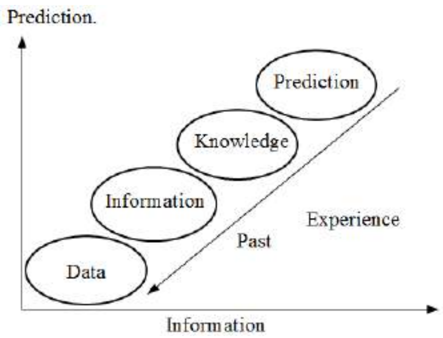


Figure 1. Scheme of the transition from unstructured data to information and to knowledge

telecom operators aim at: increase sales; assure revenue (detect and prevent revenue leakage); reduce churn and fraud; improve risk management; decrease operational cost; improve visibility into core operations, internal processes and market conditions; discern trends and establish forecasts; cross-sell/up-sell products and service plans.

Telecom systems deals with different types of Big Data (*Call Detail Data*, Subscriber (customer) data, Network (operations) data, Others sources data). Another challenge is data scale. As well Big Data influense the proces of telecom analitics management (Fig. 2)

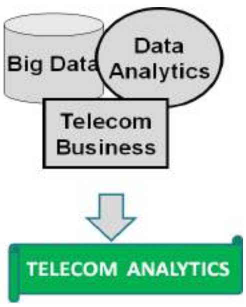


Figure 2. Proces of telecom analitics taking into the account Big Data

Telecom like any other business operates with: customers, services, processes. But, there are unique aspects, such as: network infrastructures, being used as a pipe by other business, regulations. All this shoud be tooked into the account while developing systems for telecom analitics.

Thus, urgent task is to include poverful methods and model of telecom Big Data analises and procesing into general telecom analics system. Main tasks that should be performed are Big Data systematization and stucturing, selection of data groups and identification of semantic connections between them. Such tasks ecxecution allows to gaine more effective data analises and proccesing.

II. Relevant Mathematicals Methods for Big Data

To implement the Big Data collection processes, storage, processing, analysis and forecasting different methods are used, such as: classification [3, 4]; clustering [5, 6]; machine learning [7, 8, 9]; neural networks; support vector machines; decision tree learning; etc.

In [10], it was proposed to consider a fuzzy logic approach that helps to reduce computational complexity in the process of classifying large amounts of telecom data, the efficiency is considered on the example of obtaining a general estimation of quality index of providing services by telecom operator. It is proposed to use decision-making methods based on fuzzy logic. Fuzzy expert rules are formed, which are the basis of the expert system.

Fuzzy rules in such a system can be periodically adjusted to the current status of the technical infrastructure by means of their reformation (refinement) in the process of providing services by telecom operator.

Fuzzy knowledge base (FKB) formation is considered on the example of service quality estimation. Formation algorithm in the generalized form can be presented as follows:

Initial data: measurement tables provided by telecom operators. The measurement table is a set of parameters, which we denote by $X_1...X_n$. It should be noted that $X_1...X_n$ is set of parameters by which telecom operator estimates general state of the system, but none of them characterizes the system quality cumulatively (Y).

To build FKB it is necessary to split measurement table into 3 samples:

1) Training sample with M_1 rows, where $M_1 = \{1, k\}$, which is needed to form fuzzy logical rules of knowledge base. To form FKB rules it is necessary to determine integral quality index, which is obtained using desirability function, we denote it as Y_D;
2) Test sample with M_2 rows, where $M_2 = \{k+1, n\}$, which is needed to check fuzzy logical rules quality of knowledge base;
3) Examination sample with M_3 rows, where $M_3 = \{n+1, m\}$, which is required for the final verification of the correct operation of the obtained FKB.

Generic algorithm for determining an unknown value of Y:

1) Y_D calculation using the desirability function based on M_1 data.
2) FKB formation using M_2 and obtained values of Y_{FKB}. FKB is formed by a set of rules:

$$\text{IF } X_1, X_2, X_3, \ldots, \text{ THEN } Y.$$

Such kind of approach is useful for analyzing measurement data for a number of parameters to circulate in the operator's system but there are a lot of very important factors which have no digital meaning and are able to significantly influence the final result. Taking into account this the paper presents the approach to ontology model of telecom operator Big Data development aiming to optimize the process of such data classification and systematization for its further analytics.

III. Source Data

Source data provide us with different parameters of telecom Big Data. Such parameters were divided on 3 groups according to their meaning and purpose of usage.

First group represents basic notions (objects) that are involved in the communication process:

Provider – deals with information about the service provider.

Service – describes information about the services provided by the provider, services description. It possible to distinguish three types of services: Internet, Television, and Telephony services.

Tariff plan – contains information about the tariff plans of the provider.

Account – describes personal user account information.

User – is directly information about a person who has an account (his / her name, address, etc.).

Bill – describes information about the cash bills associated with this user account. General account deals with three types of actions: get receipt, look through balance, and get invoices.

Second group includes parameters that influence an estimation of quality index of providing services:

– Connection Success Rate – the success of data connections 2G/3G;

– Connection Block Rate – percentage of locks due to overloads 2G/3G;

– Connection Drop Rate – percentage of data connection interruptions 2G/3G;

– PS Attach Success Rate – percentage of successfully completed procedures Attach 2G/3G;

– PDP Context Activation Success Rate – percentage of successful activation procedures PDP Context 2G/3G;

– Speed DL – average speed of HSDPA data transmission to the subscriber;

– Iub Congestion – 3G base station share with high overloads per Iub interface;

– Backhaul Accessibility – availability of zonal transport network;

– DNS Success Rate – successful of DNS resolution;

– DNS Response Latency – DNS resolution time.

Third group represents parameters of the monitoring system are proposed based on the analyses of measurement tables provided by telecom operators:

– COUNT_DAYS_OVER_1MB – the amount of days during which consumed traffic is over 1MB;

– COUNT_DAYS_OVER_5MB – the amount of days during which consumed traffic is over 5MB;

– DUAL_SIM_PROBABILITY – probability of dual sim usage;

– SIM_PRIORITY – the priority of sim card;

– OBLAST – district of sim card usage;

– CITY – city of sim card usage;

– N_SERVICES – amount of services used by permanent subscriber;

– etc.

The consideration of the presented data characteristics leads to the conclusion about their different nature and complexity for the operators to obtain the estimates of their activities. In suggested approach described above data are used for the elements of telecom Big Data ontology definition.

IV. Ontology Model for Telecom Data

The ontology model is developed to realize representation of telecom Big Data. Ontology allows to describe all concepts of problem domain and sets groups of classes, relations of different types, class elements, etc.

Ontology model includes such elements as [11, 12]:

– classes,
– classes attributes,
– relations set on classes,
– types of attributes values,
– class exemplars.

There are two types of relations: the relations between classes and the relations between data types. You can also classify the relations by the following types:

– Internal relations – these are the relations that are inextricably linked with the object;

– External relations are those relations that describe the connection of objects with external objects;

– Relations between instances of this class;

– The relations between instances of different classes from different parts of the hierarchy.

Telecom Big Data ontology classes and their attributes were defined using basic notions that interacts in communication process and source data parameters. Ontology includes next classes: Provider, Service, Tariff Plan, Account, User, Bill, Telephony, Internet, Television, Receipt, Balance, Invoice, Connection Success Rate, Connection Block Rate, Connection Drop Rate, PS Attach Success Rate, PDP Context Activation Success Rate, Speed DL, Iub Congestion, Backhaul Accessibility, DNS Success Rate, DNS Response Latency, SIM_PRIORITY, COUNT_DAYS_OVER_1MB, COUNT_DAYS_OVER_5MB, N_SERVICES, Oblast, City DUAL_SIM_PROBABILITY.

Classes Telephony, Internet and Television are subclasses of class Service. Class Bill is parent class for classes: Receipt, Balance, and Invoice.

Let us describe relations and attributes of basic classes (Provider, Service, Tariff Plan, Account).

The Class ”Provider” is associated with the ”Service” class with the ”has a Service” relation, with the tariff plan class relation ”has a Tariff Plan” and with the ”User” class the relation ”has a relation”. The class has data type relations:

– ”has a name” with the data type string (the name of the ”Provider”);

– ”has a date of creation” with a data type dateTime (the date of ”Provider creation”);

– ”has a description” with the data type string (additional description of ”Provider”).

The ”Service” class has the following data type relations:

– ”named” with string data type (name ”Services”);

– ”has a description” with the data type string (description of ”Services”).

The ”Account” class is related with the ”Account” and ”Tariff Plan” classes in relation to ”Has a Bill” and ”Uses the Tariff Plan” respectively. Also contains the following relations data types:

– ”has login” with data type string (Login ”Account”);
– ”has a password” with the data type string (Account password);
– ”has an email” with the data type string (registered e-mail ”Account”).

The Class ”User” is associated with the ”Account” class with the relation ”has an Account”, with the class ”Service” with respect to ”using the Service” and with the ”Provider” class, the relation ”has a relation”. In the class description, there are the following relations of data types:
– ”has a full name” with string data type (username ”User”);
– ”has passport data” with string data type (Passport data of ”User”);
– ”has address” with data type string (address of ”User”).

Described ontology was developed using one of the tools for semantic modelling – visual editor Protégé 5 [13]. Visual methods of designing ontologies help quickly and fully understand the structure of the subject domain knowledge, which is especially valuable for researchers working in the new subject domain. Protégé 5 supports all phases of the ontology life cycle in accordance with ISO / IEC 15288: 2002 [14] requirements – from the development of a semantic network and the creation of a knowledge base on its basis, to the formation of user requests to these bases in order to obtain knowledge.

The ultimate hierarchical structure of the telecom Big Data ontology was developed by means of the ontologies editor and the framework for constructing knowledge bases Protégé. A component of such ontology is represented on fig. 3.

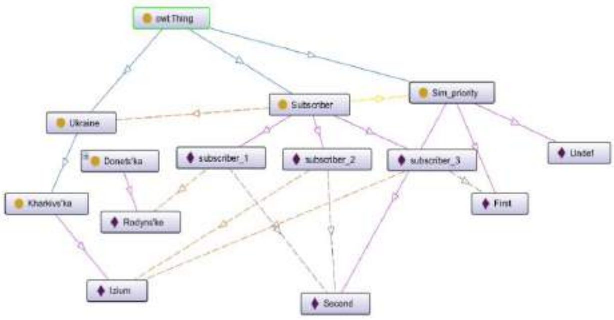


Figure 3. Component of telecom Big Data ontology

V. Conclusions

In the paper the approach to telecom operator Big Data systematizing and structuring is described. It is proposed to use ontology model for this purpose. Different measurement tables provided by telecom operators were analysed. As a result the set of parameters of telecom Big Data were identified. They formed the basis of ontology model.

Selection of Big Data groups and identification of semantic connections between them using ontology model allows to gaine more effective Big Data analises and proccesing.

Further research will be oriented on more detailed telecom Big Data ontology investigation, description of its elements and identification of relationships and regularities. As soon as ontology will be filled with all data the search engine based on logical inference will be tested.

References

[1] Y.Y. Bolyubash, “Methods and tools for processing large data resources in information systems of territorial administration”, Ph.D. dissertation, software for computer systems, Lviv Polytechnic, Lviv, Ukraine, 2017.
[2] M. Ulema. “Big Data and Telecommunications Telecom Analytics” – Tutorial II, BlackSeaCom’2016.
[3] T.O. Ayodele, “Types of Machine Learning Algorithms”, New Advances in Machine Learning. INTECH Open Access Pub., – 2010 – pp. 19–48.
[4] H.A.H. Ibrahim, S.M. Nor, A. Mohammed, A.B. Mohammed, International Journal of Computer Networks and Wireless Communications, – 2012 – 2(1):69–73.
[5] A.K. Jain, M. N. Murty, P. J. Flynn, “ACM Computing Surveys”, – 1999 – 31(3):264–323.
[6] W.A. Barbakh, Y. Wu, C. Fyfe, Studies in Computational Intelligence, Springer, 2009, 249 p.
[7] S. Russell, P. Norvig, Artificial Intelligence: A modern approach, New Jersey, Upper Saddle River, 2010, p. 1154
[8] M. Tim Jones, Artificial Intelligence: A Systems Approach, Hingham, Massachusetts, Infinity Sci. press LLC, 2008, 485 p.
[9] D. Kriesel, A Brief Introduction to Neural Networks. Available at: http://www.dkriesel.com/en/science/neural_networks (accessed 2017, Oct)
[10] Volvach, I.,Globa, L., Method of generalized quality index calculation in mobile networks, Proceedings of 14th International Conference “The Experience of Designing and Application of CAD Systems in Microelectronics”, 2017
[11] Yu. A. Zarogulko, O. I. Borovikova, An approach to constructing knowledge portals, Optoelectronics, Instrumentation and Data Processing, 2008, vol 44, no 1, pp. 75–82
[12] L. Globa, R. Novogrudska, O. Oriekhov Method of heterogeneous information resources structuring and systematizing for Internet portals development, 978-1-4673-2232-4/13/$31.00 ©2013, IEEE, EuroCon 2013, 1-4 July 2013, Zagreb, Croatia, p. 319-326
[13] Protégé – Free, open-source ontology editor and framework for building intelligent systems: Available at http://protege.stanford.edu (accessed 2017, Sept)
[14] Web Ontology Language. Available at: http://www.w3.org/TR/owl-features/ (accessed 2017, Sept)

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ДЛЯ ПРЕДСТАВЛЕНИЯ БОЛЬШИХ ДАННЫХ ТЕЛЕКОМ ОПЕРАТОРА

Глоба Л.С., Новогрудская Р.Л.
Национальный технический университет Украины "Киевский политехнический институт имени Игоря Сикорского"

Аннотация – В статье представлено подход к описанию Больших Данных телеком оператора с использованием онтологической модели. Элементы онтологической модели вводят не только описание различных атрибутов данных, но и описание семантической связи таких данных. Используя онтологическую модель становится возможным организовать поиск информации, а также реализовать логический вывод по данным. Описано специализированные параметры Больших Данных телеком оператора. Предложена оценка качества услуг предоставляемых телеком оператором с использованием нечеткой логики, а также, описана ее взаимосвязь с метаописаниями Больших Данных за счет использования онтологии.

Intellectual variation by penalty coefficients in the algorithm in constructing the contour of the enclosing structure of the heat network in the environment of the building CAD

Sorokin Oleg
Volga State University of Technology
Volgatech
Yoshkar-Ola, Russia
SorokinOlegVolgatech@yandex.ru

Sidorkina Irina
Volga State University of Technology
Volgatech
Yoshkar-Ola, Russia
igs592000@mail.ru

***Abstract*—The construction of a contour in the CAD of enclosing structures (ES) is one of the primary tasks at the stage of visualization of design procedures. Under the contour in this case we mean a closed polygon with the maximum possible coordinates for placement. The contour forms a field for the placement of elements and tracing, so the high-quality design of its boundaries allows better performance of these operations. In the CAD visualization subsystem OK, the contour is built on the basis of analysis of fuzzy information, by selecting penalty coefficients. The intellectual variation of penalties in constructing a contour allows us to introduce additional restrictions on its appearance and avoid problems of typical trace algorithms, such as: a large number of angles in the construction, non-optimal character of the problem zones, and others.**

***Keywords*—indefinite fuzzy information about the thermal circuit, engineering networks, enclosing the contour, intelligent algorithm tuning, automated design of the enclosing structure, penalty coefficients, optimization of the contour laying.**

I. Introduction

The introduction of penalty coefficients will avoid errors in the design of the circuit. The system of penalties, taking into account possible errors in the design process, is an easily customizable tool and allows to improve the design quality, due to the most accurate contour tracing and allowance for forbidden zones. The problem of constructing a contour is NP-complete and can not be solved by polynomial algorithms[1], since the adjustment is performed on the basis of varying the values of penalties depending on the type of situation, when tracing the contour. Each problem situation in the trace is given a certain initial weight, which then varies, depending on the conditions for constructing the contour. The change in the values of the penalty system will ensure the solution of the NP-complete tracing task of the circuit, since the initial data of the contour requirements do not allow accurate tracing by its classical algorithms.[2]

II. The relevance of the algorithm using penalty coefficients in CAD ES

At the stage of constructing the enclosing contour during the design of the engineering network design, a number of problems arise that need to be solved. This stage includes the tasks of: designing the technological environment, selecting the necessary design of the scheme, and arranging the heat equipment. In the course of solving each of these problems, it is necessary to choose the optimal solutions, and as a consequence, it is necessary to compare a lot of possible options and choose the best ones. These tasks, as a rule, are complex and to find the optimum it is expedient to implement the system, based on the use of intelligent algorithms, on the basis of penalty coefficients.[3] To solve the problem of thermal circuit tracing, it is necessary to develop an information system based on mathematical methods for optimization and models for describing existing objects. The result is the total heat flow, the information about which is used in the future to make decisions on the re-arrangement of heaters. The identification of situations in which the thermal circuit is visualized ultimately has only two options: the absence of problem zones (freezing zones) or their presence.The task of intelligent components with penalty coefficients on CAD thermal utilities, to provide the most optimal variant Layout, trace a connection point with the heat, so as to bring the course content of the page as possible to the user. The task of intelligent components with penalty coefficients on CAD thermal utilities, to provide the most optimal variant Layout, trace a connection point with the heat, so as to bring the course content of the page as possible to the user. The result of the algorithm is to visualize the nature of heat flow for a specific situation, taking into account the algorithm of identification of situations and decision-making.[4]

The result of the processing of the measuring information received from multiple sensors - recognition options situations subsequent trace heaters. Hierarchy overlay flows represents a fusion of the two fronts - warm from the heaters and the cold of the contour. The result is the total heat flow, which is then used for making decisions about repackaging heaters. Identifying situations thermal circuit eventually has only two options: no problem zones (zones freezing), or their availability.[5] Define the steps of an improved algorithm for identifying situations and making decisions using penalty coefficients:

- Step 1. Based on the generated a matrix of constraints using the linear programming technique, we form confidence intervals.
- Step 2. We form the constraint matrix for the contour, which determines the boundaries of the effect of the heat flux adjusted on the basis of penalty coefficients.
- Step 3. The matrix "code-solution" is checked for the existence of an adequate solution.
- Step 4. If a solution is found, then by the operation code we decide to visualize the heat flow in a certain way.
- Step 5. If no solution is found or if there are several solutions (fuzzy situation), then using the probability matrix is the solution with the maximum probability.
- Step 6. If there is no single-valued situation when the "code-solution" matrix is polled, the most appropriate frame sector is selected from the value memory by the binary relationship code from the value memory, where the most appropriate situation is then identified by the confidence test.

Thus, the function of choosing a new situation is realized by determining (by the code of binary operations) the most suitable sector of frames in the knowledge base, and also further improved on the basis of penalty coefficients.

III. Algorithm for tracing the enclosing contour in CAD ES using penalty coefficients

When managing penalty factors, uncertainties are also taken into account:

- The type of the thermal circuit. This type of information is fundamental, as it will have an effect on most others. In case of incompleteness of information on the type of the heat circuit, its prediction can be difficult, since the requirements for its construction are part of the terms of reference and are considered, as a rule, separately from the design system, as incoming input information. Inaccuracies of the type of the heat circuit can be corrected by forecasting means, including, in the case when mistakes are made in the construction at the initial design stage. In this case, a message will be issued indicating which construction rules are violated and a variant of their correction will be offered.
- The nature of the distribution of temperature. This information is the most voluminous and, as a rule, embedded in the database of the design system. Nevertheless, for different types of construction[6], with a strong influence of the external environment, the nature of the distribution of the temperature fields inside the structure can not be predetermined and requires the introduction of forecasting elements. An example can be situations where a lot of sources influence the outline of a building, including the external environment, timely calculations for the implementation of interactivity can be used on the basis of probability coefficients. In other words, when calculations of the temperature distribution inside the circuit take a long time, predicted values are used for instantaneous visualization, which allow us to evaluate the design quality at the initial stage, without the need for waiting for complex calculations.
- Mutual influence of heat flows. Information on the mutual influence of heat fluxes is necessary, first of all, to resolve the contradictions in the predicted data in those cases when several heat fluxes affect the same area of the contour at once. Accounting for the mutual influence of heat flows, for example, on heating elements is required when there are several heating elements in one closed loop.

The algorithm used is iterative and is used on the basis of local optimization. The first step of the algorithm is to create an initial approximation, which is included in the range of allowed values, based on the specified constraints in the trace algorithm.[3] In the case when many restrictions are added to the trace task, penalty functions can be selected in the target function, and the initial solution is outside the range of acceptable values. Penalties can be described by the formulas:

$$F'(x) = \Theta(x) + X0(x), \tag{1}$$

$$F(x) = \{ 0 \ , \ yzx = G \sum_{i=1}^{m} \lambda i(mi(x) - ci), \ yzx \notin G \tag{2}$$

where λ_i are penalty coefficients. According to these formulas, the last iteration of optimization is the minimal solution of the function $F(x)$, provided that it belongs to the admissible values $\omega \in G$. In this case, the introduction of penalty coefficients allows tracing the contour with the maximum accuracy, and also with the introduction of an additional unconditional optimization, criteria. It should also be noted that the use of the apparatus of penalty functions allows us to compare the values of the function $F(x)$ that do not belong to the admissible set G, and also in the case when all restrictions are not observed. Further optimization of the circuit can be expressed as:

$$X'' = arg_{x \in k}^{min} F(x) \tag{3}$$

Where k is the number of the row. This formula not allow to completely describe the process of obtaining the initial solution, in which some overlaps are not assigned to the main contour. To implement the initial solution, one k_0 value is entered into each row ki, which corresponds to the cancellation of the selection of the coordinate not belonging to the contour. In addition, it is necessary to take into account the limitations for the initial solution:

$$| \omega_i - \omega_{0i} |> 0, \ i \in (1, N) \tag{4}$$

Also, the function describing the penalty coefficients should be added:

$$F(x) = F'(x) + \sum_{i=1}^{N} \lambda_{i0} * \vartheta(\omega_i, \omega_{0i}) \tag{5}$$

The cycle of operation of the algorithm ends when a local minimum is reached at the N^{th} stage, or when a solution that satisfies the range of admissible values is obtained.[7]

Let's single out the main iterations of the algorithm for tracing the enclosing contour:

- Determination of the initial approximation ω_i
- Formation of the initial solution ω_0
- Going to the next iteration
- Check if the limit is reached on iterations of N if a transition is made to item 7
- Formation of a neighborhood of points in a series.
- Finding the minimum of the function $F(x)$
- If the value does not belong to the set G, go to step 3.
- Completion of the algorithm.

The first step defines the allowable range for all contour boundaries. In step 2, ω_0 is chosen as the initial approximation. Steps 3 and 4 implement the transition to the next optimization step, and also perform a check on the maximum number of iterations of N. Then, in Step 5, the initial solution for the next i-th step becomes the decision taken as locally optimal. At Step 6, a neighborhood of the permissible values for the i-th stage is formed. In Step 7, the reassignment is performed to determine the optimum relative to the neighborhood formed in Step 6. In the case where the solution is possible, the algorithm terminates, otherwise the next iteration is performed.

The task of reassigning a solution $O_i(\omega_i)$ is to determine the number of path paths[8] that can be deleted and rebuilt. If after the i stage, the resulting trace solution is not optimal, then for a certain set of connections j the definition of a new neighborhood oi (ω_i) is performed, which solves the problem of achieving the global optimum, by removing a part of the traces and then re-tracing. To obtain an acceptable solution based on fuzzy information, only pipelines $j = Q$ need to be rebuilt, for which some constraints (4) are not satisfied. Since condition (4) is not fulfilled at the initial stage of the construction, the minimum iteration step is usually chosen for the first iteration of the algorithm, and for each subsequent iteration an increase in step is required, since the number of traces satisfying the constraints grows. It should also be noted that the penalty coefficients on the initial iterations of the algorithm should also have a higher weight, since at the initial stage the number of traces satisfying condition (4) is minimal. The calculation of the penalty coefficients using the variation at the initial iterations of the algorithm will take the form:

$$F_m(x) = F'(x) + \sum_{i=1}^{N} \frac{N+z}{N} \lambda_{i0} * \vartheta(\omega_i, \omega_{0i}) \tag{6}$$

Where z is the number of traces that do not satisfy conditions (4)

The next step, to solve the task of reassignment, is based on the sequential assignment of traces, from the set of permissible ones and adding to the contour. The path is selected by assigning the corresponding trace to all permissible positions in the loop, and then fixing it in the most optimal position.

$$\omega_i' = arg\ min * F(\omega_{ij}) \tag{7}$$

ω_i' is the best position for the trace. $F(\omega_{ij})$ is the value of the criterion for assigning the i-th iteration to the contour ω_{ij}. The use of the reassignment method for tracing a contour makes it possible to find the optimum most accurately with respect to a certain region.

In the case when loop correction is performed on the basis of penalty coefficients, it is also necessary to additionally include new rules for logical inference in the knowledge base, which also require correction:

- Rule1. IF the coordinate of the heat flow distribution of the heater AND the coordinate of the heat flow of the utility network contour are equal to TH, calculate the resultant heat flux.
- Rule2. IF the coordinate of the heat flow of the circuit of the utility network has changed based on the coefficients of penalties, then we calculate the resulting heat flux.
- Rule3. IF the heat flow co-ordinate of the utility network contour was not recalculated automatically after correction based on penalty coefficients, then we calculate the resulting heat flux again.
- Rule 4. If the heat flux of layer N has a temperature less than $N + 1$, then an adjustment is necessary, provided that the check is performed that the contour has not been corrected.
- Rule 5. IF a region of the problem area is found, then build a minimal vector to the circuit of the utility networks to find the problem area.
- Rule 6. IF the nature of the contour is corrected, then rebuild the minimum vector.

Recognition of the situation is necessary for the possibility of setting temperature environment outside the contour, as well as granting councils in the case where the heat flow from the heater can heat the entire circuit. Elements of Learning Technologies Engineering Design of heating networks should be offered in this case, the best option arrangement, to perform communication with a heat tracing point and visualize the resulting heat flow from the heaters in the building loop.[9] The final stage of the work is to visualize the nature of heat flow for a specific situation, taking into account the algorithm of identification of situations and decision-making adaptive presentation. The goal of adaptive presentation technology is to adapt the content of a hypermedia page to the objectives, knowledge and other information stored in the user model. In a system with adaptive presentation, the pages are not static, but generated or assembled from pieces for each user. For example, expert users receive more detailed and in-depth information, while novices receive more additional explanation. Using fuzzy logic and methods for the synthesis and identification of situations in the electronic course, possible to apply the algorithm of identification of situations and decision-making, based on adaptive representation of information that takes into account all the possible outcomes in the data visualization. Using complex algorithms, the data will allow the most accurate visualization of heat flows in CAD interface ES, which will increase the accuracy of calculations and impact in the future on the quality of the layout of the heating elements in the design of the circuit.[10]

IV. Conclusion

An intelligent algorithm for constructing a contour in the CAD ES visualization subsystem of enclosing structures is proposed, based on penalty coefficients, which allows to optimally solve the NP-complete tracing problem even on the basis of fuzzy information. The algorithm is built on the basis of the principles of local optimization with the transition to the problem of unconditional optimization, which allows you to take the admissible values that are not included in the range of possible ones. The use of penalty coefficients allows to organize the optimization process, to form an optimum with respect to a certain set of data, and also to perform further optimization on the basis of reassignments.

References

[1] Sorokin O.L., Sidorkina I.G NP-zadacha balansirovki nagruzochnoi sposobnosti seti v SAPR [NP-task of balancing the load capacity of the network in the CAD] Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem OSTIS-2015 [Open semantic technologies for the design of intellectual systems OSTIS-2015],2015, pp. 585-588.

[2] Vdovichenko I. Osobennosti sozdaniya bazy znanii standartnykh metodov provedeniya ekspertizy [Features of creating a knowledge base of standard techniques for conducting an examination] Zhurnal peredovykh tekhnologii [Journal of Advanced Technologies],2010, No. 2, vol.6. pp. 13-16.

[3] Sorokin O.L., Sidorkina I.G. Modul' opredeleniya statsionarnogo rezhima v SAPR naruzhnykh inzhenernykh setei [The module for determining the stationary mode in the CAD of external engineering networks] IS i IT-intellektual'naya CAD 2015 [IS & IT-Intellectual CAD 2015],2015 No.1. pp.70-75.

[4] Stetsko A.A. Sistema modelirovaniya i proektirovaniya trafika telekommunikatsionnykh setei v usloviyakh neopredelennosti [System for modeling and designing traffic of telecommunication networks in conditions of uncertainty] Informatika, sistemy iskusstvennogo intellekta i modelirovanie tekhnicheskikh sistem [Informatics, artificial intelligence systems and modeling of technical systems], 2006. No2. pp.109-111.

[5] Vashkevich N.P., Dubinin V.N. Struktura i funktsional'nye vozmozhnosti SAPR tokarnoi obrabotki [The structure and functionality of the CAD turning] Novye informatsionnye tekhnologii i sistemy [New information technologies and systems], 2004. p.143-144.

[6] Soloviev V.I. Metody optimal'nykh reshenii [Metody optimal'nykh reshenii] Praktikum uchebnoe posobie [Workshop manual], 2017. pp.50-55.

[7] Nomerchuk A.Y Metody upravleniya avtomatizirovannymi sistemami obrabotki dannykh na osnove nechetkoi logiki [Methods of automated data processing systems control based on fuzzy logic] Inzhenernyi SFU [Engineering SFU], 2013. pp 56-60

[8] Fokin K.F. Teplotekhnicheskii ograzhdayushchie chasti zdanii [Thermal engineering enclosing parts of the buildings] Izdatel'stvo Avok Press [Avoc- PRESS], 2006. p.256.

[9] Lebedev B.K. Intellektual'nye protsedury sinteza topologii SBIS [Intellectual procedures for the synthesis of VLSI topologies] Izdatel'stvo TRTU [TRTU publishing], 2003, pp.60–68.

[10] Kureichik V.V Osobennosti sredy analogovogo proektirovaniya VIRTUOSO [Features of the analog design environment VIRTUOSO] Intellektual'nye SAPR [Intelligent CAD], 2006. p. 105

ИНТЕЛЛЕКТУАЛЬНОЕ ВАРЬИРОВАНИЕ ШТРАФНЫМИ КОЭФФИЦИЕНТАМИ В АЛГОРИТМЕ ПРИ ПОСТРОЕНИИ КОНТУРА ОГРАЖДАЮЩЕЙ КОНСТРУКЦИИ ТЕПЛОВОЙ СЕТИ В СРЕДЕ СТРОИТЕЛЬНОЙ САПР

Сорокин О., Сидоркина И.

Аннотация - Построение контура в САПР ограждающих конструкций (ОК) является одной из первоочередных задач на этапе визуализации проектных процедур. Под контуром в данном случае понимается замкнутый многоугольник с максимально возможными координатами для размещения. Контур образует поле для размещения элементов и трассировки, поэтому качественное проектирование его границ позволяет более качественно выполнять данные операции. В подсистеме визуализации САПР ОК контур строится на основе анализа нечеткой информации, путем подбора штрафных коэффициентов. Интеллектуальное варьирование штрафов при построении контура позволить ввести дополнительные ограничения на его вид и избежать проблем типовых алгоритмов трассировки, таких как: большое количество углов при построении, неоптимальный характер проблемных зон и других.

Ключевые слова - неопределенная нечеткая информация о тепловом контуре, инженерные сети, ограждающий контур, интеллектуальная настройка алгоритма, автоматизированное проектирование ограждающей конструкции, штрафные коэффициенты, оптимизация прокладки контура.

Intelligent Data Analysis: from Theory to Practice

Tatur M.M.
Belarusian State University
of Informatics and Radioelectronics
Minsk, Republic of Belarus
tatur@bsuir.by

Iskra N.A.
Belarusian State University
of Informatics and Radioelectronics
Minsk, Republic of Belarus
niskra@bsuir.by

***Abstract*—In this paper common issues in practical intelligent data analysis are addressed. Recommendations for productive practical applications and development of intelligent data analysis systems are proposed.**

***Keywords*—data analysis, structured data, knowledge acquisition, statistical analysis, machine learning, data visualization**

I. Introduction

The term "data" in the common sense is a very broad concept, and "Intelligent Data Analysis", if desired, can be used almost in any area of information processing. Therefore we initially want to concretize (restrict) the concept of "data", "data analysis", and further "intelligent data analysis".

In this paper by "data" we refer to the ***structured data***, i.e. vector of measured parameters describing an object, phenomenon or event. For example:

- While paying at the store you "leave" data about the time of your visit, the purchased goods with the corresponding amounts. In case you use a buyer's card, this transaction is personalized (i.e. it can contain additional information such as age, sex, place of residence, birthday, etc.). Obviously, this information can potentially be used to plan and carry out marketing programs to retain clients, promotion-actions, etc.
- Each call to emergency services (911) is recorded and initiates data collection: the time, place, cause of the incident and actions to be taken, as well as the results of these actions or consequences. Such data are accumulating constantly and can be used to generalize and form useful conclusions for practice.

Under the definition above do not fall unstructured and weakly structured data, for example, signals, images, video data, texts, etc., that require an additional stage of preprocessing.

The simplest case of structured data processing (apart from counting, extracting, copying, updating, etc.) is ***statistical analysis*** [1]. For example, it is easy to calculate how many visitors to the supermarket made purchases in a certain period of time, what is the average purchase price, how much did the revenue increase or decrease in the analyzed periods, etc. Such statistical analysis requests are formed by the user and are aimed at covering the current issue (problem). Such analysis, as a rule, operates with one of the parameters, while other parameters are fixed or predefined.

However, very often in order to obtain more complex information about an object, event, or phenomenon, it is necessary to carry out a deeper analysis on data, taking into account many parameters, and in the worst case - all parameters available. Thus, in each industry there are historically gained and constantly accumulated specific knowledge, that are associated with a deep data analysis. This knowledge is represented by complex characteristics that link a number of measured parameters, as well as laws and regularities related to these parameters.

For example, in economics, the efficiency of the enterprise is characterized by the enterprise's profit index for a certain period of time. This index combines information on incomes and expenditures. Incomes and expenditures in their turn are also derivative characteristics that are calculated (extracted) from a whole series of primary data. It is even more difficult to calculate such a complex index as a labor productivity. Economic laws of breakevenness of the enterprise, a stock of circulating assets, etc. are well-known. Similar complex indexes and indicators can be found in medicine, sociology, criminology and other fields [2]. This kind of knowledge is undoubtedly the result of the intelligent data analysis carried out by specialists in particular industry, relying on previous experience, empirical observations, hypotheses formation and testing, heuristics, and the like. In this context, "intelligent data analysis" is a special case of "data analysis", and there is hardly a clear boundary between trivial (simplest) and intelligent analysis, at least this issue can be classified as rhetorical.

Note: By this we don't mean knowledge structuring, formal knowledge representation or formal knowledge management, which are usually referred to as ***artificial intelligence***.

In Fig. 1 a mnemonic diagram of data analysis is given. It can be divided into a trivial and an intelligent analysis. The meaning of the picture is that the user generates a database query (request) to obtain information (and, ultimately, knowledge) about an object or phenomenon. Based on this query, the data is analyzed: trivially and / or intelligently. If to obtain knowledge of an object or a phenomenon (for example, about the production efficiency), it will be sufficient to formally apply already acquired knowledge (for example, the formula for calculating the labor productivity, without any additional interpretation and explanation), this analysis hardly could be attributed to the intelligent type. In Fig. 1

this circumstance is marked by a dotted arrow.

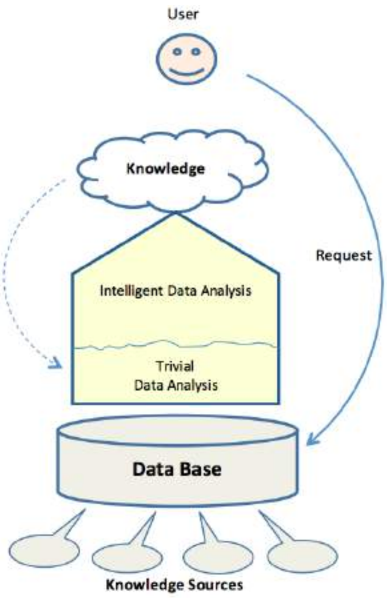


Figure 1. Place and gradation of data analysis of the certain object or phenomenon in user-oriented knowledge acquisition system.

Investigating the processes of obtaining new knowledge (hidden dependencies, regularities, trends, etc.) from structured data, developing and formalizing the mathematical apparatus, the scientific direction of the intelligent analysis of data was gradually formed. This direction has many close (rather, synonymous) names: Decision Making, Machine learning, Pattern Recognition, Data Mining, Knowledge Discovery, Advanced Data Analysis. This area is constantly updated with new terms, for example, Deep Learning, Big Data, Data Science, etc., a number of which, unfortunately, are of a mere conjunctive nature.

Today, processing and data centers, in which huge data flows are accumulated, are being created everywhere. Each major organization, each department collects, stores, protects and provides data, maintains databases up-to-date. For this purpose, software and hardware processing technologies, such as SPARK, Hadoop, Cassandra, EMC, etc., which contain tools for both trivial and intelligent analysis, are used industrially.

However, in view of our previous collaboration experience with data center representatives and direct database users [3], it turns out that in the vast majority of cases intelligent data analysis based on modern scientific achievements and computer technologies is almost never applied. As a result, databases exist and are constantly accumulating new data, but data analysis is still conducted by means of trivial methods. This, in turn, limits the value and efficiency of the useful information obtained.

Further in this paper we analyze the reasons for such state of affairs and propose an informal approaches (recommendations) to activate the practical application of the intelligent data analysis apparatus.

II. Basic Intelligent Data Analysis Terminology

Analysis of the mathematical methods and algorithms underlying the intelligent data analysis, shows that the said directions (Machine Learning, Advanced Data Mining, etc.) closely overlap. In this case, the basic concepts remain:

- feature (or informative feature) – in essence, the measured parameter, a unit of structured data;
- pattern – a vector of informative features.

Then the structured database is nothing more than a repository of patterns. And with the patterns, typical formal processing tasks can be performed: clustering, classification, ranking, forecasting, associative search, regression and some others.

These tasks can be solved by a limited list of formal methods or algorithms, such as neural network of a certain type, k-means, SVM, k-nearest neighbors, decision trees, etc. Most algorithms are based on the principle of measuring the distance between patterns using different measures (or metrics).

Some algorithms rely on the principle of differentiation or separation of patterns. These algorithms, as well as metrics for estimating the distance between patterns and ways of pattern differentiation, have long been described and sufficiently investigated in the textbook literature [4]. Moreover, most algorithms are already implemented in the form of program libraries of a number of systems and language environments, such as R, Weka, Wolfram Mathematics, Caffe, Tensorflow, cuDNN, scikit-learn, etc. It remains only to learn how to use this mathematical apparatus [5].

However, this often does not happen. The reasons for such a "modest" practical application of the mathematical intelligent data analysis apparatus is evidently due to:

- on the one hand, the problem of developing applied intelligent data analysis systems, or rather, the domestic experience of creating such systems is practically absent;
- on the other hand - the problem of a potential user who is not ready to master the achievements in the field of "Data Science" and solves pressing problems at the level of trivial data analysis.

Let's try to understand the reasons and give some advice or recommendations.

III. About Some Data Analysis "secrets" openly ...

Based on personal long-term teaching experience in one way or another related to intelligent data analysis [6], as well as on the experience of scientific papers, articles and dissertations examination and research in this field, the following observations and conclusions were made.

A. *Variety of Methods and Algorithms*

Sometimes developers are simply lost in a variety of methods and algorithms related to the field of data mining, and the terminology used, which abounds with overlays and collisions and only aggravates the problem of entering the field of research.

Recommendation. We recommend using the ontology of the Intelligent Data Analysis presented in Fig. 2. It is limited to two or three most simple for understanding tasks (for example, ranking, classification and clustering) with studying two or three most popular algorithms for solving each of the problems. Moreover, for the solution of the applied problem, in 99 cases out of 100 there is no need to develop an original method. The main thing is to master and learn how to effectively apply the already accumulated arsenal of methods and algorithms. Moreover, to create an original algorithm is half the battle, it is necessary that the created or modified algorithm shows positive winning qualities in comparison with the known ones. Carrying out an objective comparative analysis of algorithms represents an independent scientific problem, the discussion of which is beyond the scope of this thesis.

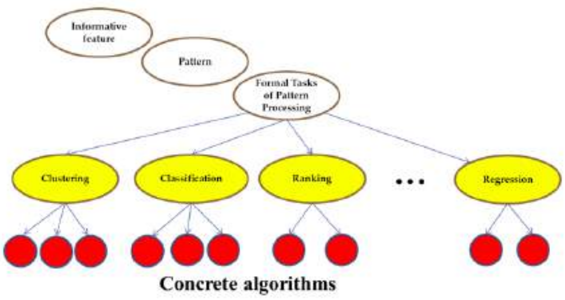


Figure 2. Key terms and typical intelligent data analysis tasks.

B. "Blind" Application

Unfortunately, some of the work of aspiring scientists is presented as a formal ("blind") application of one of the methods or algorithms of intelligent data analysis, followed by sometimes incorrect result interpretation in "their own favor".

Recommendation. The algorithmic complexity of solving an applied data analysis application usually goes beyond the single data analysis algorithm (machine learning). In general case, when solving problems from a particular domain, a chain of algorithms is used, both from the data analysis list and the conventional deterministic ones. At the same time, the task as a whole is informal, and an important role is assigned to specialists in data science who must correctly formalize the initial problem, correctly apply the mathematical (algorithmic) apparatus of the intelligent data analysis and correctly interpret the result.

C. "Small Data" Analysis

To date, a lot of data for analysis has been accumulated. But often important conclusions need to be made using only small portions (we shall call it "small data", as opposed to "Big Data"). For example, in medicine it is very hard to collect a large amounts of examples, it could be risky, expensive and, often, virtually impossible.

Recommendation. For such tasks it would be good to have metrics that would allow us to evaluate so-called "representativeness" of the training sample, and show how relevant the conclusions can be [7].

D. Data Visualization

Traditionally, the raw data and processing results are displayed as marked points on a two-dimensional plane. At the same time, it is often overlooked that in practice the number of informative features, and therefore the dimensionality of space is much higher and can reach tens and hundreds of parameters. On the other hand, data analysis begins with the visualization of data. On 2D or 3D visualization data specialist can allocate patterns, get an "insight" about the composition and internal structure of the data under examination. Very important is the task of correct and understandable visualization of the data under study, preferably without loss or with controlled losses.

There are methods of visualization that allow to compare the projections of data on the 2D or 3D plane. Insight about the data can also be obtained by analyzing the histograms, viewing the visualizations interactively. Data can be cleaned, normalized and transformed (for example, the principal data analysis), or represent data in the form of curves where the informative features are the coefficients of the curves, and it can be judged about the internal data structure by the shape of these curves [8]. In this case data will not be lost. Example of such visualization is shown in Fig. 3

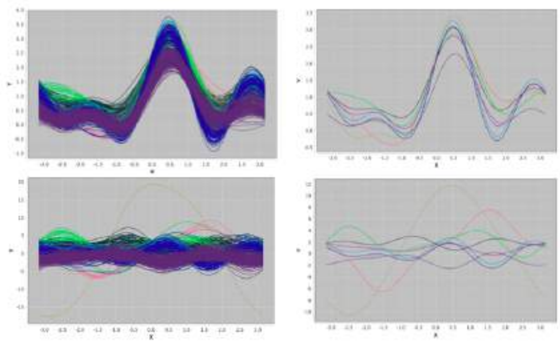


Figure 3. Andrews curves for multidimensional data visualization.

Recommendation. We see the following path in the direction of data visualization:

1) Simple methods using part of the information – the entire picture is not visible, it is difficult to draw conclusions.
2) Methods using all data as coefficients – sometimes difficult to interpret.
3) Methods using preliminary data transformation, representing the most informative aspects of data (the development of such methods can themselves be the object of intelligent data analysis: which features to choose, how to convert them, how to place and / or group images, etc.)

4) Interactive methods that allow data analysts to interact directly with real-time data in 2D or 3D, and in the form of augmented or virtual reality [9]. These methods can and should take advantage of the previous groups of methods - the simplicity and intuition of the first group, the completeness of the second and the "intelligence" of the third.

E. Practical Use

And, most importantly. Only the manager, who is highly interested in obtaining objective results, can formulate the actual task of data analysis in a specific subject area. For example, the director of a trading network is interested in knowing: "What is the reason of the outflow of customers in the current (or, certain) period?" It is assumed that there is no direct connection between this event and supplies, or the dismissal of any of the personnel. Another example, for the adoption of strategic management decisions it is necessary to evaluate and rank alternative versions of legislative initiatives (projects), relying on the statistics of previous periods, etc. In fact it can be revealed that:

- the director is not ready to formulate questions (queries) that go beyond trivial statistical analysis of data;
- the request goes beyond the possibilities of solving methods and means of data analysis, both for objective reasons (for example, an insufficiently representative database, insufficient time resources for research, etc.), and for subjective reasons (for example, the specialist's skill level in the analysis of data does not allow to solve the problem).

Recommendation. It is necessary to establish close cooperation between the user and the data analysis specialist. The user must understand the possibilities and know the objective limitations of data analysis technology. The expert analyst should know the terminology and problems of the subject area.

It is impossible to oppose the computer technology of intelligent data analysis to traditional research, i.e. intelligent analysis of data by an application specialist using trivial statistical analysis.

It must be borne in mind that the technology in question is not a silver bullet; not all tasks can be solved using intelligent data analysis, and for a number of tasks the application of this technology is unnecessary.

IV. Conclusion

We named only some of the most significant, in our opinion, aspects of intelligent data analysis. Of course, there are much more problematic issues worthy of discussion and research. These include: the problem of processing large amounts of data, the problem of providing (and / or evaluating) the representativeness of the data sample, as well as the reliability of the results of intelligent data analysis, the problem of the technical implementation of research and application systems.

A team of engineers and scientists specializing in this field has been working at the computer department of the Belarusian State University of Informatics and Radioelectronics for more than 10 years. At the department there is an innovative enterprise OOO "Intelligent processors". The team's assets include the creation of the first Belarusian neuro-like (neuronegal) computer (2010) [10], more than 30 publications on the research topic. In the period 2016-2020, research on the "Intelligent computing system for processing large data" is being carried out, with the goal of creating a tool for building applied data analysis systems. Our competencies include the development of algorithms and software to meet the customer's requirements, including participation in the formalization of the formulation of the problem, the choice of solutions and the interpretation of the results.

We are currently moving from academic research to practical applications and are starting to work with potential users from the organizations of the Ministry of Emergency Situations, the State Customs Committee, the State Committee for Forensic Expertise, the Center for System Analysis and Strategic Studies of the National Academy of Sciences of the Republic of Belarus.

References

[1] P. Buhlmann, P. Drineas, M. Kane, Mark van der Laan,"Handbook of Big Data," CRC Press, 2016, 464 p.

[2] S. Pyne, B.P. Rao, S.B. Rao, "Big Data Analytics: Methods and Applications," Springer, 2016, 276 p.

[3] A.V. Zhovna, V.M. Prorovskii, M.V. Khodin, N.D. Chistyakov, T.A. Kornacheva, "Analiz obstanovki s chrezvychainymi situatsiyami v Respublike Belarus' v 2016 godu [Analysis of the situation with emergency situations in the Republic of Belarus in 2016]," Chrezvychainye situatsii: preduprezhdenie i likvidatsiya [Emergency situations: prevention and elimination], No1(41), 2017, pp. 24–31.

[4] D. Cielen, A. Meysman, M. Ali, "Introducing data science: Big Data, Machine Learning, and more", Manning Publications, 2016, 325 p.

[5] A. Muller, S.Guido, "Introduction to Machine Learning with Python: A Guide for Data Scientists," O'Reilly Media, Inc., 2016, 394 p.

[6] A.I. Demidchuk, D.Yu. Pertsev, M.M. Tatur, D.V. Krishtal', "Intellektual'naya obrabotka bol'shikh ob"emov dannykh na osnove tekhnologii MPI i CUDA [Intellectual processing of big data volumes based on the MPI and CUDA technology]," Minsk, BGUIR, 2017, 60 p.

[7] V.V. Iskra, N.A. Iskra, M.M. Tatur, "Vliyanie statisticheskikh kharakteristik obuchayushchei vyborki na ee reprezentativnost'[The influence of the statistical characteristics of the training sample on its representativeness]," Nauchno-tekhnicheskii zhurnal Iskusstvennyi intellekt [Scientific-technical journal Artificial Intelligence], No4(62), Donetsk, 2013, pp. 325–332.

[8] C. García-Osorio, C. Fyfe, "Visualization of High-Dimensional Data via Orthogonal Curves," Journal of Universal Computer Science, 11 (11), 2015, pp. 1806-–1819.

[9] M. Huh, K. Kiyeol, "Visualization of multidimensional data using modifications of the Grand Tour," Journal of Applied Statistics, 29.5, 2012, pp. 721–728.

[10] D.N. Odinets, D.A. Lavnikevich, D.I. Samal', V.V. Starovoitov, "Postroenie informatsionnykh intellektual'nykh sistem na osnove neiropodobnogo komp'yutera [Building information intelligence systems based on a neuron-like computer]," Minsk, 2013, 43 p.

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ: ОТ ТЕОРИИ К ПРАКТИКЕ

Татур М.М., Искра Н.А.

В настоящей работе рассматриваются проблемы практического применения интеллектуального анализа данных. Даются рекомендации по созданию прикладных систем интеллектуального анализа данных.

Associative Processor as Means of Disjoint Sets Representation and Dynamic Connectivity Problem Solving

Nikolai L. Verenik
and Mikhail M. Tatur
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
nick.verenik@gmail.com
tatur@i-proc.com

***Abstract*—The article briefly describes data structures and algorithms for working with disjoint sets. The use of an associative processor with an original architecture for representing disjoint sets and solving the dynamic connectivity problem with the help of adapted quick-find algorithm is proposed. The presented approach allows to achieve constant running time of basic union-find operations and linear complexity for processing sequence of union-find operations in any order.**

***Keywords*—disjoint set, union-find, quick-find, parallel computing, vector processor, associative memory**

I. Dynamic connectivity problem

We are given sequence of pairs of integers where every integer represents an object of some type and the pair p-q means "p is connected to q". Assume that the relation p-q is *equivalent*, i.e. has the following properties:

- *reflexivity* (p is connected to p);
- *symmetry* (if p is connected to q, then q is connected to p);
- *transitivity* (if p is connected to q and q is connected to r, then p is connected to r).

The problem is to develop an efficient data structure for storing sufficient information about the input pairs to be able to decide at any time whether an arbitrary pair of objects p and q is connected. The example of processing a sequence of pairs is presented in Fig. 1.

When processing each new pair of objects from the input, we need to determine whether it represents a new connection, and then integrate the information about the detected connection into the data structure to be able to test connections in the future.

This problem is actually a fundamental computational task that arises in a variety of applications, from percolation in physical chemistry to connectivity in large communications networks. These are few of subject areas where the data structure of such kind can serve to represent objects [1]:

- computers in a network;
- friends in a social network;
- transistors in a chip;
- pixels in a digital image;
- variable-name-equivalence problem in some programming languages;
- metallic sites in a composite system;
- elements in a mathematical set etc.

The difficulty with these applications is a possible need to process millions of objects and billions or more of connections. Thus, the following properties are common for the given problem:

- number of objects N can be huge;
- number of operations M can be huge;
- find and union commands are called in random order.

The dynamic connectivity problem can be reduced to the following two abstract operations that we need to implement:

1) Connected(p, q) - determines whether the pair of objects p and q is connected.
2) Union(p, q) - replaces the sets containing the objects p and q by their union. We define *connected components* as maximal set of objects connected with each other.

Thus, the procedure for processing the input sequence of pairs of integers can look like this:

for all (p-q) **do**
    if not Connected(p, q) **then**
        Union(p, q)
    end if
end for

II. Disjoint sets and operations on them

The generalization of the connectivity problem is *disjoint-set data structure* (or union-find data structure). Such a structure maintains a collection $S = \{S_1, S_2, ..., S_k\}$ of disjoint dynamic sets, where each set is identified by a *representative*, which is a certain element of the set. The element that used as a representative depends on the applied problem and doesn't always matter.

The data structure for disjoint sets should support the following operations [2]:

- Make-Set(p) creates a new set consisting of one element (and thus representative) that is p. Since sets are

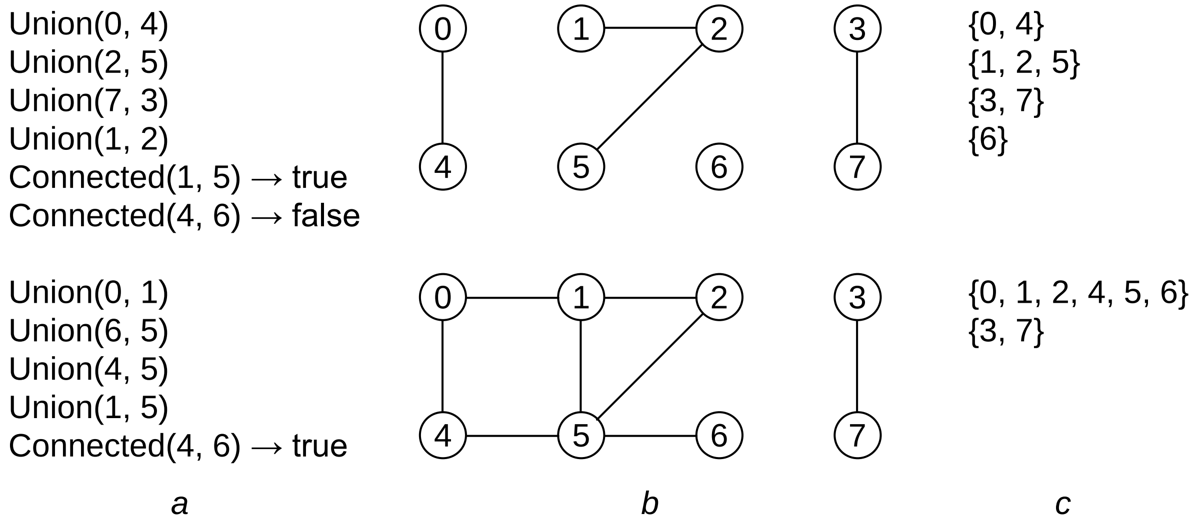


Figure 1. An example of processing a sequence of pairs of integers. The sequence of the executed commands (a); graphical representation of the data structure (b); connected components (c)

disjoint, it is required that p is not already in some other set.

- UNION(p, q) unites dynamic sets that contain p and q (denoted by S_p and S_q) into a new set. Assume that the two sets are disjoint prior to the operation. The representative of the resulting set is any element of $S_p \cup S_q$, although many implementations of UNION choose the representative of either S_p or S_q as the new representative. Since we require all sets to be disjoint, UNION operation must conceptually destroy the sets S_p and S_q, removing them from the collection S. In practice, the elements of one of the sets are absorbed into the other set.
- FIND-SET(p) returns a pointer to the representative of the (unique) set which contains the element p.

The problem of implementing disjoint-sets data structure has been well developed by a number of researchers. The convention that both elements and sets will be identified by integer values between 0 and $N - 1$, so a simple linear array $id[]$ is used as basic data structure to represent the sets. Initially, we start with N sets, each element in its own set, so we initialize $id[i]$ to i for all i from 0 to $N - 1$. For each element i, we keep the information needed by FIND-SET(p) method to determine the set containing element i using various algorithm-dependent strategies. As a result CONNECTED are reduced to simple check FIND-SET(p)== FIND-SET(q).

The running time of the disjoint-set data structures is usually analyzed in terms of two parameters: N, the number of MAKE-SET operations, and M, the total number of MAKE-SET, UNION and FIND-SET operations. Assume that N MAKE-SET operations are the first N operations performed (during the initialization process), and the number of UNION operations cannot exceed $N - 1$ (since the sets are disjoint by definition). Table I provides a summary of the most widely known algorithms with their worst-case running time cost [1].

III. QUICK-FIND ALGORITHM IMPLEMENTATION USING ASSOCIATIVE PROCESSOR

Consider the *quick-find* algorithm. The basis of this algorithm is an array of integers $id[]$ with the property that pair of elements p and q are connected if and only if the $id[p]$ and $id[q]$ values of array are equal. This method is called quick-find because FIND-SET(p) just needs to return $id[p]$ to complete the operation. Then to implement UNION(p, q) it is enough to replace all entries in the array corresponding to both sets by the same value, yet for that we need to go through the whole array.

Fig. 2 shows an example of processing the sequence of pairs of integers used in Fig. 1. The code is quite straightforward:

```
UNION(p, q)
r ← id[p]
for i = 0 to N − 1 do
    if id[i] = r then
        id[i] ← id[q]
    end if
end for
```

```
MAKE-SET(p)
id[p] ← p
```

```
FIND-SET(p)
return id[p]
```

As far as we can see the main disadvantage of the quick-find algorithm is UNION operation which requires seeing the entire data array $id[]$. It will take N^2 accesses to the array in order to

Table I
M UNION-FIND OPERATIONS ON A SET OF N OBJECTS

Algorithm	MAKE-SET	UNION	FIND-SET	Worst-case time
Quick-find (QF)	N	N	1	MN
Quick-union (QU)	N	N	N	MN
Weighted QU (WQU)	N	$\lg N$	$\lg N$	$N + M \lg N$
QU + path compression (QUPC)	N	$\lg N$	$\lg N$	$N + M \lg N$
WQU + path compression (WQUPC)	N	$\lg N$	$\lg N$	$N + M \lg^* N$

*Iterated logarithm

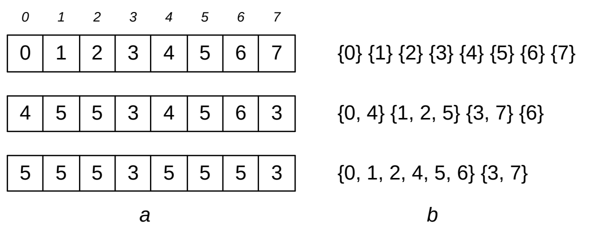


Figure 2. Representation of data in the quick-find algorithm. Elements in data array (a); connected components (b)

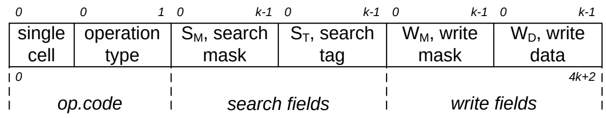


Figure 3. The associative processor command

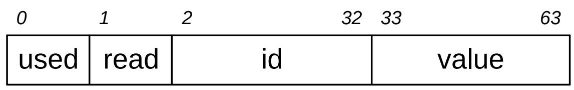


Figure 4. Interpretation of the memory cell data (bitwidth $k = 64$ bits)

process a sequence of N union operations for a structure with N objects, what is prohibitively expensive. That is the reason why quick-find can not be used to solve real-life problems with a huge number of objects and better algorithms (WQUPC, see Table I) are used. However, if we assume that the operation of writing to the data array can be performed synchronously for all elements in constant time, then the efficiency of this simple algorithm will drastically change.

In the articles [3] the architecture of the developed associative processor was considered. In essence, it is a SIMD class processor in which all memory is evenly distributed between a set of PE. Each PE is responsible for accessing its memory cells, implemented as a function of comparing the contents of the cell with the input word. As a result, the entire sequence of simple PEs can synchronously perform operations on all memory cells or over a selected set of associative memory words. We can consider the processor of this type as an intelligent memory with a specific interface for data access. Like an ordinary memory, all the functionality that the processor provides is reading and writing data.

Fig. 3 shows the only command of the associative processor. To address memory cells instead of a fixed address a search operation is used where the input parameters are the S_M and S_T fields, the mask and the search tag respectively. The PE compares the contents of each cell with the value of the search tag S_T according to the bit mask S_M. In case of equality, the cell is considered active and the selected binary operation (field *"operation type"*) is applied to it, in the simplest case corresponding to assignment . The write operation is also not applied to all bits in the memory cell, but only to the marked by W_M mask. The bit values are taken from the W_D field. The *"single cell"* command bit is used when it is necessary to guarantee the activation of only one memory cell, which is necessary, for example, to implement sequential readout of data from the processor memory.

It is proposed to use the associative processor of this type as a basis data structure when solving the dynamic connectivity problem by the quick-find algorithm.

The data format of the processor's memory cell is shown in Fig. 4. It is not used by associative processor itself as processor operates only at bit-level without resorting to upper-level abstractions. Thus, this is only one of the options for interpreting the cell data by a programmer, chosen for the convenience of solving a particular problem. The *"used"* bit is used to identify free memory cells so we can initialize them one by one. The *"read"* bit for a given problem is not used, but is reserved for implementing data readout from the processor. The *"id"* field corresponds to the element identifier or array index in the basic implementation of the quick-find algorithm. The *"value"* field contains the representative of the set that contains the element. With the 64-bit memory cell specified in the example, it is possible to address $2^{31} = 2$ billion objects.

All necessary operations for working with disjoint sets data structure can be implemented at the application level. Table II presents an example of such an implementation for the selected data format (Fig. 4). For example, INIT reset S_M mask field to zero so it won't be used (hence it doesn't matter what we send in S_T field) and all memory cells will be addressed. Similarly, W_M mask field is set to 0xFF to fill these addressed cells

Table II

OPERATIONS FOR WORKING WITH A DISJOINT SETS DATA STRUCTURE IMPLEMENTED USING THE ASSOCIATIVE PROCESSOR

	s.cell	op.type	S_M				S_T				W_M				W_D			
			used	read	id	value	used	read	id	value	used	read	id	value	used	read	id	value
1	0	00	0	0	0x00	0x00	-	-	-	-	1	1	0xFF	0xFF	0	0	0x00	0x00
2	1	00	1	0	0x00	0x00	0	-	-	-	1	0	0xFF	0xFF	1	-	p	p
3	1	00	1	0	0xFF	0x00	1	-	p	-	0	0	0x00	0x00	-	-	-	-
4	0	00	1	0	0x00	0xFF	1	-	-	id[p]	0	0	0x00	0xFF	-	-	-	id[q]

[1]INIT() fills memory cells with zeros.

[2]MAKE-SET(p) writes the element p to the first available memory cell.

[3]FIND-SET(p) returns the last addressed memory cell that is $id[p]$.

[4]UNION2($id[p], id[q]$) writes $id[q]$ value to all memory cells where $id[p]$ is stored.

with zeros (see W_D). The other thing to mention is that union method is a bit different from the initial one (here it is called UNION2). The method takes the two representatives of the sets to unite instead of the original pair of elements identifiers.

Listing 1. Quick-find algorithm using the associative processor (C++)

```
void main() {
  Init();
  for (int i = 0; i < n; ++i) MakeSet(i);
  while (std::cin >> p >> q) {
    int r1 = FindSet(p);
    int r2 = FindSet(q);
    if (r1 != r2) Union2(r1, r2);
  }
}
```

CONCLUSION

A weighted quick-union algorithm is generally used for solving disjoint-set problems. It has number of various modifications (such as path compression) and as a result worst-case running time $O(N + M \lg N)$ is achieved. Analysis can be improved to $O(N + M\alpha(M, N))$ [4] [5], where $\alpha(M, N)$ – a very slowly growing function. In practice, the running time is almost linear.

On the other hand the presented associative processor allows us to achieve constant running time for basic operations (see Table I). As a result, the cost of performing M consecutive operations is $\Theta(M)$. Basically, the running time depends solely on the number of processed M operations. Thus, the most important characteristic is the command processing time which depends on actual hardware implementation.

The running time to process single operation, or the time to process N k-bit words in the associative computing system is given by [6]:

$$T = k \times t \times (\frac{N}{n} + K),$$

where t - time of associative memory cycle; n - number of PE; K - factor of complexity of performing a single elementary operation (the number of sequential steps accessing memory). Thus, the time of processing T is constant and depends on the value of N/n, i.e. the number of memory cells per PE. After fixing the value of N/n (GPU threads configuration or scheme adjustment for PE on FPGA) the time of processing remains constant regardless of the total amount of memory being processed.

In the developed software simulation model based on the GPU cluster, a considerable amount of time is spent not so much on the processor but on the exchange of messages between the processor and the host program. The next stage of development is planned to move to a hardware prototype of the processor and possible modification of the architecture in order to reduce the costs associated with the data transmitting.

REFERENCES

[1] R. Sedgewick and K. Wayne, Algorithms. 4th ed. Addison-Wesley, 2011. 992 p.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, Introduction to algorithms. 3rd ed. London, England: The MIT Press, 2009. 1312 p.

[3] N. L. Verenik, A. I. Girel, Y. N. Seitkulov, M. M. Tatur and H. P. Razhkova, "Cognitive information processing based on a parallel processor," 10th International Conference on Digital Technologies (DT '2014), pp.356–360, 2014.

[4] R. E. Tarjan, Data Structures and Network Algorithms. Society for Industrial and Applied Mathematics. CBMS-NSF Regional Conference Series in Applied Mathematics (Book 44), 1987. 140 p.

[5] A. V. Aho, J. E. Hopcroft and Jeffrey D. Ullman, The Design and Analysis of Computer Algorithms. Addison-Wesley, 1974. 470 p.

[6] B. Tsilker and S. Orlov, Organizatsiya EVM i sistem: Uchebnik dlya vuzov [Organization of computers and systems: Students textbook], St. Petersburg, Piter, 2007. 668 p.

ПРЕДСТАВЛЕНИЕ СИСТЕМЫ НЕПЕРЕСЕКАЮЩИХСЯ МНОЖЕСТВ И РЕШЕНИЕ ЗАДАЧИ СВЯЗНОСТИ СРЕДСТВАМИ АССОЦИАТИВНОГО ПРОЦЕССОРА

Вереник Н. Л., Татур М. М.

Белорусский государственный университет информатики и радиоэлектроники, г. Минск, Республика Беларусь

В статье кратко рассмотрены структуры данных и алгоритмы для работы с системой непересекающихся множеств. Предлагается использовать ассоциативный процессор с оригинальной архитектурой для представления непересекающихся множеств и решения задачи связности с помощью адаптированного алгоритма быстрого поиска. Представленный подход позволяет достичь константного времени выполнения базовых операция поиска и объединения и линейной сложности для обработки произвольной последовательности операций поиска и объединения.

Matrix-represented Constraints Satisfaction Methods: Practical Aspects of Their Implementation

Zuenko A.A., Oleinik Yu.A.
Institute of Informatics and Mathematical Modelling
KSC, the Russian Academy of Sciences
Apatity, Russia
zuenko@iimm.ru
yoleynik@iimm.ru

***Abstract*—The paper proposes an original approach to solving the problem of ineffective processing of qualitative constraints of a subject domain in the framework of constraint programming technology. The approach is based on the use of specialized matrix-like structures, providing a "compressed" representation of constraints over finite domains, as well as using author's inference algorithms on these structures. Compared to the prototypes using the typical representation of multi-place relations in a form of tables, the techniques make it possible to more efficiently reduce the search space. The paper presents practical aspects of implementation of user-developed types of constraints and corresponding algorithms-propagators with the help of constraint programming libraries. The algorithms performance has been assessed to clearly demonstrate the advantages of representation and processing of qualitative constraints of a subject domain by means of the above matrix structures.**

***Keywords*—constraint satisfaction problem, constraint programming, constraint propogation, matrix-like representation of constraints, qualitative constraints**

I. Introduction

According to [6] the *constraint satisfaction problem* (CSP) consists of three components: X, D, C.

X – a set of variables $\{X_1, X_2, \ldots, X_n\}$.

D – a set of domains $\{D_1, D_2, \ldots, D_n\}$ where D_i is the domain of variable X_i.

C – a set of constraints $\{C_1, C_2, \ldots, C_m\}$ that specify allowable combinations of the values of variables.

Each domain D_i describes a set of the admissible values $\{v_1, \ldots, v_k\}$ for variable X_i. Each constraint is a pair $< scope, rel >$ where $scope$ – is a set of variables which participate in the constraint and rel – is the relation defininining admissible combinations of values, which the variables from $scope$ can take on.

Constraints can be presented either explicitly, i.e. by enumeration of all the admissible combinations of the values for a set of variables specified, or implicitly, i.e. as an abstract relation supporting two operations: checking if a tuple is an element of the given relation, and enumeration of all the elements of the relation. The second way, in fact, requires specifying the characteristic function of the given relation.

Each *state* in a CSP is defined by an assignment of values to some (partial assignment) or to all the variables (complete assignment):$\{X_i = v_i, X_j = v_j, \ldots\}$. The *solution* of a CSP is complete assignment which satisfies all the constraints.

As an example, consider a CSP. Let $X = \{X_1, X_2\}$. We assume that $D_1 = D_2 = \{a, b, c\}$. Let a set C consists of an only constraint, that is, $C = \{C_1\}$. Constraint C_1 describes that fact that the values X_1 and X_2 must have different values.

The given constraint can be expressed implicitly, that is:

$$C_1 =<< X_1, X_2 >, X_1 \neq X_2 > . \quad (1)$$

The same constraint can be expressed explicitly, that is:

$$C_1 =<< X_1, X_2 >, \{< a, b >, < a, c >, < b, a >, < b, c >, < c, a >, < c, b >\} > . \quad (2)$$

Note, that constraint (2) can be expressed in a more compressed way:

$$C_1 =<< X_1, X_2 >, \{a\} \times \{b, c\} \cup \cup \{b\} \times \{a, c\} \cup \{c\} \times \{a, b\} > . \quad (3)$$

There is a Table in Fig. 1a, which vividly represents expression (2). Figure 1b shows a matrix corresponding to expression (3). In fact, in case of a matrix representation (Fig. 1b), the sign

X_1	X_2
a	b
a	c
b	a
b	c
c	a
c	b

a)

X_1	X_2
{a}	{b, c}
{b}	{a,c}
{c}	{a,b}

b)

Figure 1. The tabular constraint representation (a); the constraint representation in the form of specialized matrix (b).

of operation $\times$ (Cartesian product) between the components of one row is omitted, and the sign of operation $\cup$ between rows (union of sets) is not written explicitly. In [4], the similar representation is referred to as an "compressed" representation of the relation.

Unlike the article mentioned above, this paper concerns two types of matrix structures to represent constraints: C-systems and D-systems. In [3], the set-theoretical operations with the given structures are introduced. The similar structures are also used in [8] to solve pattern recognition and knowledge base compression problems.

Expression (3) can be represented in the form of the C-system:

$$C_1[X_1X_2] = \begin{bmatrix} \{a\} & \{b,c\} \\ \{b\} & \{a,c\} \\ \{c\} & \{a,b\} \end{bmatrix}. \quad (4)$$

D-systems allow calculating the complement of the C-systems: a complement is taken for each component-set.

Let's assume that we have a constraint $C_1[X_1X_2]$ meaning that $X_1 \neq X_2$. It is necessary to express the constraint $X_1 = X_2$. Then it is possible to represent D-system as follows:

$$\overline{C_1}[X_1X_2] = \left] \begin{matrix} \{b,c\} & \{a\} \\ \{a,c\} & \{b\} \\ \{a,b\} & \{c\} \end{matrix} \right[. \quad (5)$$

The D-system representation is equivalent to the expression:

$$\overline{C_1} =<< X_1, X_2 >, \{[(D_1\backslash\{a\}) \times D_2 \cup D_1 \times \times (D_2\backslash\{b,c\})] \cap [(D_1\backslash\{b\}) \times D_2 \cup D_1 \times (D_2\backslash\{a,c\})] \cap \cap [(D_1\backslash\{c\}) \times D_2 \cup D_1 \times (D_2\backslash\{a,b\})] > . \quad (6)$$

or

$$\overline{C_1} =<< X_1, X_2 >, \{[\{b,c\} \times D_2 \cup D_1 \times \{a\}] \cap \cap [\{a,c\} \times D_2 \cup D_1 \times \{b\})] \cap \cap [\{a,b\} \times D_2 \cup D_1 \times \{c\}] > . \quad (7)$$

The D-system represented can also be expressed as an intersection of three C-systems of the same scheme $< X_1, X_2 >$ namely $\overline{C_1} = K_1[X_1X_2] \cap K_2[X_1X_2] \cap K_3[X_1X_2]$:

$$\begin{bmatrix} \{b,c\} & * \\ * & \{a\} \end{bmatrix} \cap \begin{bmatrix} \{a,c\} & * \\ * & \{b\} \end{bmatrix} \cap \begin{bmatrix} \{a,b\} & * \\ * & \{c\} \end{bmatrix}. \quad (8)$$

In specifying C-systems, a designation "$*$" (a complete component) may be used, which is equivalent to the indication of domain of the corresponding variable.

There is one more type of dummy components – a empty component (designated as "$\varnothing$"), that is a component containing no value.

Now we shall try to answer the question: "When is the representation and handling of a CSP in a kind of C- and D-systems capable to ensure the highest computing performance". In other words: "In what cases should CSPs be represented and processed as the C- and D-systems?"

Most constraint programming environments are mainly oriented on processing of numerical constraints, which are specified by means of the base set of arithmetical operations, binary relations equal/unequal, more/less, built-in functions, etc. for which specialized procedures-propagators are developed. This lack of balance between tools used in quantitative and qualitative constraints processing can be seen even at a level of languages used to define constraints, which are used by different programming libraries. Thus, for instance, in the Choco library [2], it is impossible to directly specify the constraint $x = "a"$, without having substituted a symbol $"a"$ for a number.

The authors' studies showed that processing of qualitative constraints represented in the form of logical expressions and rules, is not sufficiently effective in the systems like those mentioned above, and cannot be implemented for comprehensible time even at a rather small dimension of problem.

The paper presents practical aspects of implementation of user-developed types of constraints and corresponding algorithms-propagators with the help of specialized constraint programming libraries (the Choco library taken as an example). The algorithms performance has been assessed to clearly demonstrate the advantages of representation and processing of qualitative constraints of a subject domain by means of the above matrix structures.

II. Defining CSPs and algorithms of their solution by means of modern tools of constraint programming (the Choco library taken as an example)

The main techniques used in the solution of CSPs can be classified into three classes [7]. Class 1 includes different variants of backtracking search algorithms, which construct a solution by means of extension of partial instantiation step by step, using various heuristics and applying intelligent strategies to recover from the dead ends of a search tree. Class 2 includes algorithms of constraints propagation which eliminate some non-solution elements from the search space, decreasing the dimension of problem. The algorithms themselves do not form the solution because they eliminate not all the non-solution elements. They are used either to pre-process the problem before another type of algorithm is applied, or interwoven with steps of another kind of algorithms to boost its performance. Finally, the structure-driven algorithms exploit the structure of the primal or dual graph of the problem. There are very different algorithms in this class, including ones which decompose the initial CSP into loosely-coupled subproblems, which can be solved by methods from the previous two classes. Hence, structure-based methods can be also coupled with some other types of algorithms.

Constraints programming environments allow usage of built-in types of constraints and the algorithms of their satisfaction and make it possible to develop original types of constraints, methods of their propagation, as well as construct original search strategies.

The following libraries are most widely used in constraints programming: Choco and JaCoP for Java, as well as GeCode and Z3 for C++.

To implement the original algorithms of inference on constraints that are presented as specialized matrix structures, choice was made of the Choco library. Choco library is the open source software created to define and solve constraint satisfaction problems [5].

To describe CSPs by means of the Choco library, the following basic abstractions are used:

- Model;
- Variable;
- Constraint;
- Propagator;
- Search Strategy;
- Solver;
- Solution.

A. Model.

Abstraction "*Model*" is presented by a special class of the Choco library, on the basis of which all the further defining of the CSP is formed.

B. Variable.

In the library Choco, variables are represented by specialized classes, depending on their type. There are determined four different types of variables in the library:

- **Boolean variable** is a variable with two possible values: true/false (0/1).
- **Integer variable** is a variable taking values from a set of integers. The domain of an integer variable can be specified as an interval $[a, b]$ (bounded domain). Such a representation consumes a small amount of memory, but does not allow processing the gaps in domains. In other words, it is impossible to eliminate an inadmissible value being inside the interval. Another way to specify domains of integer variables is to explicitly enumerate all possible values of a variable (enumerated domain). In so doing, the values should be linearly ordered.
- **Set variable** is a variable whose values are sets of integers. The variable of this type is specified by two sets, i.e. by the upper and lower boundary. It can take the values which are the subsets of the upper boundary and necessarily includes the lower boundary.
- **Real variable** is a variable taking values from the specified interval with the specified precision. At present this type of variables is supported rather poorly in the library.

The variables are added into the defining of a CSP either by means of the methods of class "*Model*", which associates the variables with the corresponding model, or by means of classes implementing the interfaces of creating the variables of corresponding types. The interfaces associate the variables with the model specified in the class constructor. Thus, variables cannot exist by themselves and should be necessarily associated with the model.

C. Constraint.

A certain logical formula which specifies admissible combinations of values of variables, is referred as a constraint. In the Choco library, a constraint is defined by a set of variables and by propagators (filtering algorithms) which delete the values from the specified variables domains, which do not correspond to the legal assignments. The library contains a set of built-in constraints, for example, global ones, *Alldifferent* in particular, i.e. a constraint meaning that the values of variables in a solution, should differ. Also presented in it are various arithmetical constraints; constraints presented as logical formulae; as rules, etc. There are standard methods-propagators for each built-in type of constraints.

An example of an arithmetical constraint:
model.arithm(var,"=",5).

This constraint means that variable var should take the value equal to 5.

The methods of class "*Model*" allow constructing more complicated constraints.

For example:
model.or(new Constraint[]{model.arithm(var,"=",5), model.arithm(var,"=",6)}).

This constraint means that the variable var takes values 5 or 6, and so it takes, as a parameter, an array of constraints between which an OR-operation is set.

The library also allows constructing original constraints and propagators. To create original program class of constraint, it is necessary to specify propagator (one or several) and set of variables, over which the constraint is set.

An example of user-developed constraint creation:
Constraint c=new Constraint("My",new Dpropogator(vars));
"My" – string-name of the constraint;
new Dpropogator(vars) – a propagator for a constraint over variables vars (there is a possibility to specify several of these).

For the constraint to be taken into account in solution, it is necessary to call a method "*post*()" after the constraint has been created, otherwise, it will not be taken into account in solution.

```
Constraint C1 = model.arithm(var,"=",5);
Constraint C2 = model.arithm(var,"=",6);
Constraint C3 = model.or(new Constraint[]{C1,C2});
```

Figure 2. Example of constraint creation by means of Choco library.

There are three constraints created on Fig. 2 but the method "*post*()" has been called in no one yet. If *C*1.*post*() is to be called, there will be only value "5" for variable *var* in the solutions. Thus, if it is necessary to add constrain "*var* = 5 or *var* = 6" into the solution, the method *post*() should be called only for constraint *C*3.

D. Propagator.

Abstraction "*Propagator*" is specified in the Choco library as class. Constructing class of propagator, it is necessary to define two main methods: a method *propagate* and a method *isEntailed*. The method *propagate* implements the logic of the constraint propagation. During the propagation, the method will be called repeatedly till it causes changes in variable domains. The method *isEntailed* is called at the end of propagation. The method *isEntailed* can return three parameters:

- *ESat.TRUE* – the constraint is completely satisfied;
- *ESat.UNDEFINED* – the constraint status failed to be determined (the propogation ends but the constraint has not been satisfied);
- *ESat.FALSE* – the constraint cannot be satisfied, backtrack is necessary.

Also, in the course of propagation there may appear contradiction exception. In this case, the propagation is considered to be completed ahead of the schedule, and the propagation result is considered to be equivalent to *ESat.FALSE*.

E. Search strategies.

If the final solution has not been reached with the help of propagators, the search space is further studied in accordance with a certain search strategy. In fact, the search strategy defines the way the CSP solution should be constructed.

The Choco Version 4.0.0 constructs a binary search tree (for example, if assignment "$x = 5$" cannot be extended to a solution, then "$x \neq 5$" is considered). The search strategy like this, which implements backtracking search, is typical for CSPs. However, other search strategies shouldn't be neglected.

The strategy correctly selected for a certain problem allows the solution to be generated much faster. The types of strategies are related to the types of variables, i.e. each type is supplied with a specific set of strategies. Like in case with constraints, the Choco library provides built-in program classes for popular strategies. However, if necessary, user-developed search strategy may be constructed. To describe a user-developed strategy, it is necessary to additionally define two (three, optionally) classes: a strategy to select a particular variable from a set of the CSP variables (Class 1), a strategy to select a particular value of the specified variable from its domain (Class 2), and, optionally, a class implementing the strategy to select the branch of a search tree (Class3).

Let's characterize Class 3 separately. It is necessary to define two basic methods here. The first one, *apply*(*IntVar*, *int*), is the method for processing a pair <*a variable*, *its value*> obtained as a result of the variable selection strategy application and the variable value selection strategy application. The second, *unapply*(*IntVar*,*int*), is the method specifying how modifications introduced by the method *apply* should be droped. If the search-tree branch selection strategy is not determined, the method *apply* assigns the selected value to the specified variable, trying to extend the partial assignment to a complete one. Having worked out this variant, the method *unapply* eliminates the considered value from the domain of the corresponding variable. The search proceeds in alternative directions.

F. Solver.

"*Solver*" is an abstraction presented by the class, which keeps the stages of solution process, search strategies and the CSP solutions if those have been reached. The Solver type object is the field of class "*Model*". After the necessary solver parameters have been specified, its method *solve*() is called to start searching.

G. Solution.

A "*Solution*" is an abstraction presented by the Choco library class, which serves as a storage of a complete or partial assignment.

III. The original approach to implementation of inference procedures

Based on the Choco library, the original classes have been developed, extending the functional of the basic library in order to represent and solve the CSPs in the form of a D-systems set. Unlike the standard constraints of the Choco library and their propagators working with assignments of variables, the propagator for the D-systems works with the system components containing several values. In the process of inference, the D-system is reduced and the amount of the information processed decreases with each iteration.

Due to the requirements concerning the length of the paper, no detailed consideration is given to the methods of inference. Given in [9], [10] are the particular techniques to be used to solve the CSPs. These techniques are based on matrix representation of constraints over finite domains.

To explain the approach based on the similar inference, we shall consider that the CSP constraints may be represented in the form of a D-systems set. In the practical problems, it is a set of the C- and D-systems, numerical constraints, as well as that of global constraints [1], [7].

So, let's consider the affirmations which allow implementation of the equivalent CSP transformations for the case under consideration (constraints propagation). The aim of transformations is to reduce a CSP to a simpler form, with the less number of D-systems, the less number of rows of D-system, columns (attributes) of D-system, values in the attributes domains, values in separate components, etc.

- **Affirmation 1**. If, at least, one tuple (row) of the D-system is empty (all components of the tuple are empty), the D-system is empty (the corresponding system of constraints is inconsistent, the CSP has no solution).
- **Affirmation 2**. If all the components of an attribute are empty, the attribute can be eliminated from the D-system (all the components in the corresponding column are removed) and the pair "the eliminated attribute - its domain" is included into the partial solution.
- **Affirmation 3**. If in the D-system there is a tuple (row) containing only one nonempty component, all the values not included into this component, are deleted from the corresponding domain.
- **Affirmation 4**. If a tuple of the D-system contains, at least, one complete component, this tuple is removed (one can remove the corresponding constraint from the system of constraints).
- **Affirmation 5**. If the component of an attribute of the D-system contains the value not belonging to the corresponding domain, this value is deleted from the component.

IV. Comparison between the algorithms of Choco library and the procedures developed

To determine the effectiveness of the classes and algorithms developed we used the problem of placing n chess queens on an $n \times n$ chessboard so that no two queens threaten each other ("N-Queens problem"), due to the simplicity of its scaling.

The aim of the given example was not to generate solutions of the "N-Queens problem" (one or all possible) as fast as possible. Moreover, the authors realized that, by means of standard numerical constraints, it is possible to define the given problem more implicitly. Taking into account of the chessboard symmetry and constraints propagation techniques based on the interval analysis, allow the solutions to be generated faster than in the procedure described in the study presented. The aim of the analysis made is to demonstrate that qualitative dependencies processing by means of modern constraint programming environments, by the Choco library in particular, is less effective than that by the matrix representation in this paper.

That is why the "N-Queens problem" will be described in the form of a set of qualitative constraints and a comparison will be made between the propagation algorithms performance for the two cases. In the first case, the problem is formulated in the form of logical expressions in the Choco language and standard classes-propagators are used. In the second case, the constraints are formalized in the form of a set of the matrix-like structures suggested. Original constraint propagation algorithms are proposed, on the basis of which own classes-propagators extending the base functional of the Choco library are developed.

Thus, the algorithms compared differ in the qualitative constraints representation and the propagators used. After the propagation has stopped, the search tree branching strategy, typical for Choco, is applied.

Example (“N-Queens” problem). Consider a simplified variant of “N-Queens” problem. In this example, the chessboard size is $n \times n$. It is necessary to find possible variants of four queens placing.

Let’s associate the i-th horizontal with variable X_i. Then each variable (attribute) will be defined in the domain as $\{a, b, c, d\}$, where a, b, c, d are the labels of the verticals. As an example, let’s formulate the constraint “two queens placed on horizontals 1 and 2 are not threatened by each other” in the form of the D-system:

$$\begin{array}{c} \\ \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{cc} X_1 & X_2 \\ \{a,b,c,d\} & \{a,b,c,d\} \\ \left]\begin{array}{c} \{b,c,d\} \\ \{a,c,d\} \\ \{a,b,d\} \\ \{a,b,c\} \end{array}\right. & \left.\begin{array}{c} \{c,d\} \\ \{d\} \\ \{a\} \\ \{a,b\} \end{array}\right[\end{array} \tag{9}$$

In particular, the first row of the given D-system shows that if a queen is on the field a_1(the intersection of the first horizontal and the first vertical), then, in the second horizontal, other queen can occupy fields c_2 and d_2 only. In the language of logic, it is expressed as follows:

$$\begin{array}{l} (x_1 = a) \rightarrow ((x_2 = c) \vee (x_2 = d)) \\ \text{or } \neg(x_1 = a) \vee (x_2 = c) \vee (x_2 = d) \\ \text{or } \neg(x_1 = b, c, d) \vee (x_2 = c) \vee (x_2 = d). \end{array} \tag{10}$$

Comparing different pairs of horizontals, it is possible to write out all the constraints on the inter-relative positioning of all the 4 queens. In our case, the number of constraints is calculated by the formula: 4 C_4^2 (for each pair of horizontals, four constraints are formed, the total number of pairs is – C_4^2), that is $\frac{4^2 \cdot (4-1)}{2}$. For a board of $n \times n$ in dimension, the number of constraints is calculated by the formula: $\frac{n^2 \cdot (n-1)}{2}$. So, the CSP considered can be expressed by a D-systems set (like the D-system shown above) describing to the admissible positions of pairs of queens relative to each other.

In the Choco library, the D-system can be represented as a set of built-in logical constraints of the library. For example, the D-system

$$\begin{array}{c} \\ \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{cc} X_1 & X_2 \\ \{b,c,d\} & \{a,b,c,d\} \\ \left]\begin{array}{c} \{c,d\} \\ \{b,d\} \\ \{b,c\} \end{array}\right. & \left.\begin{array}{c} \{d\} \\ \{a\} \\ \{a,b\} \end{array}\right[\end{array} \tag{11}$$

can be described, using the Choco library by encoding the values of variables by integers (Fig. 3).

```
Model testmodel=new Model ("Test");
int []x = {2,3,4},y = {1,2,3,4} ;
IntVar X = testmodel.intVar("X",x);
IntVar Y = testmodel.intVar("Y",y);
Constraint node11=testmodel.or(testmodel.arithm (X,"=",3),
                                testmodel.arithm(X,"=",4));
Constraint node12=testmodel.arithm(Y,"=",4);
Constraint node21=testmodel.or(testmodel.arithm(X,"=",2),
                                testmodel.arithm(X,"=",4));
Constraint node22=testmodel.arithm (Y,"=",1);
Constraint node31=testmodel.or(testmodel.arithm(X,"=",2),
                                testmodel.arithm(X,"=",3));
Constraint node32=testmodel.or(testmodel.arithm (Y,"=",1),
                                testmodel.arithm(Y,"=",2));
Constraint firstrow=testmodel.or(node11, node12);
Constraint secondrow=testmodel.or(node11, node12);
Constraint thirdrow=testmodel.or(node31, node32);
testmodel.and(firstrow,secondrow,thirdrow).post();
```

Figure 3. D-system described by the means of Choco library.

It is clear from Fig. 3 that in order to describe even such a small D-system, a significant number of constraints is required: one constraint per each value of the component (cell) of the D-system, one constraint per each component, constraints uniting components of a row, and one common constraint uniting rows of a system. As the system increases in size in such a representation, the number of constraint also increases essentially, which will take much more time to reach the solution. That is why an original representation of the D-system has been developed to specify it by only one constraint.

Presented below are comparative plots of time required for solutions of "N-Queens problems" for both representations specified. The units of measure for a vertical scale of plots are milliseconds, i.e. the time to solve the problem. However, the time assigned to the solution, has been limited to 2 minutes (120000 msec). Plots in Fig. 4 show the time spent for reaching the first solution of the "N-Queens problem". The plots are broken when the testing computer memory is exhausted. The plots in Fig. 5 show the time spent for reaching all the solutions of the "N-Queens problem". These are broken in the points in which the program was not able to reach all the solutions at the time assigned.

Figure 4. Reaching the first solution of the "N-Queens problem" (ms/N).

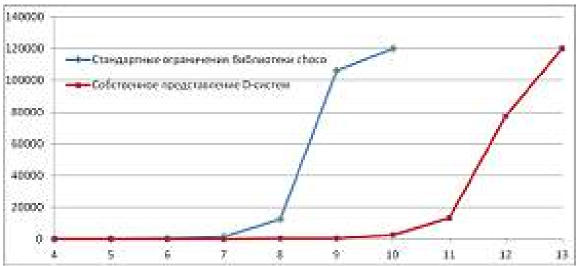


Figure 5. Reaching all the solutions of the "N-Queens problem" (ms/N).

V. Conclusion

The studies have demonstrated that the matrix constraint representation proposed by the authors, as well as the original methods of their propagation are quite suitable for practical application. Moreover, in case of qualitative constraint modelling, the application of the approach proposed gives an essential gain in time against the algorithms of qualitative constraint propagation, which are built in the Choco library. In particular, solving the "N-Queens problem", when it was required to reach all the solutions under the time limit of two minutes, the approach proposed permitted processing the search space of 12^{12} in dimension. In doing so, the standard tools of the Choco library did not give a possibility to study the search space of greater than 9^9 in dimension. When the task was to achieve even one solution for the same two minutes, the time for the standard tools of the Choco library to reach a solution was not enough even when the search space was of 19^{19} in dimension. The methods proposed allowed the authors to reach a solution even for the search space of 76^{76} in dimension.

The authors would like to thank the Russian Foundation for Basic Research (grants No 16-07-00562-a, No 16-07-00377-a, No 16-07-00313-a, No 16-07-00273-a, No 18-07-00615-a) for help in partial of this research.

References

[1] Bartak, R. Constraint Programming: In Pursuit of the Holy Grail. *Proceedings of the Week of Doctoral Students (WDS99)*, Part IV, 1999, pp. 555–564.

[2] Jussien N., Rochart G., Lorca X. Choco: an Open Source Java Constraint Programming Library. *CPAIOR'08 Workshop on Open-Source Software for Integer and Contraint Programming (OSSICP'08)*, 2008, pp. 1-10.

[3] Kulik B.A. A Logic Programming System Based on Cortege Algebra. *Journal of Computer and Systems Sciences International*, 1995, vol. 33, No. 2, pp. 159-170.

[4] More T. Axioms and theorems for a theory of arrays. *IBM Journal of Research and Development*, 1973, No. 17(2), pp. 135–175.

[5] Prud'homme Ch., Fages J-G, Lorca X. Choco Solver Documentation Release 4.0.0. Sep 13, 2016, 37 p.

[6] Russel S., Norvig P. *Artificial Intelligence: A Modern Approach. 3rd edition*. Prentice Hall, 2010, 1132 p.

[7] Ruttkay Zs. Constraint satisfaction a survey. *CWI Quarterly*, 1998, vol. 11, pp. 163–214.

[8] Zakrevskij. A. Integrated Model of Inductive-Deductive Inference Based on Finite Predicates and Implicative Regularities. *Diagnostic Test Approaches to Machine Learning and Commonsense Reasoning Systems IGI Global*, pp. 1-12.

[9] Zuenko A.A. Vyvod na ogranicheniyakh s primeneniem matrichnogo predstavleniya konechnykh predikatov [Constraint inference based on the matrix representation of finite predicates]. *Iskusstvennyi intellekt i prinyatie reshenii* [*Artificial Intelligence and Decision Making*], 2014, No. 3, pp. 21-31.

[10] Zuenko A.A., Lomov P.A., Oleinik A.G. Primenenie metodov rasprostraneniya ogranichenii dlya uskoreniya obrabotki zaprosov k ontologiyam [Application of constraint propagation techniques to speed up processing of queries to ontologies]. *Trudy SPIIRAN* [*SPIIRAS Proceedings*], 2017, No. 1(50), pp. 112-136.

МЕТОДЫ УДОВЛЕТВОРЕНИЯ ОГРАНИЧЕНИЙ, ПРЕДСТАВЛЕННЫХ В МАТРИЧНОЙ ФОРМЕ: ПРАКТИЧЕСКИЕ АСПЕКТЫ ИХ РЕАЛИЗАЦИИ

Зуенко А.А., Олейник Ю.А.

В работе предлагается оригинальный подход к решению проблемы недостаточной эффективности обработки качественных ограничений предметной области в рамках технологии программирования в ограничениях. Подход основан на применении специализированных матрицеподобных структур, обеспечивающих “сжатое” представление ограничений над конечными доменами, а также авторских алгоритмов вывода на данных структурах. По сравнению с прототипами, использующими стандартное представление многоместных отношений в виде таблиц, разработанные методы позволяют более эффективно сокращать пространство поиска. В работе представлены практические аспекты создания пользовательских типов ограничений и алгоритмов их распространения с помощью библиотек программирования в ограничениях. Также было проведено сравнение быстродействия различных алгоритмов, наглядно демонстрирующее преимущества использования описанных матрицеподобных структур для представления и обработки качественных ограничений.

Ontological Resources for Representing Security Domain in Information-Analytical System*

Natalia Loukachevitch
Research Computing Center
Lomonosov Moscow State University
Moscow, Russia
louk_nat@mail.ru

Boris Dobrov
Research Computing Center
Lomonosov Moscow State University
Moscow, Russia
dobrov_bv@mail.ru

***Abstract*—The paper presents the approach to the description of the broad domain of national security as a thesaurus for automatic document processing. The created Security thesaurus has the representation model of the RuThes thesaurus.The Security thesaurus includes terminology related to social, national and religious conflicts, extremism and terrorism, information security. It is used in a specialized information-analytical system and for automatic text categorization according to several categorization schemes. The information-retrieval system provides several search instruments including word, phrase and concept search, category and facet search. It also supports the creation of analytical reports.**

***Keywords*—Security domain, thesaurus, text categorization, automatic document processing**

I. Introduction

National security is the important concept for describing the social phenomenon of the protection of the vital interests of the individual, society and the state against dangers and threats. The National Security Strategy of the Russian Federation ensures the implementation of constitutional rights and freedoms of citizens of the Russian Federation, decent quality and standard of living, sovereignty, independence, state and territorial integrity, sustainable social and economic development of the Russian Federation. National security includes the defense of the country and all types of security provided for by the Constitution of the Russian Federation and the legislation of the Russian Federation, primarily state, public, information, environmental, economic, transport, energy and personal security [22].

In accordance with this, the priorities and prospects of scientific and technological development include the counteraction to technogenic, biogenic, sociocultural threats, terrorism and ideological extremism, as well as cyber threats and other sources of danger for society, economy and the state.

One of important sources of information for forming counteraction against existing threats is the analysis of textual data including news reports, analytical papers, messages in social networks. In this paper we consider the approach to description of the broad domain of national security as a thesaurus for automatic document processing. The created Security thesaurus has the representation model of RuThes thesaurus [9]. We use the Security thesaurus in a specialized information-analytical system and for automatic text categorization of documents according to several categorization schemes, including Threats categories, Values categories, Regional problem categories and others. The information-retrieval system provides several search instruments including word, phrase and concept search, category and facet search. It also supports the creation of analytical reports.

*The work is supported by the Russian Foundation for Basic Research (project 16-29-09606) and the Ministry of Education and Science of the Russian Federation (project N 14.601.21.0018)

II. Related Work

In literature, several directions of natural language processing for national and international security issues have been studied.

Many works are devoted to the analysis of extremist messages in social networks. The part of such research is devoted to the activities of ISIL in Twitter. It is known that this organization actively works with a variety of social media [16], [20]. In particular, ISIL and associated organizations support a large number of Twitter accounts in order to spread their ideas in several languages. In the work [6], an analysis of the posts of people who joined or attempted to join ISIL is given. In most cases, they expressed strong anti-American and anti-Western attitudes long before joining ISIL.

The paper [2] describes a corpus containing 100 texts written by Islamists. In particular, fragments concerning war, non-believers, and punishments (64 texts) have been extracted from the collections of hadith (Islamic religious texts). Also, the corpus contains messages from Islamist blogs, as well as articles from the Islamist magazine Inspire. These texts have been labeled at several levels, including syntactic, time, and referential annotations.

Since one of the prerequisites of supporting terrorism is a sharply negative attitude toward certain phenomena or groups of people, so-called "hate" statements require special attention. In the work [15], blog posts are classified not only into classes of expressed emotions (positive, negative, neutral), but also according to the type of action discussed (negative - *attack, attack, bloodshed, cruelty*, or positive - *help, support*). Kwok and Wang [7] describe the collected balanced set of 24.5 thousand tweets, which are classified into racist and normal ones. Waseem and Hovy [19] created a corpus of 16,000 tweets, in which 3.3 thousand tweets are labeled as sexist, and 1.9 thousand as racist.

In work [13], it is pointed out that the automatic detection of statements that incite hatred towards some groups of the population (which is one of the signs of extremism) is complicated by the following factors: such a message cannot always be detected by a simple set of keywords, because some words are intentionally distorted (to avoid detection); any fixed lists of offensive words constantly require adding; hate statements can be written in a completely fine literary language; offense and hatred may cross the boundaries of the sentence, when the object of the sentence is in another sentence; use of sarcasm. Schmidt and Wiegand [17] provide an overview of current approaches to detecting hate messages. It is pointed out that such approaches are based on the application of supervised machine learning methods based on several groups of features.

In the sphere of natural language processing for information security, Lim et al. [8] discuss the construction of a database for annotated malware texts. The annotation framework is based on the MAEC vocabulary for defining malware characteristics [5], along with a database consisting of annotated APT reports. The authors plan to use the database to construct models that can potentially help cybersecurity researchers in their data collection and analytics efforts. Gorokhov et al. [3] study application of convolutional neural network for anomaly detection in e-mail data.

Several projects are devoted to global monitoring of events occurring in the world to understand and respond to global problems, for example, International Crisis Early Warning System (ICEWS), maintained by Lockheed Martin, and Global Data on Events Language and Tone (GDELT) [14], [18]. Hand-coded and automated event data have been used to anticipate conflict escalation [14]. When combined with statistical and agent-based models, ICEWS claims a forecasting accuracy of 80%. GDELT has been used to track, e.g., wildlife crime and the rise of hate speech following the U.K. Brexit vote [18].

III. RuThes Model of Representation

We use the model of the RuThes thesaurus for knowledge representation in the security domain. The structure of RuThes is based on three traditions of developing computer resources for document procssing: information-retrieval thesauri, WordNet-like thesauri, and formal ontologies [11]. The RuThes thesaurus is created in form of a linguistic ontology, which concepts are based on senses of really existing words and phrases. There exist several large Russian thesauri presented in the same format:

- RuThes thesaurus comprising words and phrases of literary Russian together with terms of so-called sociopolitical domain (see below) [9];
- RuThes-lite, a published version of RuThes[1], can be obtained for non-commercial purposes [10];
- Sociopolitical Thesaurus comprising lexical items and terms from the sociopolitical domain. The sociopolitical domain is a broad domain describing everyday life of modern society and uniting many professionals domains, such as politics, law, economy, international relations, finances, military affairs, arts and others. Terms of this domain are usually known not only professional, but also ordinary people [11]. The Sociopolitical thesaurus can exist and be used separately. At the same time it is included as a part into three larger thesauri: RuThes, OENT ontology, and the Security Thesaurus;
- Ontology on Natural Sciences and Technology (OENT) includes terms of mathematics, physics, chemistry, geology, astronomy etc., terms of technological domains (oil and gas, power stations, cosmic technologies, aircrafts, etc.). It also contains the Sociopolitical thesaurus as a part because scientific and technological problems can be discussed together with political, economical, industrial and other issues [1];
- Security Thesaurus is an extension of the RuThes thesaurus and includes terminology related to social, national and religious conflicts, extremism and terrorism, information security.

The Table 1 contains quantitative characteristics of the abovementioned resources.

Table I
RuThes-like Thesauri

Thesaurus	Number of concepts	Number of Text Entries	Number of Conc. Relations
RuThes	55,275	170,130	226,743
RuThes-lite	31,540	111,559	128,866
Sociopolitical Thesaurus	41,426	121,292	161,523
OENT	94,103	262,955	376,223
Security Thesaurus	66,810	236,321	271,297

The Security thesaurus is created in form of a linguistic ontology, which concepts are based on senses of really existing words and phrases. Each concept has a unique name and is associated with a set of text entries, the senses of which correspond to the concept. Text entries of a specific concept can comprise single words of different parts of speech, including ambiguous ones, and phrases that can be either idiomatic or compositional groups. Large rows of synonyms and term variants are collected to provide better recognition of concepts in texts.

Fig. 1-3 show the interface of thesaurus developing. The upper left form contains list of concepts, the lower left form shows text entries for the highlighted concept. The right upper

[1] http://www.labinform.ru/pub/ruthes/index.htm

form presents the relations of the highlighted concept, and the lower right form shows text entries for a related concept. It can be seen (Fig. 1-3) that concepts are provided with numerous text variants extracted from real texts, for example, *import dependence* concept (Fig.1) can be expressed also as *dependence on import* or *dependence on import goods*. Fig. 3 shows different text variants to express *zero-day vulnerability* and *vulnerability attack* concepts.

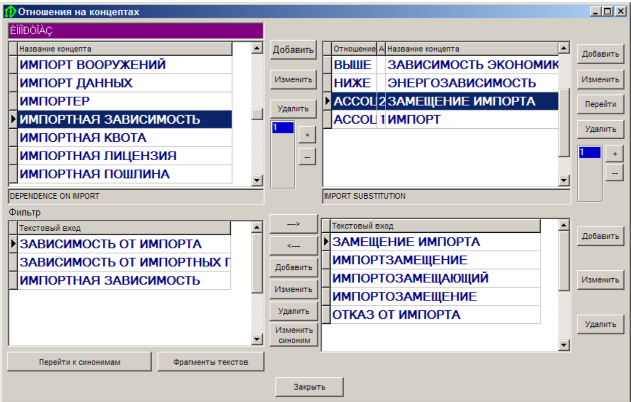


Figure 1. Example of economic threat description.

There are four basic types of relationships between concepts. The first type of relation is the class-subclass relationship, has the properties of transitivity and inheritance.

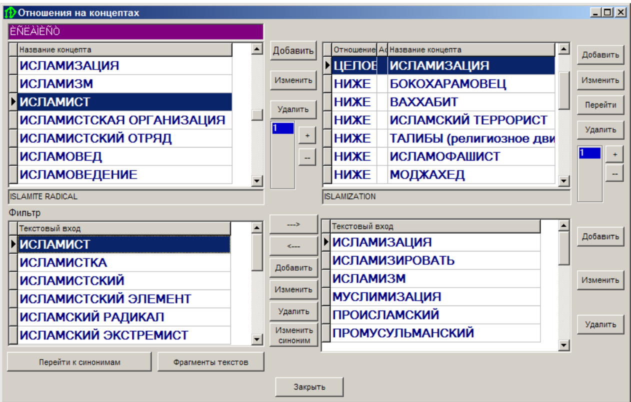


Figure 2. Example of state threat description.

The second type of relations is the part-whole relation. It is used not only to describe the physical parts, but also to other internal attributes, such as properties or roles for situations. An important condition for establishing this relationship is that the concept-parts must be rigidly connected with their whole, that is, each example of the concept-part must, throughout its entire existence, be part of the concept of the whole, and not relate to anything else. Under these conditions, it is possible to rely on the transitivity property of the part-whole relation, which is very important for automatic logical inference in the process of automatic text processing [11].

For example, concept *halal* (sanctioned in Islam) is described as a part of concept *Shariat*, and concept *Shariat* is presented as a part *Islam* concept. *Islamist radical* is presented as a part of *Islamization* (Fig. 2).

The third type of relations, called the asymmetric association $asc_1 - asc_2$, connects two concepts that cannot be related by the relationships discussed above, but when one of which does not exist without the existence of another, for example, the *import dependence* concept can exist if only the *import* concept exists (Fig.1). The *vulnerability attack* concept can appear if only the *computer vulnerability* concept exists (Fig.3)

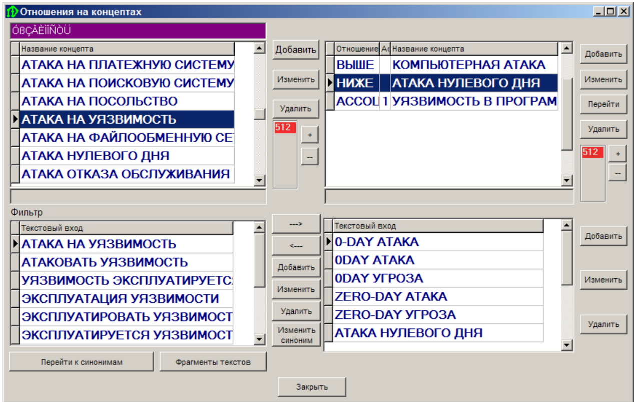


Figure 3. Example of information threat description.

The last type of relationships is the symmetrical association links concepts that are very similar in meaning, but which seems difficult to combine in one concept.

Thus, the system of the thesaurus relations describes the most significant relationship of concepts. Fig 4. presents the hierarchy of concepts under the *threat* concept. There can be seen such threats as military confrontation, cyber threats, radiation threats (nuclear war, radioactive contamination, nuclear terrorism), terroristic threat, etc.

IV. Thesaurus-Based Text Categorization

The mainstream technology in automatic text categorization is the machine learning approach. This approach presupposes that consistent training data of sufficient volume is available for training algorithms. However, in a new complicated text categorization task, even a system of categories can be absent and should be created.

In such conditions, knowledge-based methods of automatic text categorization based on manual rules of category assignment are more appropriate. Working in a broad domain, it is necessary to use thesaurus support in rule description because the thesaurus allows for operating not with single words and expressions but with concepts and thesaurus substructures [11] to describe inference of a specific category from a text.

In the security domain, the following categorization schemes have been created:

- Threat categories, describing existing national security threats (188 categories, 5-level hierarchy);

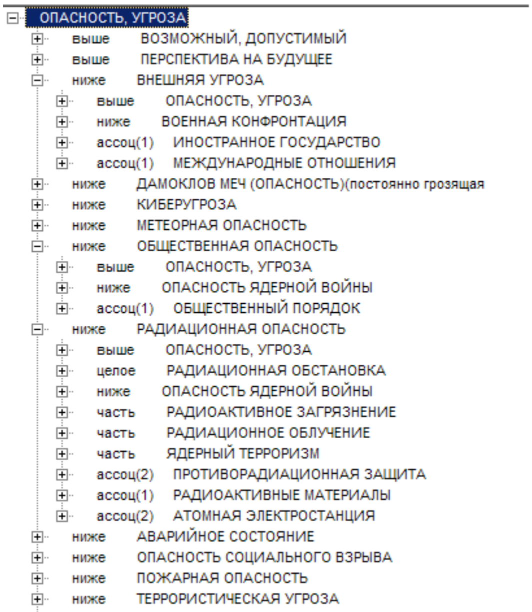


Figure 4. Hierarchy of the threat concept.

- Values categories presenting individual, social, and state values, for example, freedom, democracy, human rights, family values, etc. (109 categories, 4-level hierarchy);
- Categories of ethno-confessional relations (94 categories, 4-level hierarchy);
- Regional problems categories (84 categories, 2-level hierarchy);
- Region facets (325 facets: subjects of the Russian Federation and foreign states).

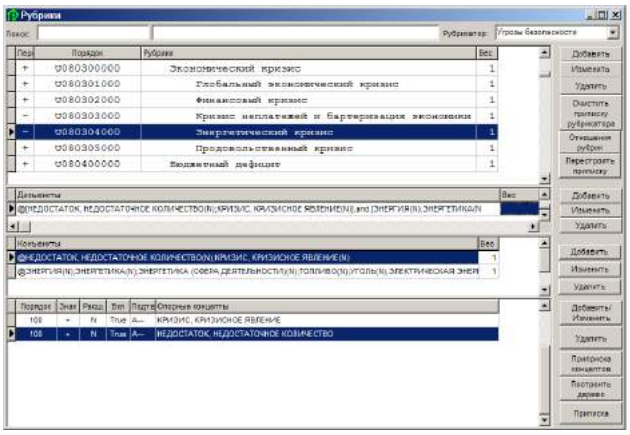


Figure 5. Example of description of a threat for text categorization system.

A. Using Thesaurus to Describe the Category Contents

In our case, for each category, a Boolean expression over thesaurus concepts is created. The syntax of the expression allows using the thesaurus hierarchies to provide appropriate coverage of the description. Here the transitivity of class-subclass and part-whole relations is significantly exploited.

In a special interface, an expert can edit the description of a category, delete extra concepts, add new concepts, or change their expansion scope. The expansion scope determines what related concepts to an initial concept should be used in the category profile:

- no expansion, only concept text entries are used (N);
- all lower concepts, including subclasses, parts and dependent concepts (Y) obtained from the concept hierarchy;
- only subclasses from the concept hierarchy (L);
- all directly related lower concepts (W);
- all directly related lower concepts without subclasses (V).

Fig. 5 presents the description of the category "Energy crisis" from the Threats subject heading scheme. This is conjunction of two disjunctions. The first one contains crisis-related concepts, including *crisis, lack, deficiency*. The second one comprises energy-related concepts.

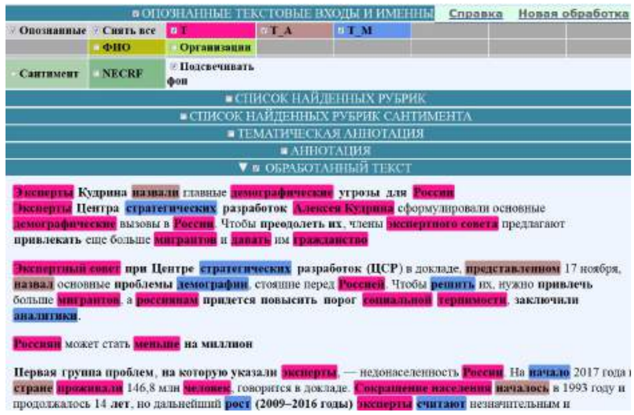


Figure 6. Thesaurus terms found in a text. Brown and blue boxes show ambiguous terms, which should be disambiguated

B. Categorization of Texts according to Subject Headings

The main stages of thesaurus-based text categorization include:

- Tokenization and lemmatization; that is, the transfer of word forms to lemmas (dictionary word forms);
- Matching with the thesaurus based on the lemma representation of the document. Fig. 6 shows the term coverage of news text " Kudrin's experts named the main demographic threats for Russia"[2]);
- Disambiguation of ambiguous text entries. Brown and blue boxes on Fig. 6 highlight ambiguous terms. For example, Russian word *demografiya (demography)* can mean *demographic situation* or *demographic science*;
- Grouping semantically related concepts into so called thematic nodes. This provides better determination of concept weights, which calculated on the basis of the concept frequency in the given document and the significance of the corresponding thematic node. Fig. 7 (in the

[2]https://www.rbc.ru/economics/17/11/2017/5a0eb1d39a79470f724250b4

center) demonstrate such thematic nodes for the above-mentioned document about the demographic threats. The important thematic node about the demographic situation (*demographic situation, life expectancy, natural population decline, age structure of the population, population aging*... etc.) ;

- Forming the conceptual index of the document. The conceptual index consists of concepts found in the document and their assigned weights. The weight of a concept accounts for the significance of the corresponding thematic node and the frequency of the concept in the document. In the example text, the important threat "population aging" was explicitly mentioned only once in the text, and it could obtain a too low frequency-based weight, but with the support of the main thematic node "demographic situation", its weight is considerably higher;
- Calculation of category weights in dependence of concepts included into the rules of the inference for this category. Fig. 6 (upper part) shows the categories found in the mentioned document including "Depopulation", "Population aging", "Fertility decline";
- The results of document processing, including the word index, the conceptual index, the calculated categories, etc. are loaded into an information-analytical system.

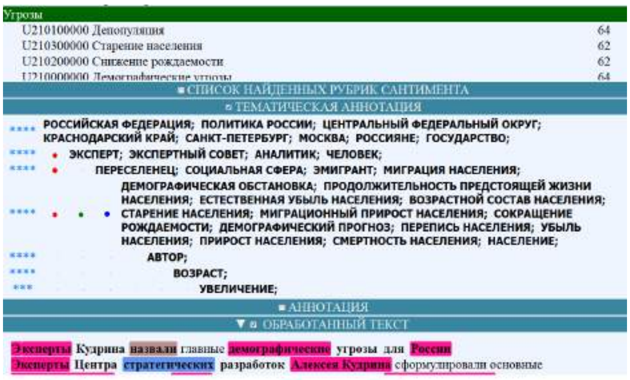


Figure 7. The upper part shows the threat categories found for a text. The central part presents thematic nodes of related concepts such as "demographic situation" node.

V. Document Searching and Categorization Schemes in Information-Analytical System

As a result of single document processing, a large number of entities of different nature (terms, named entities, sentiment expressions, etc) are extracted. The found entities are attached to word positions. The extracted entities, their weights and text positions are loaded into the search indexes of the information-analytical system.

Two types of search indexes are used: inverse and direct indexes, which are stored in special noSQL databases. The inverse index stores information organized from vocabulary elements (E_i) to documents and used for document searching:

$$[E_i[d_j, rank(E_i, d_j), [positions_{ijk}]]] \quad (1)$$

.

The direct index stores the same information, but is organized according to the documents:

$$[d_j[E_i, rank(E_i, d_j), [positions_{ijk}]]] \quad (2)$$

The direct index is used, for example, to highlight the relevant text fragments, which makes it possible to speed up the process of selecting the documents needed by a user. Also, the direct index is the main support tool for advanced search engine analysis instruments, when, for the initial query, a group of documents is received and it is required to select the most significant elements in this sample.

NoSQL database NearIdx 9.0 is used as a search engine, developed by LLC "Laboratory for Information Research" in cooperation with the Research Computing Center of Lomonosov Moscow State University. The search time for simple queries on the inverse index for a collection of 10 million documents on standard modern computers is 1-2 seconds.

The documents can be searched according to different search criteria. Also analytical reports can be formed. An analytical report usually contains document fragments related to the given topic with additional subcategorization. Similar fragments are clustered. The screen interface allows for editing, removing etc. fragments to obtain a more meaningful report. Fig. 9 presents the analytical report on "Demographic threat" category with additional regional subcategorization. It can be seen that the document fragments do not contain words *demography* and *threat*.

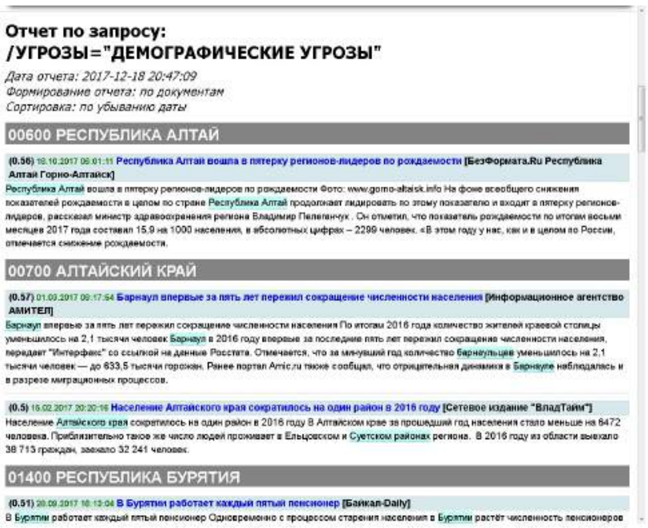


Figure 8. Analytical report created for "Demographic Threat" category

VI. Conclusion

In this paper we considered the approach to description of the broad domain of national security as a thesarus for automatic document processing. The created Security thesaurus includes terminology related to social, national and religious conflicts, extremism and terrorism, information security. The

Security thesaurus has the representation model of RuThes thesaurus. Its structure is based on three traditions of developing computer resources for document processing: information-retrieval thesauri, WordNet-like thesauri, and formal ontologies. We continue the development of the Security thesaurus.

We use the Security thesaurus in a specialized information-analytical system and for automatic text categorization of documents according to several categorization schemes, including Threats categories, Values categories, and Regional problem categories. The information-retrieval system provides several search instruments including word, phrase and concept search, category and facet search, useful for analytics. It also supports the creation of analytical reports.

References

[1] B.Dobrov and N. Loukachevitch, "Development of Linguistic Ontology on Natural Sciences and Technology", in Proc. of Linguistic Resources and Evaluation Conference LREC-2006, 2006.

[2] M. Finlayson, J. Halverson, and S. Corman, "The N2 corpus: A semantically annotated collection of Islamist extremist stories", in Proc. of Linguistic Resources and Evaluation Conference LREC-2014, pp. 896-902.

[3] O. Gorokhov O., M. Petrovskiy and I. Mashechkin, "Convolutional Neural Networks for Unsupervised Anomaly Detection in Text Data", in Int. Conf. on Intelligent Data Engineering and Automated Learning, Springer, Cham, 2017. pp. 500-507.

[4] R. Killam, P. Cook and N. Stakhanova, "Android Malware Classification through Analysis of String Literals", in First Workshop on Text Analytics for Cybersecurity and Online Safety TA-COS, 2016.

[5] I. Kirillov, D. Beck, P. Chase, R. Martin, " Malware Attribute Enumeration and Characterization", The MITRE Corporation, Tech. Rep., 2010

[6] E. Kohlmann, L. Alkhouri L,"Profiles of foreign fighters in syria and iraq", CTC Sentinel, September, 2014.

[7] I. Kwok, and Y. Wang, "Locate the hate: Detecting tweets against blacks", in Proc. of AAAI-2013, 2013.

[8] S. Lim, K. Swee, A.Muis, W., Lu, and C. Ong, "MalwareTextDB: A Database for Annotated Malware Articles", in Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1, 2017.

[9] N. Loukachevitch, and B. Dobrov, “RuThes linguistic ontology vs. Russian wordnets“, in Proc. of Global WordNet Conf. GWC-2014, 2014.

[10] N. Loukachevitch, B. Dobrov and I. Chetviorkin, ’ RuThes-Lite, a publicly available version of Thesaurus of Russian language RuThes", Computational Linguistics and Intellectual Technologies: Papers from the Annual Intern.l Conf. “Dialogue-2014”, 2014, pp. 340-349.

[11] N. Loukachevitch and B. Dobrov, "The Sociopolitical Thesaurus as a resource for automatic document processing in Russian", Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication 21.2, 2015, pp. 237-262.

[12] M.Saleem, KP. Dillon, S. Benesch, D. Ruths, "A Web of Hate: Tackling Hateful Speech in Online Social Spaces", in Proc. of 1st Workshop on Text Analytics for Cybersecurity and Online Safety TA-COS 2016, 2016.

[13] Ch. Nobata, J. Tetreault, A. Thomas, M. Yashar, and Yi Chang, "Abusive language detection in online user content", in Proc. of the 25th International Conference on World Wide Web, 2016, pp. 145–153.

[14] S. O’brien, "Crisis early warning and decision support: Contemporary approaches and thoughts on future research", International studies review, 12(1), 2010, pp. 87-104.

[15] S. Sela, T. Kuflik, G. Mesch. "Changes in the discourse of online hate blogs: The effect of Barack Obama’s election in 2008", First Monday, V. 17, N. 11, 2012.

[16] S. Shane, and B.Hubbard, "Isis displaying a deft command of varied media", New York Times 31.

[17] A. Schmidt, and M. Wiegand. 2017. "A survey on hate speech detection using natural language processing",in Proc. of SocialNLP 2017, 2017.

[18] W. Wang, R. Kennedy, D. Lazer, and N. Ramakrishnan, "Growing pains for global monitoring of societal events", Science, 353(6307), 2016, pp. 1502-1503.

[19] W. Zeerak and D. Hovy. "Hateful symbols or hateful people? predictive features for hate speech detection on twitter". in Proc. of NAACL-HLT, 2016, pp. 88–93.

[20] G.Weimann, New Terrorism and New Media,Wilson Center Common Labs, 2014.

[21] J. Qiao, H. Jie, Y. Yuexiang, and J. Lin, "Analyzing Malware by Abstracting the Frequent Itemsets in API Call Sequencesm", in 12th IEEE Intern. Conf. on Trust, Security and Privacy in Computing and Communications. IEEE, 2013, pp. 265–270. https://doi.org/10.1109/TrustCom.2013.36.

[22] The National Security Strategy of the Russian Federation, approved by decree by the President of Russian Federation N 683, 31st December of 2015.

ПРЕДСТАВЛЕНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ "БЕЗОПАСНОСТЬ"В ФОРМЕ ТЕЗАУРУСА ДЛЯ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ДОКУМЕНТОВ

Лукашевич Н.В., Добров Б.В.

НИВЦ МГУ имени М.В. Ломоносова

Национальная безопасность является важной концепцией для описания социального феномена защиты жизненных интересов личности, общества и государства от опасностей и угроз. Стратегия национальной безопасности Российской Федерации определяет концепцию «национальной безопасности» как состояние защиты личности, общества и государства от внутренних и внешних угроз, обеспечивающих реализацию конституционных прав и свобод граждан Российской Федерации, достойного качества и уровня жизни, суверенитета, независимости, государственной и территориальной целостности, устойчивого социально-экономического развития Российской Федерации. Национальная безопасность включает защиту страны, а также обеспечение государственной, общественной, информационной, экологической, экономической, транспортной, энергетической и личной безопасности.

Один из важных источников информации для формирования противодействия существующим угрозам - анализ текстовых данных, включая новостные сообщения, аналитические документы, сообщения в социальных сетях. В этой статье мы рассматриваем подход к описанию широкой области национальной безопасности как тезаурус для автоматической обработки документов. Созданный Тезаурус по безопасности имеет модель представления тезауруса РуТез, фактически это расширение тезауруса РуТез и включает терминологию, связанную с социальными, национальными и религиозными конфликтами, экстремизмом и терроризмом, информационной безопасностью.

Мы используем Тезаурус по безопасности в специализированной информационно-аналитической системе и для автоматической текстовой классификации документов в соответствии с несколькими рубрикаторами, включая рубрикаторы угроз, ценностей и региональных проблем. Информационно-поисковая система имеет несколько инструментов поиска, включая поиск по словам, фразам и понятиям, поиск категорий и фасетов. Она также поддерживает создание аналитических отчетов.

Hybrid approach to knowledge extraction: textual analysis and evaluations of experts

Boris A. Kobrinskii and Nikolay A. Blagosklonov
Institute of Modern Information Technologies in Medicine
Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences
Moscow, Russia
kba_05@mail.ru
nblagosklonov@gmail.com

***Abstract*—The problem of extracting knowledge from medical experts for rare diseases causes difficulties due to the need to involve additional information about the clinical manifestations in this pathology. The report will consider a complex two-stage system of knowledge extraction. First, from literary sources using textological analysis. Then an expert evaluation of the information extracted from the literature with the addition of certainty factors.**

***Keywords*—knowledge extraction, textological analysis, textological card, certainty factor, cognitive science world model, expert world model**

I. Introduction

Extraction of knowledge for intelligent medical systems supporting the adoption of diagnostic solutions in hereditary diseases is always characterized by considerable difficulties. This is due to the fuzzy of the changes observed in this pathology and the great similarity of a number of diseases. The complexity of differential diagnostics of orphan (rare) metabolic diseases is determined by genetic heterogeneity, clinical polymorphism (multivariate character space), constant progression (ie, continuous development) of diseases and absence of pronounced clinical signs at an early age. At the same time, early and accurate diagnosis is crucial for the timely prescription of pathogenetic therapy. At present, special attention is drawn to such a subject area as lysosomal storage diseases, in particular mucopolysaccharidosis, sphingolipidosis (GM1- and GM2-gangliosidoses, Gaucher disease, galactosialidosis, Farber granulomatosis, leukodystrophy, Niemann-Pick disease, etc.), mucolipidosis (I cell disease, etc.), glycoproteinoses (fucosidosis, sialidosis, mannosidosis, Pompe disease, Danone's disease, etc.) [1].

Accordingly, the purpose of this study is to extract knowledge about the occurrence and diagnostic significance of clinical signs of metabolic hereditary diseases for the subsequent constraction of an expert system for supporting the diagnostic process in different age periods.

The proposed two-stage procedure involves the knowledge extraction in consecutive order. At the first stage, a textological analysis of special literary sources (articles, monographs, etc.) is carried out to identify the signs (symptoms) characterizing the clinical manifestations of diseases in different age groups using a linguistic scale for assessing the degree of expression of qualitative features, at the second stage – communicative methods of work with experts. To date, expert diagnostic systems for genetic diseases have been built on the basis of a dialogue between a cognitive scientist and an expert-physician and did not include a preliminary textological analysis of literary sources. It should also be noted that, unlike [2], the search for knowledge about the manifestations of diseases is carried out in manuals, monographs, articles, but not as a result of analysis of clinical guidelines and technological maps that determine the necessary patient examinations and the procedure for the actions in the process of diagnosis and treatment.

II. Formation of textological cards

To fix the results of textological analysis, a configuration of data representation in a textological card was developed (Table 1). For each nosological form of the disease, "textological cards" were formed, including a list of signs with their degree of severity and frequency of occurrence in different age periods. As an example, Table 2 presents a textological card for one of the lysosomal orphan diseases – GM1-gangliosidosis.

Table I
Format of data representation in a textological card

Symptom	Intensity	Age group	Frequency of manifestation	Source (author)

The extraction of cognitive knowledge from texts is based on the identification of certain semantic fragments [3], in this case the signs and symptoms of hereditary diseases occurring at a certain age in the disease under consideration. A textological card allows one to present knowledge in any structured document from any number of special (professional) publications with the indication of their authors. The authors' names also provide additional information for experts, since they allow us to indirectly assess the reliability of the fragments of knowledge extracted from the text in each area of medicine (neurology, orthopedics, cardiology, etc.). This is of great importance for the polysystemic diseases under consideration.

To assess the frequency of manifestation of features, a discrete (quantized) linguistic scale was developed that includes 5 levels [4]:

- 5 - very frequently (in 90 - 100%),
- 4 - frequently (in 70 - 80%),
- 3 - average (in 50%),
- 2 - rarely (in 20 - 30%),
- 1 - very rarely (in 10%),

Table II
TEXTOLOGICAL CARD OF GM1 GANGLIOSIDOSIS

Symptom	**Intensity**	**Age group**[1]	**Frequency of manifestation**	**Source (author)**[2]
low height	strongly	1	frequently	1
coarse facial features	weakly	1	average	1, 2, 3, 4, 5
	strongly	2	frequently	
macroglossia	strongly	1	frequently	2
	strongly	2	frequently	
multiple mucopoly saccharidose-like dysostosis	weakly	1	average	2, 4, 5
	strongly	2	frequently	
	weakly	3	frequently	
stiffness of the joints	moderate	2	frequently	1
thickening of wrist joints	-	2	-	1
contractures of the elbow and knee joints	-	-	-	1
deformation of the hand like a "clawed foot"	-	3	-	1
kyphosis	strongly	2	frequently	1
spacity of muscles	strongly	3	-	1
	weakly	4	rarely	2, 6, 7
	weakly	5	rarely	
hypotonicity	strongly	1	frequently	1, 3, 5
	strongly	2	frequently	
ataxia	strongly	3	frequently	6
	weakly	4	rarely	2, 7, 8
	weakly	5	rarely	
convulsions	strongly	1	frequently	1, 2, 5, 6
	strongly	2	frequently	
	strongly	3	frequently	
backward mental development	strongly	1	frequently	1, 2, 3, 5
	strongly	2	frequently	
	strongly	3	frequently	6
neurogical disorders	strongly	2	-	1
	strongly	3	frequently	2, 6
	strongly	4	frequently	6, 7, 8, 9
	strongly	5	frequently	
cherry-red spots on retina	-	1	average	2, 3, 5, 6
	-	2	average	
	-	3	rarely	
corneal opacity	strongly	4	frequently	2
	strongly	5	frequently	
blindness	strongly	1	-	6
	strongly	2	-	
swelling	strongly	1	frequently	2
	strongly	2	frequently	
hepatomegaly	moderate	1	frequently	1, 2, 3, 4, 5, 6
	moderate	2	frequently	
splenomegaly	moderate	1	frequently	1, 2, 3, 4, 5, 6
	moderate	2	frequently	
cardiomyopathy	-	1	frequently	4
	-	2	frequently	

[1] 1 - at birth, 2 - 0-1 years, 3 - 1-3 years, 4 - 4-6 years, 5 - over 6 years

In specific textological cards only the required scale levels are used. In the example of such levels there are only three (frequently, average, rarely).

III. WORLD MODELS OF THE COGNITIVE SCIENTIST AND EXPERT

B.M. Velichkovsky and M.S. Kapitsa [5] consider that the text in natural language is only a conductor of meaning, and the author's intention and knowledge lie in the secondary (semantic) structure of the text that is tuned over the natural text. Earlier it was noted that understanding is the formation of the "second text", that is, the semantic or conceptual structure [6]. Extraction of knowledge from texts is considered precisely as a task of understanding and highlighting the meaning of the text [7]. In the terminology of artificial intelligence is an attempt to recreate the semantic structure in the process of model formation. That is, it is the first step in structuring knowledge.

It should be specially noted that the semantic structures that the cognitive scientist distinguishes from the text include not only the author's world model, but also the cognitive science world model, with its interpretation of similarity, but not the identity of the language definitions of the phenomena described, in this case manifestations of diseases, having a fuzzy character.

The fuzzy of the world is particularly pronounced fuzziness of the changes occurring in the body of the sick person. At the same time, in the practice of assessing clinical manifestations,

[2] The following sources are written in Russian and English languages were used in the preparation of textological card of GM1 gangliosidosis:
1. K. Jones, M. Jones, M. del Campo. Smith's Recognizable Patterns of Human Malformation. 7th Edition, Saunders, 2013, 1016 p.
2. D.L. Kasper, A.S. Fauci, S.L. Hauser, D.L. Longo, J.L. Jameson, J. Loscalzo. Harrison's Principles of Internal Medicine (19th ed.), McGraw-Hill Professional, 2015, 3000 p.
3. R.E. Berman. *Pediatriya. Rukovodstvo. V 8 knigakh. Bolezni ploda i novorozhdennogo, vrozhdennye narusheniya obmena veshchestv* [*Pediatrics. Guide. In 8 books. Diseases of the fetus and newborn, congenital metabolic disorders*], Moscow, Meditsina, 1991. 528 p.
4. A. Hinek, S. Zhang, A.C. Smith, J.W. Callahan. Impaired Elastic-Fiber Assembly by Fibroblasts from Patients with Either Morquio B Disease or Infantile GM1-Gangliosidosis Is Linked to Deficiency in the 67-kD Spliced Variant of b –Galactosidase. American Journal of Human Genetics, 2000, vol. 67, no 1, pp. 23-36.
5. S.V. Serkov , E.Yu. Zaharova, G.N. Levitskij, I.N. Pronin. MRT pri pozdnei infantil'noi (yuvenil'noi) forme GM1-gangliozidoza. Nablyudenie iz praktiki [MRI of the Late Juvenile Form of GM1-gangliosidosis (A Case Report)]. *Meditsinskaya vizualizatsiya* [*Medical visualization*], 2006, no 1, pp. 123-127.
6. J. Campdelacreu, E. Muñoz, B. Gómez, T. Pujol, A. Chabás, E. Tolosa. Generalised dystonia with an abnormal magnetic resonance imaging signal in the basal ganglia: A case of adult-onset GM1 gangliosidosis. Movement Disorders, 2002, vol. 17, no 5, pp. 1095-1097.
7. M. Hirayama, Y. Kitagawab, S. Yamamotob, A. Tokudaa, T. Mutoha, T. Hamano, T.Aita, M. Kuriyama. GM1 gangliosidosis type 3 with severe jaw-closing impairment. Journal of the Neurological Sciences, 1997, vol. 152, no 1, pp. 99-101.
8. U. Muthane, Y. Chickabasaviah, C. Kaneski, S.K. Shankar, G. Narayanappa, R. Christopher, S.S. Govindappa. Clinical features of adult GM1 gangliosidosis: Report of three Indian patients and review of 40 cases. Movement Disorders, 2004, vol. 19, no 11, pp. 1334-1341.
9. E. Roze, E. Paschke, N. Lopez, T. Eck, K. Yoshida, A. Maurel-Ollivier, D. Doummar, C. Caillaud, D. Galanaud, T. Billette de Villemeur, M. Vidailhet, A. Roubergue. Dystonia and Parkinsonism in GM1 type 3 gangliosidosis. Movement Disorders, 2005, vol. 20, no 10, pp. 1366-1369.

there is a fuzzy of the concepts used (attributes) themselves, and of referring them to a certain class. In the process of chronic disease symptoms, as a rule, undergo change, the vagueness of which is differently estimated by each observer of the patient's doctors. In addition, personal characteristics of the patient influenced the manifestation of disease and its dynamics. Thus, fuzzy is determined by the clinical patterns, the variety of known descriptions, the age characteristics of the patient, the knowledge of the expert physicians (involved in the creation of intellectual systems), and their subjective preferences based on past events that have been observed in the past and known from other studies.

The initial fuzzy of descriptions of the observed signs are determined by a number of reasons: transitional states of pathological changes, differences in the manifestations of the disease at different ages of patients within a particular age interval (for example, 1 to 3 years, 4 to 6 years), a rare opportunity to observe patients with orphan diseases. This leads to varying degrees of uncertainty in assessments that characterize the severity of clinical manifestations.

According to Zadeh's granulation principle [8], the degree of granulation of information must correspond to the permissible level of inaccuracy in solving a particular problem, the ability to operate data, knowledge at various levels of detail. The term "granulation" encompasses the processes of composition (formation of larger granules) and decomposition (formation of smaller granules). In this paper, a hierarchy of attributes is used that corresponds to the concept of granularity within the subjective-objective evaluation of characteristics by physicians.

The additional complexity of the process of formation of a textological card is connected with the combination of medical knowledge from publications in different languages and descriptions relating to different ethnic groups with their peculiarities of external manifestations. However, in general, textological cards provide a high-bulk view of the manifestations of the disease in different age periods of life. And the cognitive scientist must be able to choose or reject individual sources.

The structuring of the fragments of the text includes the compilation of a dictionary of signs of different levels – from the lower to the meta signs (for example, the contractures of the joints of the hand – the multiple contractures of the joints of the limbs – multiple dysostosis), taking into account their expression according to the fuzzy linguistic scale. The severity of signs is also represented by a linguistic scale, which includes 7 levels: very strong, strong, moderately, weakly, very weakly, normal value (absence of a pathological sign), and also -1 - impossibility of the manifestation of a pathological sign at a given age. Such a variant of the scale makes it possible to distinguish normal signs characterized by 0, and absence of sign (-1).

At the second stage of knowledge extraction, the cognitive scientist discusses textological cards with the expert. Then, the final variant of the description of the clinical manifestations of the disease, including the level of expression of signs and experts' certainty factors, is chosen. A hybrid structure of process of knowledge extraction by a cognitive scientist from texts and from an expert is presented in Figure 1. As a basis, we took the scheme of extracting knowledge from special texts in the book of T.A. Gavrilova and V.F. Khoroshevsky [7]. Let us consider the existing differences. In the proposed variant, the cognitive scientist structures the text using a linguistic scale. Than the expert, analyze descriptions of signs by different authors in a textological card and, using his knowledge, forms (synthesizes) the final version of the set of symptoms for a specific disease, which are accompanied by certainty factors in the diagnostic significance of these signs, taking into account their severity.

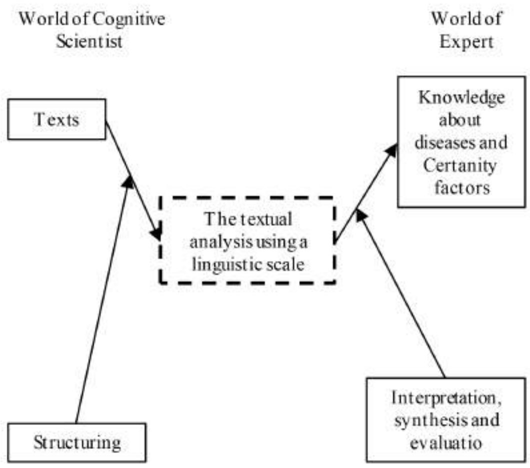


Figure 1. Scheme of complex knowledge extraction by a cognitive scientist – from texts and from an expert

IV. Work with expert

The views of physicians of different specialties complement each other. Their experience and knowledge are reflected in publications. It allows to represent a multidimensional picture of the disease in a textological card. This is especially important for orphan metabolic diseases due to the fact that there is a gradual, but constantly increasing damage to various morpho-functional systems of the body. And different physicians can have their own viewpoint on the intensity of this process. Also, this is due to visual observation or the use of special instruments. And for the evaluation of mental disorders, various psychological tests can be used. The cooperation of experts in conjunction with the questions of the cognitive scientist, including probing, allows in the process of discussion to gain additional knowledge reflecting, perhaps, the intuitive representations of medical experts.

In the process of analyzing textological cards, the cognitive scientist specifies the value of individual features in the expert. In addition, he uses various types of questions to determine the degree of reliability of the cognitive with respect to the meaning of the symptoms and their severity, especially when excluding certain symptoms encountered in literary sources. Participation of two experts allows to reveal the lack of information that can be realized in the communication process "cognitologist - two experts".

Working with two experts is particularly important in genetic diseases. As mentioned above, these are rare diseases. Therefore, each expert has in-depth knowledge of different diseases in patients of various ages. A joint dialogue with the cognitologist of two experts can be called in this case the "complementarity principle", as suggested by Niels Bohr in quantum physics. An analogue in the differential diagnosis of hereditary diseases was the group extraction of knowledge from medical experts when creating an expert system [9]. However, the new system is distinguished by the inclusion of the concept of symptom severity and descriptions by age groups, which allows to take into account the progression of the disease course.

Also, when extracting knowledge for the system being created, in the course of a collaboration discussion between the cognitive scientist and the experts, the issue of attribute connections was decided. The importance of the second expert's participation consisted in revealing in the course of the discussion of implicit relationships of attributes, primarily associative ones. In the older age groups, new connections appear, due to the appearance of previously absent features. Regardless of the participation of the cognitive scientist, in this case, the situation was simulated with precedents, which were actively exchanged by experts.

Certainty factors play an important role in the confirmation phase of diagnostic hypotheses. From the standpoint of cognitive linguistics, they reflect both deliberate and intuitive notions of the medical expert. For actuality this process, the expert was shown textological cards. This approach is based on the idea of psychosemantic methods of reconstructing the implicit (deep) knowledge inherent in the subject, which he may not realize, but which are actualized in the "mode of use" [10].In the diagnostic thinking process, an intuitive-shaped component is of great importance for an experienced physician. Therefore, in a dialogue "cognitive scientist - physician-expert" measure of confidence in each of the signs in the age range under special discussion.

V. Conclusion

The hybrid system of knowledge extraction based on the synthesis of textological analysis and expert knowledge makes it possible to extract more useful information for differential diagnostics. At the same time, work is accelerating at the stage of extracting expert knowledge in the process of analyzing prepared textological cards, since the expert receives structured information, previously selected from various literary sources.

Thus, the textological cards formed for each differentiated nosological form of the diseases allow us to visually present to the expert knowledge of the clinical manifestations of the disease with their peculiarities. Comparing them with each other and with their own ideas, the expert forms an integrated description of the disease, supplemented by confidence factors that allow him to assess his measure of confidence in certain characteristics.

References

[1] J.M. Saudubray, H. Ogier de Baulny, C. Charpentier. Clicnical approach to inherited metabolic diseases. Inborn metabolic diseases. Diagnosis and treatment. 3th ed., J. Fernandes, J.M. Saudubray, G. Van den Bergue Eds., Berlin, Springer-Verlag, 2000, pp.3-41.

[2] A.S. Starostin, I.M. Smurov, M.E. Stepanova. A production system for information extraction based on complete syntactic-semantic analysis. Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference Dialogue (2014, June 4-8), 2014, vol. 20, no. 13, pp.659-668.

[3] M.Yu. Khachay, N. Konstantinova, A. Panchenko, D. Ignatov, V.G. Labunets. Analysis of Images. Social Networks and Texts: 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9-11, 2015. Springer, 2015, 460 p.

[4] B.A. Kobrinskii. Nechetkost' i faktory uverennosti verbal'nykh i vizual'-nykh ekspertnykh znanii [Fuzzy and certainty factor of verbal and visual expert knowledge]. *Nechetkie sistemy, myagkie vychisleniya i intellektual'nye tekhnologii (NSMVIT- 2017): Trudy VII vserossiiskoi nauchno-prakticheskoi konferentsii. V.2t.*[*Fuzzy systems, soft computing and intelligent technologies (NSMVIT-2017): Proceedings of the VII All-Russian Scientific and Practical Conference. In 2 vol.*], 2017, vol. 1, pp. 83-91

[5] B.M. Velichkovsky, M.S. Kapitsa. *Psikhologicheskie problemy izucheniya intellekta* [*Psychological problems of studying intellect*] in *Intellektual'nye protsessy i ikh modelirovanie* [*Intellectual processes and their modeling*], E.P. Velikhov and A.V. Chernavskii Eds., Moscow, Nauka, 1987, pp. 120-141.

[6] S.A. Sirotko-Sibirskiy. *Smyslovoe soderzhanie teksta i ego otrazhenie v klyuchevykh slovakh* [*The semantic content of the text and its reflection in key words*], Abstract of the dissertations ... candidate of philological sciences, Leningrad, 1968.

[7] T.A. Gavrilova, V.F. Khoroshevsky. *Bazy znanii intellektual'nykh sistem* [*Bases of knowledge of intellectual systems*], St. Petersburg: Peter, 2000, 384 p.

[8] L.A. Zadeh. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy sets and systems, 1997, vol.90, pp.11-127.

[9] B.A. Kobrinskii. Retrospektivnyi analiz meditsinskikh ekspertnykh sistem [Retrospective analysis of medical expert systems]. *Novosti iskusstvennogo intellekta* [*News of Artificial Intelligence*], 2005, no 2, pp. 6-17.

[10] V.F. Petrenko. Lichnost' cheloveka – osnova ego kartiny mira [Personality of man - the basis of his picture of the world]. *Rossiiskaya assotsiatsiya iskusstvennogo intellekta* [*Russian Association of Artificial Intelligence*], 1997, pp. 9-24.

ГИБРИДНЫЙ ПОДХОД К ИЗВЛЕЧЕНИЮ ЗНАНИЙ: ТЕКСТОЛОГИЧЕСКИЙ АНАЛИЗ И ОЦЕНКИ ЭКСПЕРТОВ

Кобринский Б.А., Благосклонов Н.А.
Институт Современных Информационных Технологий в Медицине
ФИЦ "Информатика и Управление"РАН
Москва, Россия

Проблема извлечения знаний у медицинских экспертов при редких заболеваниях вызывает трудности вследствие необходимости привлечения дополнительной информации о клинических проявлениях при этой патологии. В докладе будет рассмотрена комплексная двухэтапная система извлечения знаний. Вначале из литературных источников с использованием текстологического анализа. Затем экспертная оценка информации извлеченной из литературы с присоединением факторов уверенности.

The ellipses in the geometric texts

Ksenia Naidenova
Military medical Academy
Saint-Petersburg, Russia
ksennaidd@gmail.com

Kurbatow Sergey
Scientific research centre
for electronic computer technology
Moscow, Russia
curbatow.serg@yandex.ru

***Abstract*—It is presented processing of ellipses occurring in texts of plane geometry tasks described in natural language. It is proposed a general approach to the processing of ellipses based on cognitive semantics. The ellipsis processing is based on the parallelism between syntactic structures and using geometric semantics. Some examples of the ellipsis processing with its limitations are described. The types of ellipses most commonly encountered in geometric tasks are highlighted. It is noted the ability to recognize ellipses and their resolution in the framework of cognitive semantics.**

***Keywords*—ellipse, a geometrical problem, cognitive semantics**

I. Introduction

The ambiguity of natural language caused by homonymy, has long been studied by computer linguistics. However, the ambiguity associated with the omission of a thinkable language unit (ellipsis) in the text has been actively analyzed by automatic processing natural language relatively recently [1], [2]. Although, in the theoretical linguistics, the ellipticity got enough coverage [3], [4], the restoration of ellipses in the systems of syntactic text analysis is clearly developed not enough. This is largely due, first, to the fact that eliminating the ellipticity is of a subordinate nature with respect to the actual syntactic analysis and, secondly, to the complexity of recovering ellipses. The complexity is explained by the potential need to take into account a number of contexts: current sentence, adjacent sentences, already established syntactic relations and, finally, semantics of text.

This work is divided into two parts. In the first part, it is described how to handle the ellipticity in a specific holistic system of solving plane geometry tasks in natural language description. The second part proposes a general approach to the processing of ellipses based on cognitive semantics, aimed at a much broader range of applications (both in subject areas and in different natural languages).

A. Treatment of ellipses in a system for solving geometrical tasks automatically

The general scheme of an integrated system to solve geometrical problems automatically in natural language is given in [7]. The system includes a linguistic translator, an ontology and a graphical interface for output result (drawing of NL-explanation of the solution process). An expansion system provides the interpretation of more complex tasks (to identify their logical structure) associated with language processing in the limits required by a number of modifications of the parser. This section describes an extension of the system to correctly interpret elliptical (incomplete) sentences. The correct interpretation of the ellipses is based on the parallelism between the syntactic structures (the ellipticity) and knowledge about the general concepts included in the structural schemes of ontology. Therefore, the interpretation essentially uses both levels of linguistic processing - syntactic and semantic ones.

The language translator creates a syntactic structure and determines that some of its elements violate the language rules. For example, there is no noun for adjective, pretext is at the end of sentence, number does not have a mandatory measuring unit, and so on. The basic criteria for determining ellipticity are studied by linguists [6]. Based on these criteria, recorded in the ontology, the translator identifies the fragments of syntactic tree that admittedly contain ellipticity. Next, with the use of algorithms described in short below in 1.2, the identified ellipses are restored.

Specifically, in sentence «the radius of the first circle is set to 12 cm, and the second 10», the elements «second» and «10» define the ellipticity. As a result, two syntax structures must be formed:

- The radius of the first circle is 12 cm.
- The radius of the second circle is 10 cm.

These structures are further processed by the system mechanisms of paraphrasing to obtain an ontological representation of sentence in formal terms of the subject area [7].

The concept "‘paraphrasing"’ has been proposed by the well-known Russian linguist Apresyan in [10]. In our system, we use an adaptive variant of this concept. The conception of paraphrasing assumes that any class of sentences corresponding to one and only one sense can be reduced to the simplest or canonical phrase composed only of the lexemes expressing most clearly the based concepts of the sentences. Thus "‘paraphrasing"’ is based on the following proposition in [10]: «One of the fundamental properties of human languages consists in the fact that if there are several synonyms, in the broad sense, to express some concept, then only one of them turns out to be privileged, canonical, or prototypical for expressing this concept.» [10].

In particular, such canonical concepts in the plane geometry are, for example, «point», «line», «plane» and «belong», «to lie between» and «to be congruent».

In the « Space shipyard » domain [8], the canonical concepts are «tank», «adapter», «reinforcement», «dock», «point of joining» and some others.

Thus, the rules of paraphrasing provide only one canonical form for a group of sentences having the same sense. For example, sentences "a point located on the straight line", "the straight line passing through a point", "a point belonging to the line", "a point lying on the line segment", etc. are reduced to the following canonical phrase, namely, "point belongs to straight line".

This canonical phrase is displayed in the ontological representation in the form of the following triplet « point lies line ». It is noted to emphasize that the members of the triplet (objects and relations between them) are not dependent on a particular language. Therefore, the corresponding rule of paraphrasing contains, in its left part, the objects and relations depending on the language, but, in its right part, the objects and relation are invariant with respect to the different languages.

The rules of paraphrasing are divided into two classes, the first one consists of rules in which both parts of them are some generalized syntactic structures; the second one consists of rules having the canonical descriptions in their left parts and the semantic descriptions in their right part. The second class rules can be used for transforming the ontological structures into the corresponding natural language texts. It is reasonable to apply the rules of the first class for equivalent synonymic transformations of the synthesized structures to retrieve texts in the most appropriate manner in a considered application domain.

B. Algorithm for resolving ellipticity

B. The processing algorithm is based on the ontology knowledge reflecting the semantic hierarchy of word forms in the syntactic structure and rules of natural language. To a first approximation, the algorithm can be described as follows:

- to segment syntactic structure into two segments: a complete one without ellipticity and the other containing an ellipticity (generally, there are noun groups);
- in the elliptical structure, identify elements that can be usefully matched with full syntactic structure elements to be used for resolution of ellipticity;
- in the full syntactic structure, identify candidates to replace with the elements found in the previous step;
- perform a replacement and get a complete syntactic structure.

In the example from subsection A «first» is replaced by «second» and «12» by «10» because they correspond to the same concepts of ontology. Here we have different objects and the same type of attribute (length). In sentence «the perimeter of the triangle is 37 cm and area – 20 cm», we have the same object and the different types of attributes. This seemingly simple algorithm allows to successfully recover not only geometrical ellipses, but several others, described, for example, in [2]: In sentence «twenty years of this dance form the age, forty – the history», «twenty» is replaced by «forty» and «age» is replaced by «history».

Naturally, the ontology should contain the concept «time-interval» binding hierarchically «age» and «history». It is obvious that the standard tools for editing ontology ensure that the algorithm is correctly carried out without reprogramming. More examples: «He went to the pharmacy and his brother – to the post office». «He» is substituted by «his brother» (It is a noun group: whose brother? him). Here the noun group could be: «its half-brother on the maternal line». The mapping in the ontology is formally performed for the roots of words in syntactic tree of noun groups. It is for this reason that the structure for sentence «Triangle ABC is inside the circle and the square – outside» is correctly restored.

C. Limitations

Of course, many cases of ellipticity can not be processed by the above algorithm. Example: «There are seven circles. Radius of one 5 cm, two other – 3, and the others – 10 cm.» We have multiple ellipticity in this example. A similar example from [2]: «Anemones discard tentacles, crayfishes – claws, lizards – tail». In some cases, there is an ambiguity at the comparison level. At the logic level, two options were analyzed:

1) continuation of work and eliminating ambiguity at the stage of semantic processing (canonical syntactic structures);
2) enter in to the ontology not only knowledge about the hierarchy of concepts, but the rules of preferences when choosing a candidate for replacement (substitution).

A decision on whether to choose one of the options or a combination of them is the subject of further study. In any case, the focus is on the universality of knowledge-based algorithm. Note that the algorithm described was tested not only for geometric texts but also for a number of others, in particular in the text of so-called «space shipyard»[8]. The ontology fragment, which describes the concepts of this area, specifies the hierarchy (visualization object) isA (tank, adapter, armature, solar panel, energy block, and so on) isA (specific types of objects: Tank-b, Truss-c, etc.). According to this hierarchy, sentences of the type «diameter of the first tank equals 1.7 cm and the second 5.8» or «the length of the solar panel is 15 cm and the weight 40 grams» is successfully restored.

It should be noted that the question of a clear ellipticity criteria and methods for restoring the full structure of sentences has not been fully resolved within the framework of a generally accepted linguistic theory. Further, the following cognitive approach to the resolution of ellipses provides a significant extension of the possibilities described above and focuses on self-learning ontologies for the treatment of a wide spectrum of incomplete sentences.

II. ELLIPSES PROCESSING BASED ON COGNITIVE SEMANTICS

1) General information: The cognitive approach is opposed to traditional formal one, based on Tarski and Hilbert's ideas, playing a major role in the computer paradigm of most AI research. Two main differences of the cognitive approach from the traditional formal one are to solve categorization and semantics problems. Categorization is a problem of creating concepts (categories) and structuring the conceptual human system.

Resolving ellipses in natural language texts remains one of the most difficult and unsolved tasks in linguistics, despite the abundance of proposed methods based on semantic-parsing of sentences. The syntax reveals the structure of ellipse and the similar part of the sentence without the ellipse, and the semantics deal with word values. However, as the example from (Umberto Eco. The role of the reader. Exploration in the semiotics of texts. – Moskva: Publishing House ACT: CORPUS, 2016 (page. 62)) shows, resolving ellipses is based on the understanding of context (text theme), the meanings of words and collocations: «Charles makes love with his wife twice a week. So does John».

2) Ellipsis classification in geometrical tasks: To study the typology of ellipses in geometric tasks, we used a body of texts that is listed in [9]. From a variety of sentences, we have selected several types of ellipses: ellipses using the sign «-» (ellipses with missing a predicate, ellipses with missing a verb), ellipses without using the sign «-» (ellipses with a skipped noun, with a skipped pronoun, with a skipped predicate. The characteristics of these types are analyzed, not necessarily related to the subject area (geometry) and the language (Russian).

A detailed discussion of the selected ellipses in terms of their properties and structure is given in [9]. Resolving complex cases of ellipticity requires understanding the context of geometrical tasks. To account for context, we enter the concepts of cognitive models of geometric object and action with object. The model of object includes the properties of the object, the actions as a result of which this object can be created, the actions that object can perform itself, and in which it can participate; the object model also includes the elements of which object is composed, and the elements of which it is a part. Cognitive schema will show how the object is formed and its position in space in respect to other objects. The model of action, naturally, includes objects involved in this action and its result as a geometrical configuration of objects. And all the parts of the cognitive models will be associated with the words that appear in the text body in question.

Within the proposed model, text analysis becomes cognitive-driven, and the parser plays a subordinate role in the process. In processing ellipses resolution based on cognitive models, it is possible to synthesize text that describes the geometric situation and compare the text generated with the text to be analyzed. Ontology contains theoretical knowledge in the area of plane geometry and knowledge of methods of solving plane geometry tasks of various types (computational, construction, challenge for proof). Ontology takes the burden of solving the problem and creating the drawing accompanying solution (visualization of solution). The Cognitive Analyzer of text runs incrementally and transmits the converted and meaningful text to the ontology in the language required by the ontological block.

A. The structure of cognitive models of objects and actions

The cognitive structures correspond to the semantic structures of the situations described in text. They should be aligned with the narrative structures of sentences. A word can have multiple values, but only one meaning, at least in mathematical texts. Ellipsis (omitting words, economy of text) is possible because the preceding text determines unambiguously (uniquely) the meaning of each word and situation, and the meaning remains unchanged. In cognitive model of object, the following relationships are important:

- object can perform some actions;
- object can be subjected to action of other objects;
- object can have spatial and temporal relationships (earlier, later, already built, already given) with other objects;
- object can be composed of some other objects;
- object can be a part of some other object (objects);
- object has properties, some of which (call them actant ones) are related to the actions that the object commits (intersects-intersecting, lies - lying) or the actions that are committed over it (has been given - given, has been formed – formed, cut off, embedded). Thus, the actant properties of objects are directly displayed in the morphological forms of the words describing these properties;
- the relationships between the properties of one geometrical object, between object and its parts are realized through implications: if "‘center"’, then "‘a circle"’; if "‘radius"’, then a "‘circle"’; if "‘circle"’, then "‘circumscribed about or inscribed in"’; if "‘inscribed in"’, then "‘in an object"’; if "‘bisector"’, then "‘bisector of an angle"’; if "‘bisector of angle"’, then "‘the vertex of angle from which it originates"’; if "‘bisector"’, then "‘the angle from which it comes is divided in half"’; if "‘bisector"’, then "‘it is the axis of symmetry of angle divided in half by this bisector"’.

With cognitive patterns, we associate such phenomenon as a cognitive wait for appearing certain words and narrative constructions in the text. All cognitive models can be explicitly defined based on geometric semantics, and are associated with speech parts and typical collocations with their grammatical categories at the sentence level. The creation of cognitive models of objects and actions for plane geometric tasks in the proposed approach is performed in a step-by-step mode by the use of a given text corpus. An example of creating a cognitive model "Bisector" across the text of many tasks is shown in [9].

It is a problem of considerable interest to apply plausible reasoning for the resolution of ellipses, including analogy, generalizations, specialization, use of implications, forming hypothesis and many others. Example 2 (Section 1.3) “Anemones discard tentacles, crayfishes – claws, lizards – tail” can be resolved by analogy. Anemones, crawfishes, lizards are living things, tentacles, claws, tale are parts of their bodies, then crayfishes, claws and lizards, tail can be connected by the action "discard".

III. Conclusion

The processing of ellipses is given in a specific system of plane geometry tasks with a natural language description. The ellipse resolution is based on using in parallel the syntax structures of sentences and the geometry semantics. A broader approach to ellipses processing based on cognitive semantics has been proposed. The approach gives a classification of ellipses (across a geometric text corpus) and introduces the cognitive models of geometry objects and actions. The model proposed allows to view the structure of automated analysis of geometric texts as a cognitively controlled parsing. Further research is envisaged in two directions:

1) enhancing the capacity of the existing system;
2) algorithmization and software implementation of the cognitive approach.

Acknowledgment

The work is executed at financial support of RFBR, project No 15-07-03847, "Interpretation of the technically-oriented natural language texts using application ontologies" .

References

[1] Malkovsky M. et al. Restoring ellipse as a task of automatic text processing. Program products and systems, No. 3 (107), 2014, http://www.sworld.com.ua/konfer35/365.pdf

[2] Kobzareva T., Epifanov M., Lakhuti D. Finding the antecedent of the ellipsis of a fragment containing predicate (automatic analysis of Russian sentence). The Proceedings of 15th National Conference on Artificial Intelligence wth International Participation CAI-2016, Vol. 2, pp. 56-62.

[3] Adametc P. Semanticheskaya interpretatsiya «znachimykh nulei» v russkikh predlozheniyakh [Semantic interpretation of "significant zeros" in Russian sentences]. Proceedings of papers in honor of Dean C. Worth, for his 65 anniversary. M.: East Literature, Russian Academy of Sciences. 1995. P. 9-18. (in Rus)

[4] Vardul I. To the question of the ellipse phenomenon //Invariant syntactical values and structure of sentence. M.: Nauka, 1969. P. 59-70.

[5] Minjajlov V. Detection of elliptical sentences by means of classification algorithms. In: Problems of Physics, Mathematics and Technology, No. 3 (28), 2016.

[6] Larkina, A. Ellipsis in the modern French language, The thesis of a PhD in Philological Sciences (specialization 10.02.05). 2009.

[7] Kurbatov S., Vorobyev A. Ontological Solver of Geometry Problems in natural language description. // Proceedings of 15th National Conference on Artificial Intelligence with International Participation (CAI-2016), Vol. 1, P. 56-63.

[8] Lobzin A., Khakhalin G., Kurbatov S., Litvinovich A. Integration based on natural language and image ontology in the system Text-To-Picture. The Proceedings of 8th Scientific-Practical Conference «Integrated Models and Soft Computing in Artificial Intelligence», Phizmatlit, 2015, 296-305.

[9] Processing of ellipses based on cognitive semantics, http://www.eia–dostup.ru/cognitiv.htm

[10] Yu.D. Apresyan, I.M.Boguslavskii, L.L.Iomdin i dr. Lingvisticheskoe obespechenie sistemy ETAP-2 [Lingware of the ETAP-2 system] M.: Nauka, 1989.

[11] Eco Umberto. The role of the reader. Exploration in the semiotics of texts. – Moskva: Publishing House ACT: CORPUS, 2016 (page. 62))

ЭЛЛИПСИСЫ В ГЕОМЕТРИЧЕСКИХ ТЕКСТАХ

Найденова К. А.
Военно-медицинская академия
Санкт-Петербург, Россия
Курбатов С. С.
НИЦЭВТ ОАО - АО «НИЦЭВТ»
Москва, Россия

Аннотация. В работе описана обработка эллипсисов в конкретной системе решения планиметрических задач по описанию на естественном языке и предложен общий подход к обработке эллипсисов на основе когнитивной семантики. Разрешение эллипсиса базируется на параллелизме синтаксических структур и использовании семантики геометрии. Даны примеры обработки и описаны ограничения. Выделены типы эллипсисов, наиболее часто встречающихся в геометрических задачах. Предложен подход к распознаванию эллипсисов и их разрешения в аспекте когнитивной семантики.

Ключевые слова: эллипсис, геометрические задачи, когнитивная семантика.

Knowledge acquisition based on natural language texts

Hubarevich Nastassia
Belarussian State University of Informatics and Radioelectronic
Minsk, Belarus
stasia@tut.by

Boyko Igor
Center for System Analysis and Strategic Studies of the National Academy of Sciences of Belarus
Minsk, Belarus
igor_m_boyko@hotmail.com

Semenyaka Anatoly
Belarussian State University of Informatics and Radioelectronic
Minsk, Belarus
toli44777@gmail.com

Hardzei Aliaksandr
Center for System Analysis and Strategic Studies of the National Academy of Sciences of Belarus
Minsk, Belarus
alieks2001@yahoo.com

***Abstract*—The article describes a method of computer analysis of natural language texts and automatic filling the knowledge base using OSTIS technology. The method helps to implement semantic-syntactic analysis of texts and then analyse its context based on specific subject domain ontologies.**

***Keywords*—natural language processing, universal semantic code, text understading**

A cognitive approach in the field of artificial intelligence is under intensive development and it involves modeling various aspects of text understanding [1].

According to D.Pospelov: an intellectual system understands some text if able to answer questions about its content with an accent on the deep semantics but not just pure facts. The sense of the text reflects knowledge represented by a formal language as a semantic equivalent where objects and relations between them are not limited with linguistic categories and where they represent real world objects and relations [2].

I. Introduction

Main components of natural language text analysis are syntactic and semantic.

Syntactic analysis describes a syntactic structure of the text. For Russian text analysis, we use two approaches: the dependency grammar [3] and the grammar of the direct components [4].

Semantic analysis is related to computer text understanding. There are several popular approaches to semantic analysis:

- semantic role labeling [5] represents semantic roles of words in a sentence through frames;
- entropy based frames as an extension of the role labelling [6];
- compound use of frames with the functional grammar [7];
- semantic text structuring with the generative lexicon [8], [9];

For conversion natural language texts into knowledge representation programming language like Prolog researchers use a semantic analyser preliminary trained on syntactic processing of the natural language texts [10].

The analyser integrates semantic and syntactic analysis in a single procedure or divides on two procedures with results depending on each other. For example: the technology ABBYY Compreno [11] is processing Russian texts in parallel semantically and syntactically affecting each other; in the system ETAP-3 [12], on the first step, the syntactic analyser uses word semantic features and builds the semantic structure on the next step.

Despite of existence of different approaches to the semantic-syntactic text analysis there is a list unresolved problems. The article considers some of them:

- Homonymy resolution [13];
- Synonymy resolution;
- Named entity recognition [14];
- Ellipsis recognition [15];
- Understanding different forms of the same word;
- Metaphor resolution;
- Automatic conversion of the text sense into semantic formalisms for computer processing (understanding);

The article describes an approach to the semantic-syntactic text analysis and automatic filling of the knowledge base, i.e. the implementation of the computer text understanding.

Generally, the process of the natural language text understanding, with the aid of the OSTIS-system, includes next stages [16]:

- Linguistic (graphematic) analysis determines the text components: paragraphs, sentences, words.
- Morphological and syntactic analysis define grammar relations between words in the sentence.
- Semantic analysis generates the sc-text equivalent to the input text.
- Pragmatic analysis integrates the sc-text into the OSTIS-system knowledge base. During the stage:

- •• synonymic sc-elements from the knowledge base form pairs with sc-elements of the processed text [17];
- •• the terminological system of input text notions aligns with notions of the knowledge base.
- Discourse analysis (context) of the input text. The analysis can have several levels:
  - •• analysis of neighboring phrases, sentences, and paragraphs (text analysis);
  - •• context analysis from other sources (extended text analysis). For that, the system should operate with additional information sources: visual and sound. So, a phrase "Look what the man has done" can be analysed and additional information from the visual observation will be integrated;
  - •• context analysis on the basis of information about source, author, location, and time of the text publication [18].
  - •• context analysis of the internal knowledge base having some fragment representation of the input text.

Discourse analysis applied on all stages of the text analysis.

The kernel part of the proposed approach is in building the linguistic ontology that integrates: knowledge about linguistics, rules of syntactic and semantic analysis of texts, and the specific subject domain ontology. The ontology includes knowledge about objects and their relations, i.e. gives a formal description of some fragment if a model of the world [22]. Thus, the linguistic ontology includes knowledge about the text and methods of its processing, and the subject domain ontology includes knowledge about some fragment of the real or virtual world, described in the text, and operates with knowledge about rules and methods of specified knowledge processing.

A universal linguistic interface component is built on the bases of the OSTIS technology [19] to be applied in any ostis-system. Every ostis-system consists of the knowledge base in a view of the formal model integrating all knowledge kept in the system and of knowledge processing machine uniting all program agents of the entire system.

SC-code used for formal information representation. Texts of SC-code are interested into the semantic net with basic theory-set interpretation.

The advantages of the technology are following:

- use of unified tools for representing different kinds of knowledge including meta knowledge, which allows to describe all necessary information for analysis;
- use of formalisms allowing to specify as kept notions of the knowledge base so the external computer files;
- provide the ostis-system modifiability, i.e. the ability to extend its functionality.

II. The Structure of the Knowledge Base of the Natural Language Interface

The knowledge base of the natural language interface has two parts: the language part common for the entire system and the subject part defined by the specific subject domain.

In the ostis-system the subject domain (SD) is a specific structure consisting of:

- main objects of research (OR) – primary and secondary;
- different classes of OR;
- various links where main OR are components, and other types of the links being OR itself having a different level structure;
- different classes of the links (relations);
- different classes of objects not being OR nor the links nor components of the links.

The article considers the subject domain of 'History' represented in the system with sub domains and with corresponded to them ontologies:

- **SD of artefacts** describes all historically valuable and artificially created material entities as a result of a purposeful activity;
- **SD of urban planning** describes immovable monuments of history and culture.
- **SD of persons and social communities** considers a person and all arising from his activity communities of people.
- **SD of ideas** describes compiled on the basis of purposeful activity results [23].
- **SD of historical actions and events** described according to the principles of Semantic coding. [20], [21], [22].

Every SD represented by the set of ontologies [24]:

- **Structure specification** describes roles of the notions and relations of the specific SD with other SDs;
- **Theory-set ontology** describes theory-set relations between the notions of the SD;
- **Logical ontology** includes a system of statements about the notions of the SD;
- **Terminological ontology** represents a system of main and complementary terms (names, signs) corresponding to concepts and relations of the SD and a description of constructing rules of entity terms used as elements (instances) of the concepts and relations.

The constructed knowledge base includes the own knowledge processing machine having program agents implementing logical reasoning based on a hierarchy of statements comprised in the logical ontology.

The linguistic SD represented the language part, which is the common component for all designed systems [25].

Look at a very general structure of the linguistic SD represented in the SCn-language.

SD of Russian language texts
=> *specific SD*:*
- • *SD of Russian language syntax*
- • *SD of Russian language morphology*
- • *Lexical SD of Russian language*

The knowledge-processing machine of the natural language interface
<= *decomposition of an abstract sc-agent*:*
{

- *The abstract sc-agent translating external texts into the knowledge base*
- *The abstract sc-agent verifying the knowledge base*
  <= decomposition of the abstract sc-agent:*
  {
  - *The abstract sc-agent verifying correspondence of the relations to its domains*
  - *The abstract sc-agent verifying action specification to its class*
  - *The abstract sc-agent for searching synonymic elements*
- *The abstract sc-agent for context interpretation*
  <= decomposition of the abstract sc-agent:*
  {
  - *The abstract sc-agent identifying the essence correspondent to defined criteria*
  - *The abstract sc-agent identifying relations between entities*
  }

III. Example of the approach

The example shows the analysis of the following sentences:

- Лангбард И.Г. спроектировал Дом правительства в 1934 году [Langbard I.G. designed the Government house in 1934 year].
- Лангбард И.Г. быд архитектором [Langbard I.G. was an architect].
- Лангбард был архитектором в Комиссии по делам архитектуры [Langbard was an architect in the Commission on Architecture].

Several constraints defined:

- The input of the system are simple narrative Russian sentences;
- The sentence is completed and has a sense;
- Text analysis and understanding implemented for the specific SD (History) and formally represented in knowledge base;
- The knowledge base includes entity signs denoting proper names of persons, establishments, and buildings, which used in the sentences.

Here is a step-by-step analysis of the sentence «Langbard I.G. designed the Government house in 1934 year».

Step 1 Graphematic analysis generates a set of individual words with a given order in the sentence (Fig. 12).

Step 2 The obtained fragments are compared with the samples represented in the knowledge base where they correspond to certain lexemes

Description of lexemes is stored in the linguistic knowledge base. Some of them are on the Figs. 1 - 2

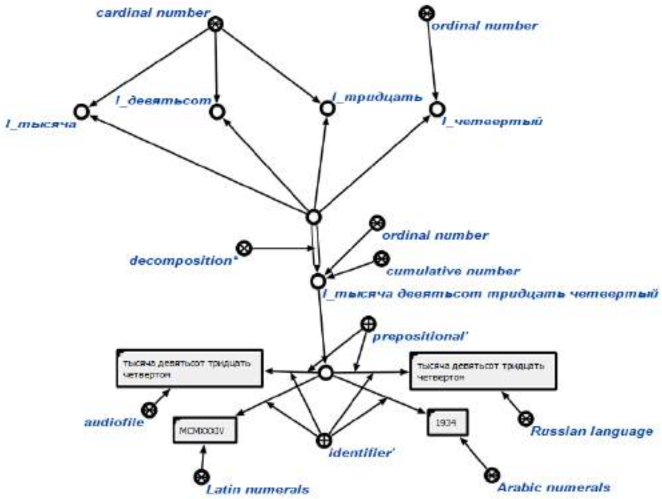


Figure 1. The lexeme '1934'

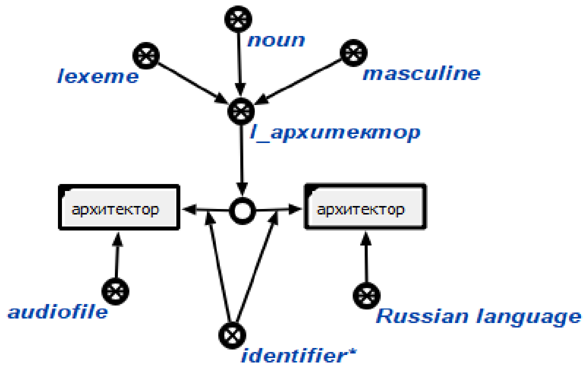


Figure 2. The lexeme 'architect'

The lexemes describing the SD notions at the same time are a part of the linguistic ontology. So, the context of the words described, including it synonyms, through the linguistic ontology.

Named entity recognition is implemented by composing together elements with identical identifiers in the knowledge base. However, naming of the real world objects is much wider of the identification in the knowledge base. Such cases are taken into consideration by the terminological ontology and composition of elements denoting the same object anyway will occur.

Step 3 For each fragment of the text formed a pair «sample-text fragment».

Step 4 The input text is analysed on interrelations between words in the sentence and a syntactic tree is constructed with use of a third-party software [26].

Step 5 The agent translates the result into a semantically equivalent structure in the knowledge base and, as a result, between the words in the sentence appear relations characterising its relationships in the sentence (Fig. 13).

Step 6 In the example, the verb "design" will be found in the historical knowledge base which, in turn, according to the Universal Semantic Code (USC) verb classifier belongs to the class “reproduce” [20], [21].

The ontology of actions is built on the USC basis. The classification of actions supports an idea that different classes of actions differs by the structure of the components.

The USC system satisfies several demands:

- Every USC string corresponds to only one sense;
- Declarative knowledge should be represented in the form of a procedural one. It is important to know not “what is an object in the system", but "what an object performs in the system";
- The means of knowledge representation are not formally separated from the means of knowledge transformation [27].

Thus, each action class has its specific for this action class of performers, mediators, initiators, objects, etc. Based on the characteristics of action classes, the agent is able to transform a verbal description of the action into its formal representation.

Step 7 Homonymy resolution is not applied to the first sample sentence so it will be applied to the two left sentences where accordingly the word ‘architect’ means an occupation and a job position.

In the result of analysis, the agent finds the word ‘architect’ in the historical domain can mean:

1) Subset of the set ‘occupation’

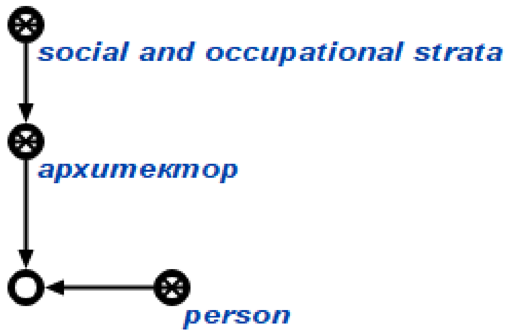


Figure 3. The instance of the class

For the sentence “N was an architect”, without further analysis of words in the sentence, the shown construction will always be formed with the meaning that someone belonging to the class ‘person’ also belongs to the occupation ‘architect’.

2) The role relation ‘architect’, whose domains are the social organisation and the person

The role relation ’architect’ means that some entity of the class ‘person’ has a role of ’architect’ in a scope of the class ‘social organisation’. This construction will be created in further sentence analysis if both relations will be found.

3) Binary relation ’architect’ with domains ‘person’ and ‘building’.

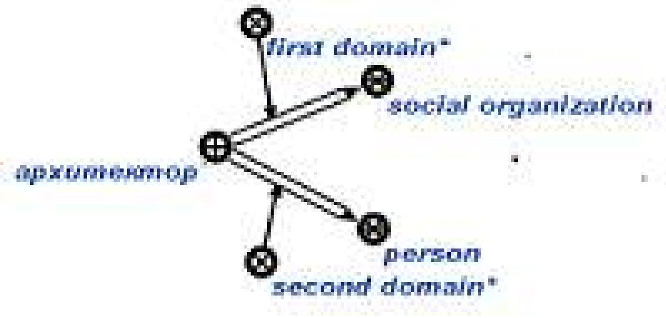


Figure 4. The role relation ’architect’

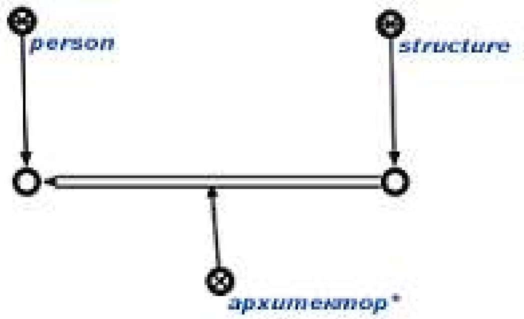


Figure 5. The role relation ’architect’

Working with first sentence the agent is able to find Langbard I.G. belongs to the set ‘architect’ and to form the correspondent structure in the knowledge base.

Figure 6. Representation of the sentence “Langbard I.G. was an architect”

For the second sentence the construction on the Fig.6 will be formed again and because the construction already exists in the knowledge base the synonymic elements will be composed together based on the identity of its identifiers and roles in the formalism.

Then, in the sentence will be found notions belonging to the class ‘person’ – Langbard I.G. and to the class ‘social organisation’– Commission on Architecture. Therefore, the next construction will be formed Fig. 7.

Figure 7. Representation of the sentence "Langbard was an architect in the Commission on Architecture"

The construction will be added to the knowledge base as a unique element and became a part of the entity description named Langbard I.G.

For the binary relation 'architect' the second data domain belonging to the class 'building' will not be found and the agent ends functioning because there are no other variants of use of the notion 'architect'.

Thus, the homonymy resolution in the scope of the system implemented when specifying notions. *Step 8* In a result of transforming the text into an equivalent formal structure, it becomes possible to derive logical results based on the data available in the sentence. It is the essence of the context analysis.

The further work of the agent for revealing the relations between entities uses the hierarchy of statements described in the history SD, while the search of statements will be limited only with concepts extracted within the semantic-syntactic analysis of the text, or resulting from contextual analysis.

Accordingly, the set of logical rules will be finite and limited with the Logical ontology of the specific SD. The effectiveness and completeness of the contextual analysis depends directly on the completeness of ontologies.

Consider the results of text contextual analysis for the proposed fragment 'Government House' within the knowledge base.

Initially, the phrase "Government House" has already been correlated in the knowledge base on history with the set 'building'. The same will happen with the term 'government'. As a result, the following structure is formed.

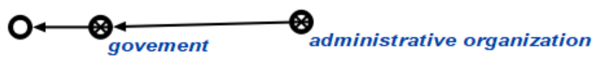


Figure 8. Determining the notions through the knowledge base on history

The construction means: there is some entity that belongs to the set 'government', which in turn is a subset of the set 'administrative organisation'.

Then, the agent analyses and selects all types of relations existing between instances of the sets 'building' and 'organisation', and alternately applies existing statements for the found relations.

Fig. 9 demonstrates the rule for analysis possible relations between the notions 'government' and 'building'.

After applying this rule, the existing structure in the knowledge base is supplemented with new information.

Fig. 10 demonstrates as the existing structure was supplemented with information that with a probability of 0.3 the government owned the building called "Government House", and with the probability 0.7 the government was located in the building.

The coefficient of probability is calculated by counting the frequency of used relations between the entities, where "1" is the total number of all found relations, and the coefficient is the fraction of each relation in the total number. The coefficient reflects the current state of the knowledge base on history, and can change with its further completion.

Further, the proposed changes should be approved and placed into the knowledge base or rejected.

Step 9 Then, the agent for identifying the entity that meets the specified criteria is activating.

Having found the date in the sentence, the agent is replacing the node without the identifier belonging to the set 'government', with the node with the identifier "Council of People's Commissars of the BSSR", composing it together with the corresponding elements in the knowledge base on the ground that in 1934 the Council of People's Commissars of the BSSR played the role of the government (Fig11)

As a result of the analysis, all formed structures will become the part of the knowledge base, and the historical system will have enough knowledge to answer following questions:

- Who designed the Government House?
- What is the Government House?
- What is the government?
- What was located in the Government House?
- Who was the owner of the Government House?
- Which organisation acted as the government in 1934?
- Who was Langbard I.G.?

IV. Conclusion

The proposed approach of the text analysis has the following advantages:

- All stages of the text analysis implemented in the same environment, what excludes the need to address compatibility issues between different solutions;
- Semantic net fragments equivalent to the text become the result of the text analysis, and text become the part of the knowledge base and this allows further processing of the received knowledge by standard methods, and also allows to organise the automatic filling of the knowledge base from texts;
- Common use of linguistic knowledge bases and any other SD allows to resolve problems of homonymy and synonymy easier and faster;
- The availability of knowledge bases for the specific SD allows use of logical ontologies for contextual analysis of the text within the given SD.

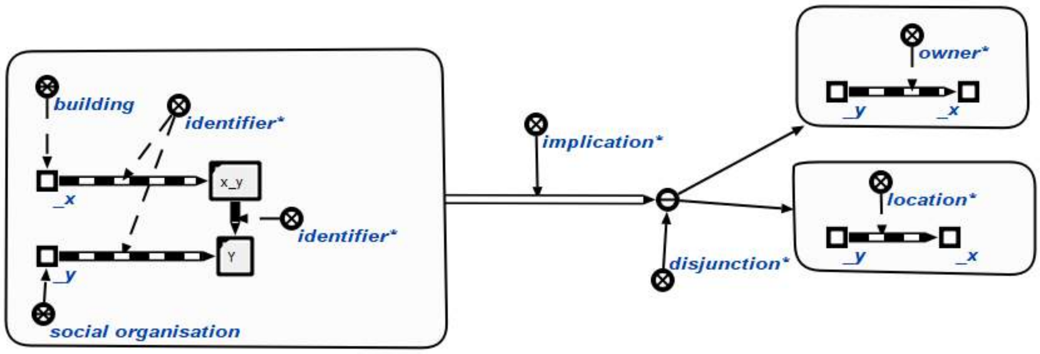


Figure 9. The logical rule found in the knowledge base

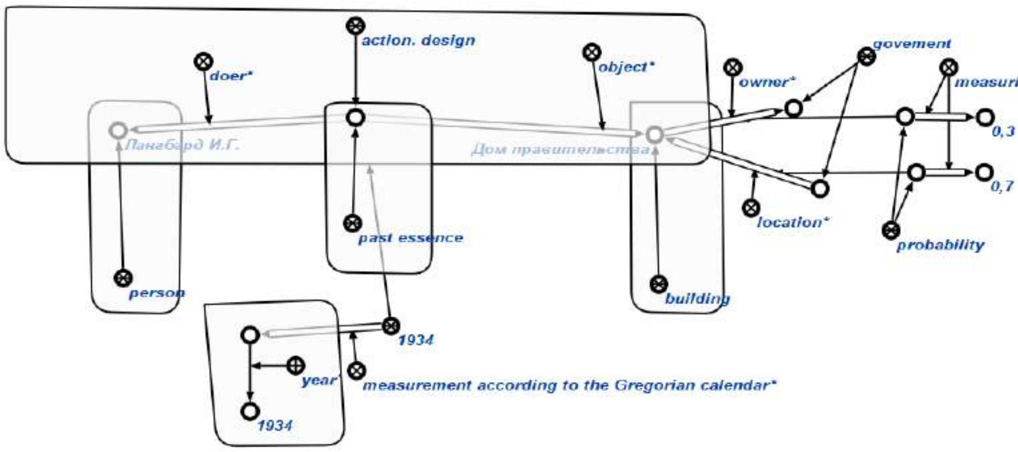


Figure 10. Functioning of the context analysis agent

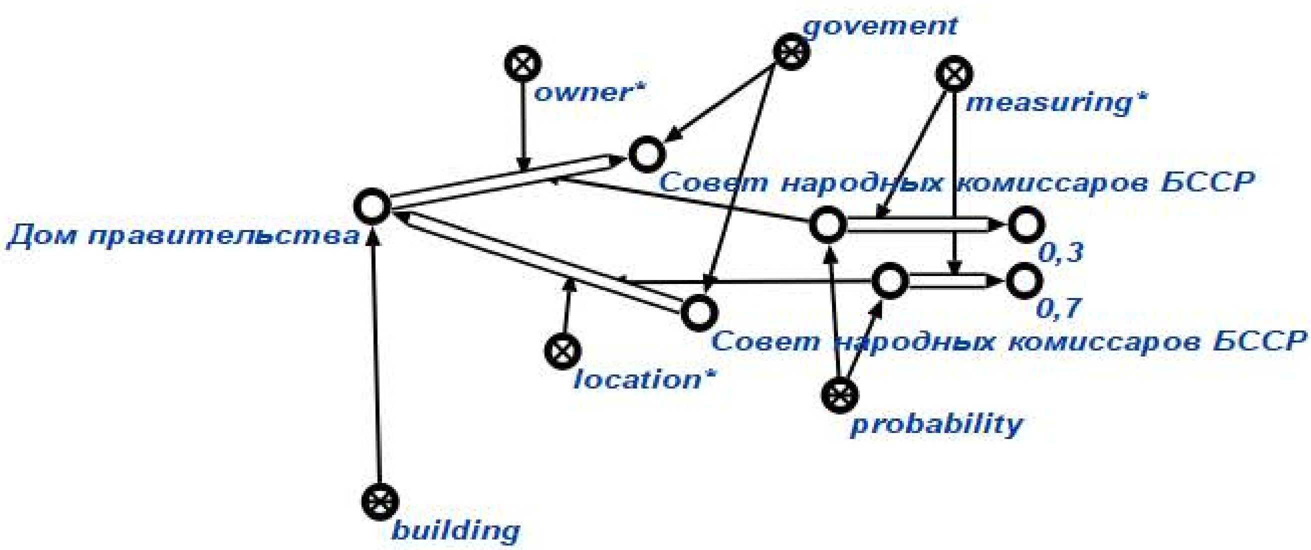


Figure 11. Synonymic elements after the agent functioning

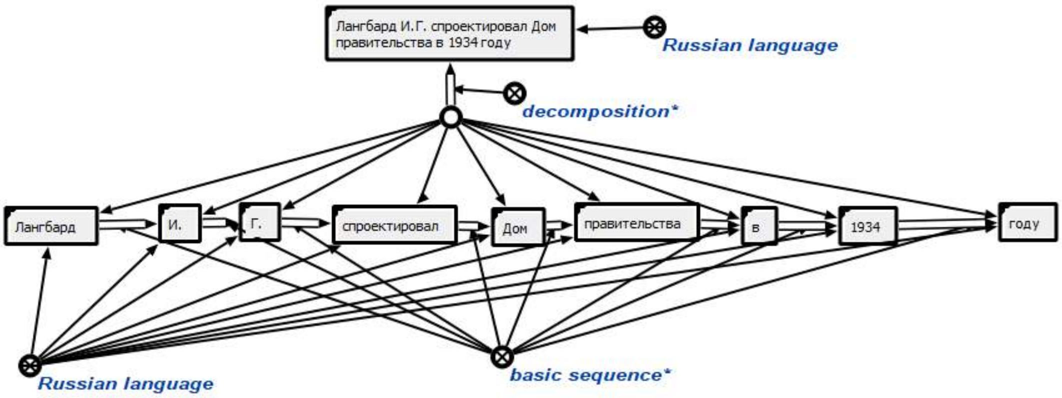


Figure 12. Graphematic analysis of the input sentence

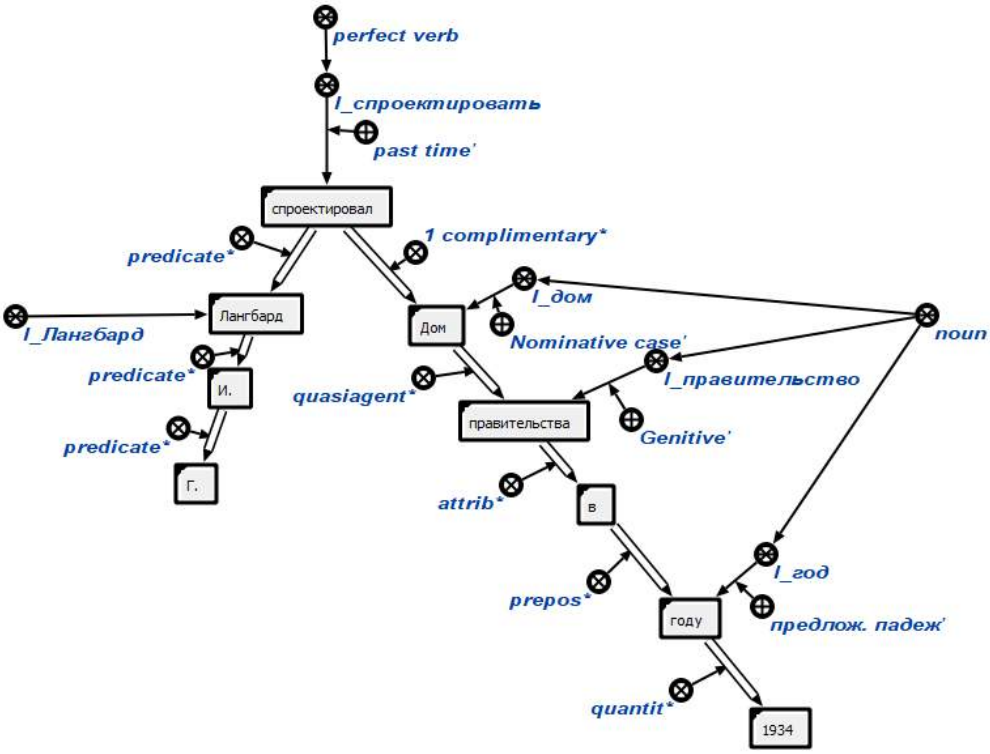


Figure 13. The result of translating the tree constructed by the ETAM service into a semantically equivalent structure in the knowledge base

REFERENCES

[1] Tarasov V.B. The Problem of Understanding: The Present and Future of Artificial Intelligence. Open semantic technologies for intelligent systems, 2015, pp. 25-42.

[2] Pospelov D.A. Ten hot spots in research on artificial intelligence./ Intellectual systems, 1996, vol. 1,No 1-4, pp. 47-56.

[3] Zolotova GA, Onipenko NK, Sidorova M. Yu. Communicative grammar of the Russian language. Moscow, Nauka, 2004. 544 p/

[4] Gladkiy A.V. Formal grammars and languages. Moscow, Nauka, 1973. 386 p.

[5] Gildea D., Jurafsky D. Automatic labelling of semantic roles. Comput. Linguist, 2002, Vol. 28, no. 3, pp. 245–288.

[6] Bharati A., Venkatapathy S., Reddy P. Inferring semantic roles using sub-categorisation frames and maximum entropy model. Proceedings of the Ninth Conference on Computational Natural Language Learning, 2005, pp. 165–168.

[7] De Busser R., Moens M.-F. Learning Generic Semantic Roles. Available at: http://www.academia.edu/265713/Learning_Generic_Semantic_Roles

[8] Claveau V., Sebillot P. From efficiency to portability: acquisition of semantic relations by semi-supervised machine learning. Proceedings of COLING'04, 20th International Conference on Computational Linguistics, 2004, pp. 61–267.

[9] Ichiro Yamada T. B. Automatic discovery of telic and agentive roles from corpus data. Proceeding of the 18th Pacific Asia Conference on Language, Information and Computation, 2004, pp. 115-126.

[10] Mooney R. J. Learning for semantic parsing. Computational Linguistics and Intelligent Text Processing: Proceedings of the 8th International Conference, 2007, pp. 311– 324.

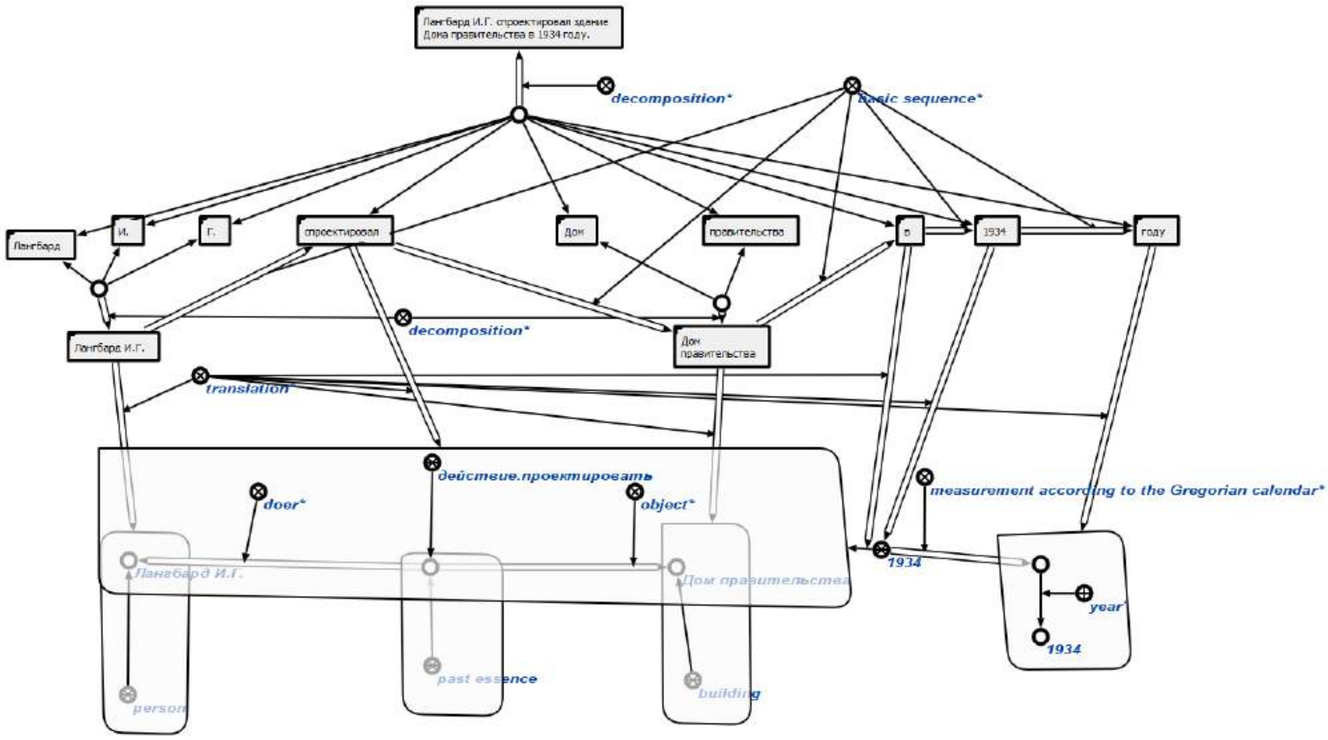


Figure 14. The result of translating the natural language text into a structure in the language of the sc-code

[11] Anisimovich K. V.,Druzhkin K. Ju Syntactic and semantic parser based on ABBYY Compreno linguistic technologies. Papers from the Annual International Conference "Dialogue" , 2012, vol. 2., pp. 91–103.
[12] Iomdin L., Petrochenkov V. ETAP parser: state of the art. Papers from the Annual International Conference "Dialogue", 2012, pp. 830–853.
[13] Anjali, M K., Babu, Anto P. 2014. Ambiguities in Natural Language. Processing International Journal of Innovative Research in Computer and Communication Engineering, 2014, Vol.2, Special Issue 5, pp. 392–394.
[14] Mozharova V.A., Lukashevich N.V. Research of signs for extracting named entities from texts in Russian. Scientific and technical information. Series 2: Information Processes and Systems, 2017, No. 5, pp. 14-21.
[15] Tonoyan-Belyaev I.A. Ellipse and the concept of conventional syntax. Bulletin of St. Petersburg University, 2007, Series 6, Issue. 3, pp. 232-236.
[16] Maksimov V.Yu., Klyshinsky ES The problem of understanding in artificial intelligence systems. New information technologies in automated systems, 2016, pp. 43-60
[17] Ivashenko V.P. Models and algorithms of knowledge integration based on homogeneous semantic network. Dokt, Diss., Minsk, 2014. 147 P.
[18] Pospelov D.A. Levels of understanding. Artificial intelligent. Book 2. Models and methods: Manual, Moskow, Radio and Communication, 1990, pp.110-115
[19] The IMS.OSTIS website. Available at: http://www.ims.ostis.net.
[20] Martynov V.V. Fundamentals of semantic coding. Experience in the representation and transformation of knowledge.Minsk, 2001. 138 p.
[21] Boyko.I.M. Semantic classification of actions for knowledge inference. Open semantic technologies for intelligent systems, 2016, pp. 115-121
[22] Hardzei A.N. Theory of the architecture automatic formation (TKAAF-2) and futher minimisation of semantic calculus. . Open semantic technologies for intelligent systems, 2014, pp.49-64
[23] Hubarevich A.V. Ontology-Based Design of Intelligent Systems in the Field of History. Open semantic technologies for intelligent systems, 2017, pp. 245-250.
[24] Davydenko I.T., Grakova N.V. Means of structuring of semantic models kazan knowledge. Open semantic technologies for intelligent systems, 2016, pp.93-107
[25] Rusetskiy K.V. Approaches to the natural language formatting in the semantic graph language of knowledge representation. Karpov Scientific Readings: a collection of scientific articles, iss. 7, part 1., pp. 255-259.
[26] Laboratory of Computational Linguistics. Available at: http://cl.iitp.ru/ru/etap3
[27] Martynov V.V. Semantic coding for representation and knowledge transformation.Moscow, Institute for Information Transmission Problems of the USSR Academy of Sciences, 16p.

ПРИОБРЕТЕНИЕ ЗНАНИЙ НА ОСНОВЕ ТЕКСТОВ ЕСТЕСТВЕННОГО ЯЗЫКА

Бойко И.М., Гордей А. Н.
Центр системного анализа и стратегических исследований НАНБ
г. Минск, Беларусь
Губаревич А.В., Семеняка А.Ф.
Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Беларусь

В статье описывается подход к машинному анализу естественно-языкового текста с последующим автоматическим наполнением базы знаний на основе технологии OSTIS. Данный подход позволяет проводить семантико-синтаксический анализ текстов с последующим анализом контекста, что достигается за счет построения онтологий конкретной предметной области.

The detection of actual research topics using co-word networks

Lande D.V., Balagura I.V, , Andrushchenko V.B.
Institute for Information Recording
of National academy of science of Ukraine
Kyiv, Ukraine
balaguraira@gmail.com

***Abstract*—The purpose of research is to compare co-word and co-author networks from Web of science and show the possibilities of scientific priorities' searching for scientific teams in computer science. The paper proposed a new algorithms for detecting actual research topics, which is calculated on the analysis of co-word and co-author networks. We offer to compare co-words networks of certain team and the main world's tendencies. The algorithm is available to search common topics for international researches of scientific teams and compares co-words networks of certain team and worlds co-word network. The algorithm determines the scientific topics that are relevant to global trends, as well as the scientific clusters that perform them. The examples of calculation of the algorithm for computer science for authors from Belarus have been proposed. The results of the research can be used to determine the scientific topics of research to address the urgent needs of any country, to correct the research of organizations.**

***Keywords*—scientometrics; co-author networks; co-word network; actual research topics;computer science; Belarus.**

I. Introduction

Researchers are always searching for the most actual problems of our society and trying to solve them. There are well-known international science resources, which are used for searching and publication main scientific results . Arrangement of publications by their significance on certain topic is very helpful in such databases. There are numbers of rating systems world-wide which are being used for publications, journals and scientists ranking. Citing indexes, eigen-factor and others are widely used for the scientific work estimation and all of them are based on the amount of a publication citing [1]. Unfortunately there are some papers, that aren't popular and aren't cited. The main point of the popularity is the actuality of topic in the world in the right moment. So, the actual research topic identification is very important for any scientific work. Project "State of Innovation" defines the most current innovative trends and the most innovative companies in the world [2]. The researchers are determined the dynamics of growth in the number of publications in the database Web of Science for assessing the origin of innovation and database Derwent World Patents Index for assessing the innovative capacity and so on. The global leaders in the number of patent research and development are China and South Korea according to the reports. For example, in information technology computing is by far the most active subsector, comprising 83 percent of IT's overall activity. Development and innovation could be identified with new research areas. The methods are based on the study of citations networks and abstracts of patent databases, citation time and frequency downloads. Another project of Web of science database is "research fronts", that consist the list of 100 "hot" papers (which is 0.1 percents of publications on the same industry and the same period, the most cited papers in a short period of two years) and 44 new "research fronts" (new research areas that marked a significant increase in citations of "hot" papers). There are used 10 research fields and co-citation maps. In the last report Research Fronts 2016 of Clarivate Analytics we could see the most actual themes in mathematics, computer science and engineering. It is mainly focus on hesitant fuzzy sets, configuration design and heat transfer analysis, the Keller-Segel chemotaxis model, partial differential equations, cloud manufacturing, the internet of things, multiple-input multiple-output systems, measurement-device-independent quantum key distribution, Li-ion batteries and bio-inspired algorithms [3]. |

Most actual and new research topics may be recognized using of all described methods. But researches of some scientific teams don't match with actual topics [4]. At this time it is expedient to find main priorities of science development and choose right topics for international cooperation. Research teams already have some scientific heritage and haven't to change all topics.

The purpose of research is to compare co-word and co-author networks from Web of science and show the possibilities of scientific priorities' searching for scientific teams in computer science. We propose to use the methods of co-author and co-word network analyzes on the base of international resources. It will allow to analyze main key-words and topics of Ukrainian researches and to find main points, which is most close to most actual international researches.

Co-authors network is a network structure where nodes are scientists and links are co-authorship, size of nodes and width of lines are depends of network characteristics and common paper numbers. A co-word network are built on occurred pairs of terms and shows their interconnections. According to the algorithm from paper [5] terms will be extracted using frequency characteristic in abstracts. Co-word and co-author networks could be used for identification and description of scientific groups and research topics, the most communicative

Figure 1. Co-author network

researchers and main principles of science communication. For the analysis we have to use main principles and instruments of complex networks, that are described in many works [2-9]. Co-word networks often are used to describe the rsearch topic statistically for identifying themes and trends in a subject domain [10,11].

Co-word networks research teams publications allow to find common "narrowed" line of research with a clearly defined system of concepts (terms); common terminology may differ in detail from the general in a separate sciences; reduce the noise information that facilitates the work of experts in the knowledge that forms the model domain. In [8] co-word analysis was employed to reveal patterns and trends in the research by measuring the association strength of terms representative of relevant publications.

We use bibliographic information from Web of science. As an example the analysis of publications of computer science for Belarus authors are considered. Web of science is the biggest and high quality research databases. Lots of articles dedicated to the scientometric analysis of this resource.

A. Development of the algorithm

We assume that crossing of key-words from scientific team papers with whole co-word network of research area could show topics and authors, which are the most close for international researches and their work could be adjust to higher level. And we could call them actual topics for the scientific team.

So, proposed algorithm of detection of priority scientific topics consists of next stage:

- Detecting of the area of researching.
- Co-author networks building and calculating of main characteristics, extracting scientific teams and the most communicative authors of country [12, 13].
- Co-word networks building and main words groups, matching with technical dictionaries [14].
- Co-word networks building for all researching area.
- Calculating of $(A \cap B)$ for co-word networks.
- Searching of $(A \cap B)$ in co-word networks of our teams and matching with co-author networks, searching the teams with the crossing of words from the international networks.
- Describing research area of possible cooperation with foreign partners, building international co-author networks, searching for partners.

B. Co-author and co-word network analysis

Web of science is the main scene for the international representation of researches produced by Clarivate Analytics company. About 50000 Belarus papers from the Core Collection since 1975 were indexed by database. Optics,

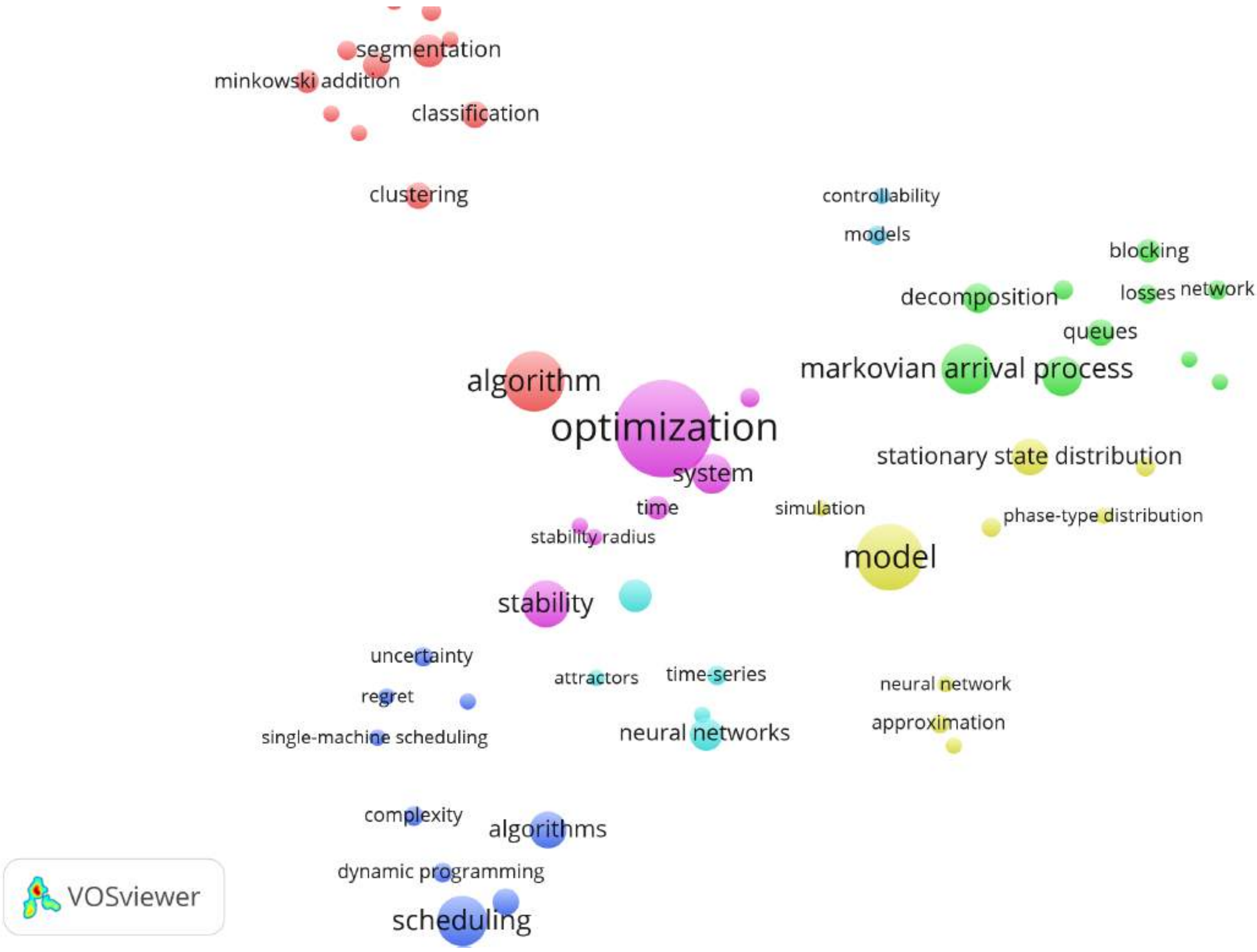


Figure 2. Co-word network

applied physics, material science, engineering are main areas of the publications from Belarus in Web of science. Web of science contains about 964 documents of computer science with Belarus scientists in 1997-2017 years. This part was analyzed with VOSviewer and Pajec tools. The co-author, co-word, co-citation networks was build. (Fig. 1, Fig.2).

According to proposed algorithm:

- We selected The computer science in Web of science data base.
- We chose the country Belarus, so have to build co-author network and search for clusters (Fig.1). Input data file was .txt format.
- Co-word network of the country team consist of 2357 key-words and shown on Fig.2 The distance between the nodes is inversely proportional to the number of coincidental occurrences of terms. Main terms according to clusters are optimization, stability, system; algorithm, clustering, segmentation, classification; model, stationary state distribution, neural networks, approximation, simulation; markovian arival process, decomposition, blocking and others.
- Co-word networks building for all researching area. About 20 thousand words were used: complex networks, neural networks, quantum computers, recognition, statistical analysis, big data, simulation and others. Most papers dedicated to interdisciplinary applications.
- We used own software and detect 100 common words for Belarus and world publications. Among them: neural networks, algorithm, optimization, clustering, system, process, programming models and others.
- Last steps could be used by experts for describing actual research topics according to key-words. For more detail description we could use descriptions of main scientific teams from co-author networks and technical classifications. Also the co-word map we could use for the base of ontology for different scientific area.

Previously in [15], the testing of the algorithm was presented on the example of the analysis of the term networks on scientific publications in physics in the Scopus database of the research team of the Kiev National University Taras Shevchenko. Scopus contains more Ukrainian and Belorussian articles than Web of science. And there are more possibilities for searching common research topics among separate organizations. So for complex analysis of actual topics of the country it will be useful to work with all databases, which are available. The results of scientometric analysis is depends on input data and it is necessary to choose the most full database for detecting future collaboration. For example, if we choose searching for Ukrainian and Belorussian cooperation we could use the biggest national resources and transform terms to one language. Also widely used in Ukraine and Belarus are Google Schoolar, Arxiv, Index Copernicus, РИНЦ and even such resources as Research Gate. But Web of science and Scopus could be used for the results revising as the most authoritative resources for international collaboration.

II. Conclusion

The algorithms, which based on co-word and co-author analysis for detecting actual topics and priorities, are proposed. We offer to compare co-words networks of certain team and the main world's tendencies. The algorithm is for activation of international research and popularity of papers witch is depend on topics. Using of the algorithm could help scientists to get more authority in the world. The results of the research can be used to determine the scientific topics of research to address the urgent needs of any country, to correct the research of organizations. The co-author and co-word networks of research output for computer science from Belarus were built, which shown the possibility for detecting the most popular research topics. Detected key-words are neural networks, algorithm, optimization, clastering, system, process, programing models and others. A further research need to use and compare several databases for increasing number of papers and searching scientific teams.

References

[1] Milojevic S., Radicchi F., Bar-IIan J. " Citation success index – An intuitive pair-wise journal comparison metric ", Journal of Informetrics, vol. 11, issue 1, pp. 223-231, 2017.

[2] " State of Innovation Report 2017 ", Clarivate Analytics, 143 p., 2017.

[3] " Research fronts 2016 ", Institutes of Science and Development, Chinese Academy of Sciences, The National Science Library, Chinese Academy of Sciences, Clarivate Analytics, 100 p., 2016.

[4] Newman M., Strogatz S., Watts D. " Random graphs with arbitrary degree distributions and their applications ", Physical review E, vol. 2, issue 64, pp.116, 2001.

[5] Lande D. , Snarskii A. , Yagunova E. , Pronoza E. "Volskaya Hierarchies of Terms on the Euromaidan Events: Networks and Respondents " Perception 12th International Workshop on Natural Language Processing and Cognitive Science NLPCS 2015 Proceedings, 22-24 September 2015, pp. 127–139, materials, Krakow, Poland., 22-24 September, 2015.

[6] Albert R., Albert-Laszlo Barabasi " Statistical mechanics of complex networks", Reviews of Modern Physics, vol. 74, pp 47-97, 2002.

[7] Bolshakova H. , Vorontsov K. , Efremova N. , Klyshinskiy E., Loukachevitch N., Sapin A. " Automatic treatment texts on natural language and analysis of data: training allowance. ", in Russian, NIU HSE, 269 p. , 2017.

[8] Wiesner A., Morbach J., Marquardt W. " Information integration in chemical process engineering based on semantic technologies ", Computers & Chemical Engineering, vol. 4, issue 35, pp. 692–708, 2011.

[9] Evangelos Pafilis, Sune P. Frankild, Julia Schnetzer, Lucia Fanini, Sarah Faulwetter, Christina Pavloudi, Katerina Vasileiadou, Patrick Leary, Jennifer Hammock, Katja Schulz, Cynthia Sims Parr, Christos Arvanitidis, and Lars Juhl Jensen " ENVIRONMENTS and EOL: identification of Environment Ontology terms in text and the annotation of the Encyclopedia of Life Bioinformatics " Bioinformatics (Oxford), vol. 11 , issue 31, pp. 1872-1874, 2015.

[10] Choi Jaewoo, Kang Woonsun " Themes and Trends in Korean Educational Technology Research: A Social Network Analysis of Keywords " Procedia - Social and Behavioral Sciences , vol. 131, pp.. 171–176, 2014.

[11] Xiuwen Chen, Jianming Chen, Dengsheng Wu, Yongjia Xie, Jing Li " Themes and Trends in Korean Educational Technology Research: A Social Network Analysis of Keywords "Procedia Computer Science, vol. 91, pp.. 547–555, 2016.

[12] Franceschini Fiorenzo, Maisano Domenico, Mastrogiacomo Luca " Empirical analysis and classification of database errors in Scopus and Web of Science " Journal of Informetrics, vol. 10,issue 4, pp.. 933–953, 2016.

[13] Balagura I. V., Lande D. V. " Linguistic studies of research collaboration using the «Ukrainika naukova» abstracts database ", Data recording, storage and processing, vol. 16, issue 3, pp. 45–53, 2014.

[14] Lande D. V., Andrushchenko V. B., Balagura I. V. " Formation of the subject area on the base of wikipedia service ", Open Semantic Technologies for Intelligent Systems, pp. 211–2014, materials, Minsk, February,2017.

[15] Lande D. V., Balagura I. V., Pohorilyi S. D., Dubchak N. A. "Scientometric analysis of co-word networks taken from publications of a scientific team ", Data recording, storage and processing, vol. 19, issue 1, pp. 34–39, 2017.

ОПРЕДЕЛЕНИЕ АКТУАЛЬНЫХ НАУЧНЫХ НАПРАВЛЕНИЙ С ПОМОЩЬЮ СЕТЕЙ ТЕРМИНОВ

Ландэ Д.В., Андрущенко В.Б., Балагура И.В.
Институт проблем регистрации информации
Национальной академии наук Украины
Киев, Украина

Цель исследования - сравнить сети терминов и соавторов, построенные на основе реферативной базы данных Web of science, показать возможности поиска научных приоритетов по информатике. В статье предложен алгоритм поиска актуальных направлений с помощью анализа сетей авторов и терминов. Предлагается сравнить сети терминов научных коллективов и основные мировые тенденции. Алгоритм заключается в поиске общих тем для международных исследований научных коллективов, при этом происходит сравнение сетей терминов коллектива и общей для научного направления. Выполняется определение научных тем актуальных на мировом уровне, а также научные коллективы, которые их выполняют. Рассмотрен пример использования алгоритма для авторов из Беларуси в области информатики. Результаты исследования могут быть использованы для определения перспективных научных тем в стране и отдельных организациях на основе реферативных и научных баз данных.

An approach to speech ambiguities eliminating using semantically-acoustical analysis

Zahariev V.A.
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
zahariev@bsuir.by

Azarov E.S.
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
azarov@bsuir.by

Rusetski K.V.
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
rusetski@bsuir.by

***Abstract*—An approach to the problem of elimination of ambiguities in speech messages by application of semantically-acoustical analysis is presented in this paper. Authors propose the architecture of the intelligent system that implements this principle. According to this principle the direct transition from speaking to meaning of given phrase is possible with the help of digital signal processing techniques, as well as knowledge formalization methods using semantics networks (semantically-acoustical analysis). A prototype of intelligent system to resolve speech ambiguities of a certain type (homonyms and paronyms) based on the tools provided by the OSTIS technology and GUSLY signal processing framework has been implemented. The main advantages of the proposed solution in comparison to the standard automatic speech recognition systems and possible ways of further development for natural language understanding problem are also reported in this paper.**

***Keywords*—speech processing, semantic technologies, acoustic analysis, semantic analysis, semantically-acoustical analysis, instantaneous harmonic analysis, natural language understanding, ostis-system, SC-code**

I. Introduction

Speaking is the one of the most natural and effective forms of communication between people. This fact explains the significant interest of researchers to the development and practical use of speech interfaces to provide human-machine interaction in the modern communicational, multimedia and intelligent systems [10], [9]. The majority of scientific publications in this direction is devoted to the issues of basic technologies development. This technologies are the components of the speech interface, such as text-to-speech synthesis (TTS), as well as speech-to-text (STT) recognition [4], [8].

Especial interest cause the tasks that are associated with understanding the sense of the natural linguistic message in the context of the practical use of the speech interface as a part of the intelligent systems.(NLU) [33], [18], [30]. In terms of intelligent systems, it is interesting to investigate the integration of the message's information content into the current state of the intellectual system of knowledge base for the purpose of further processing by the knowledge machine and intelligent agents.

Most modern systems of sense understanding are built on the basis of three-tier architecture, when the voice message consistently passes the stages of acoustic analysis of speech signal, then linguistic analysis. The linguistic analysis results in the textual form of the original message are presented, and only then semantic analysis of the message is performed. However, from works on psycholinguistics and cognitive psychology It is known, that processes of perception and comprehension in human consciousness proceed continuously [21], [27], and in general it is not necessary to bring speech message preliminary to a text form to perform semantic analysis of its contents. Oral and written forms of speech with equal success can be processed by sensory and cognitive systems of the person [22]. Therefore, the question of creation of approaches and methods for systems in which the direct transition from processing of the message in a speech form to the analysis of its semantic content is actual.

II. Problem statement

The classical three-tier approach (acoustic, linguistic, semantic analysis), in case of solving the problem of comprehension of speech messages, has a number of significant disadvantages:

- introduction of intermediate stage: conversion of speech signal to text, entails extra costs associated with the need of linguistic processing, thereby increasing the overall computational complexity of the algorithm.
- the presence of a text processing stage causes additional errors and distortions due to the limitations and incomplete of correspondence linguistic models to the process used to navigate to the textual representation of information on different stages of transformation (phoneme-to-morpheme, morpheme-to-word, word-to-phrase, etc.) [13]
- an approach when we translate a speech signal into text, we could lose some of the information that may be important for understanding the meaning of the message, such as volume, duration, intonation, pauses between words that may not always appear in the text clearly expressed with punctuation marks, etc. This problem is especially relevant when analyzing messages that are not complete sentences, but can be interpreted by the listener. For example, in everyday speech a sentence consisting of only a sound [ah] depending on the volume, intonation and duration of sounding can express pain, wonder, question, act as a conjunction or a particle («ah,

leave him...» - «а... оставь его», «ah, who is it?» - «а... кто это?», «ah, and if we did a different...» - «а если бы сделали по-другому...») [26].
- the translation of the sound signal into the text makes it impossible to analyze audio, which are not only speech messages, but also carry potentially important information for the system, for example:
  - information about conditional signals, issued by objects of external environment, in particular, equipment on manufacture, cars on road, etc;
  - sounds that can correspond to emergency situations or alarm signal (rumbling, clang, hiss, explosions, etc.);
  - other sounds that potentially carry information about the state of the environment of the automated system.

  The lack of this type of signal analysis greatly limits the ability of automated systems that are oriented on work in a constantly changing environment, including environment that is difficult to predict.

A. The problems in the analysis of voice messages

Until now, many problems related to the understanding of natural language, and, in particular, speech messages, remain unresolved. These problems can be divided into two groups: (1) problems caused by the properties of natural language and inherent to the understanding of natural language messages presented in any form, both textual and verbal; (2) problems inherent to the directly understanding of speech messages.

Problems of analysis and understanding of natural language texts are widely considered in the literature [3] and their full review goes beyond the scope of this paper. In more detail, consider those that will be partially resolved within the framework of the approach proposed in the work. Such problems include:
- the problem of solution for homonyms identifying [23];
- pronoun identification problem [34];
- the problem of proper names identifiaction (for example, the word «Slava» at the beginning of a sentence can mean both a diminutive form of the name «Vyacheslav», and a common noun);
- the problem of homographs (zAmok-zamOk);
- the problem of understanding the different forms of the same word (which in some cases may coincide with the forms of other words);
- the problem of understanding terms that consist of two or more words («acceleration of free fall»).

In addition, there are a number of additional problems related directly to the characteristics of the speech signal:
- the problem of paronyms (similar in sounding words) (koza-kazak-kazan-Kazan', dictant-dictat, postel'-pastel', etc.);
- the problem of proper names resolution is complicated by the lack of capitalization;
- the speech signal is more complex in terms of presentation and processing than textual information, because of the greater variability and expressiveness of oral speech compared to the written form. This fact, in particular, is connected with the rich intonational (prosodic) possibilities of oral speech. The intonation is formed by melody (change of frequency) of speech, intensity (loudness) of speech, duration, increase or slowing down the tempo [28]. A large role is played by the place of logical stress, the degree of clarity of pronunciation, the presence or absence of pauses. The speech has such an intonational variety that it can convey the whole wealth of human experiences, moods and emotions.

Obviously, the listed problems cannot be solved without an analysis of the context of the use of a particular word, while the context in general can be quite broad and go beyond one sentence.

Analysis of the current state of publications about the understanding of oral speech shows that to solve the problem of contextual recognition in the absence of a linguistic stage of processing, approaches based on popular machine learning techniques based on the neural network classification (Connectionist Temporal Classification - CTC) based on recurrent neural networks with Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) and Deep Neural Network (DNN) methods [1], [5], [16]. And in this case, we mean the recognition of individual phonemes, morphemes or words in the speech stream. However, in these works, questions of the semantic processing of such information and integration into the intellectual system are not posed and are not considered, which leaves this issue open and indicates that there are unsolved problems.

III. Semantically-acoustical analysis

In this paper, problems that were mentioned above are proposed to be solved by performing a semantically-acoustical analysis. This process involves the primary analysis of a voice message using special signal processing techniques. In the course of their application, words are isolated from the stream of individual «acoustical pattern». This acoustical pattern will correspond to certain nodes (signs of concrete entities or concepts) in the semantic network. It is assumed that the results of the acoustic analysis phase will be iteratively corrected considering the information stored in the knowledge base of the system, including the semantic analysis of context-sensitive information.

By «acoustical pattern» we mean a fragment of the speech signal, usually represented in some parametric form, which corresponds in duration to the phonetic word in the speech stream. One of the hypotheses put forward in the work is the assumption that from the point of view of formalizing the semantics of a message one can work with a signal at the level of the whole word (and not its separate parts, such as a phoneme or morpheme used in classical speech recognition systems on a linguistic basis stage of processing). And a word, in this case, is the minimal sense distinctive unit of speech, i.e. denotes the concept or specific entity represented in the knowledge base of the system.

In this work, at the current level of development of the proposed approach, the selection and analysis of words is proposed to be carried out based on recognition systems in comparison with the vocabulary of standards and further determining the degree of certainty (correspondence) of the allocated "acoustic pattern" to the reference of this pattern – the content of the node of the semantic network in the knowledge base. The use of such a measure as the degree of certainty instead of the immediate value of the likelihood of the pattern matching to the reference is due to the fact that the concept of probability in this case can not be used in the full sense correctly, because the standards presented in the dictionary do not represent a complete group of events, the sum of the probabilities of which gives unity. The measure of the degree of confidence in this case is rather a measure of the proximity of the selected signal fragment to the reference signal in the selected parametric space.

In the general case, in order to carry out this kind of analysis, the following information may be required in the proposed approach:

- a set of standards for comparison with the fragments of the speech signal and their specification;
- the context of the message being analyzed (from whom the message was received, in what external conditions, what other sounds are present on the background, etc.);
- a set of rules for the transition from fragments of a voice message to semantically equivalent constructs in the knowledge base;
- the semantic specification of the concepts that make up such constructions.

For speech analysis will be used a model based on a hybrid representation of a speech signal, which allows the most adequate representation of any fragments of the speech signal of a different nature of sound formation [15]. Vocalized and unvoiced fragments of the signal refer to separate parts of the model: periodic (harmonic) and aperiodic (noise).

Mathematically, the basic idea of the model can be formalized in the following form:

$$s(n) = h(n) + r(n), \quad n = \overline{0,..,N-1} \tag{1}$$

where $s(n)$ – input speech signal, $h(n)$ – harmonic component, $r(n)$ – noise component of the signal, n and N – current signal reference number and the total duration of the analysis fragment, respectively. The harmonic component can be represented by the following expression:

$$h(n) = \sum_{k=1}^{K} G_k(n) \sum_{c=1}^{C} A_k^c(n) \cos_k^c n + \phi_k^c(0)) \tag{2}$$

where G_k – gain coefficient on the basis of the spectral envelope, c is the number of sinusoidal signal components for each harmonic, A_k^c – instantaneous amplitude of the c-th component and k-th harmonic, f_k^c and $\phi_k^c(0)$ – frequency and initial phase of the c-th component of the k-th harmonic, e_k is the excitation signal of the k harmonic. The amplitudes A_k^C are normalized in order to provide the sum of the energy of the harmonics equal to $\sum_{c=1}^{C}[A_k^c]^2 = 1$ for k = 1,...,K.

In this case, the aperiodic component is modeled in the whole frequency band, as it is observed in the spectrum of the real speech signal [12]. This effect is achieved by applying the technique of signal analysis through synthesis and subtraction of the harmonic part from the original signal:

$$r(n) \quad = \quad \begin{cases} \max(s(n), h(n)) - h(n), & s(n) > 0 \\ \min(s(n), h(n)) - h(n), & s(n) < 0 \end{cases} \tag{3}$$

Thus, for a single frame of the signal with the number m and the length of N samples, a characteristic vector is formed, including the coefficients of the model $\mathbf{x_m} = [G_k, A_k^c, f_k^c, K, C]$. And the acoustic pattern of one word is a sequence of such characteristic vectors: $\mathbf{X} = (\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_M})^T$.

The model parameters are proposed to be estimated using the original method of instantaneous harmonic analysis, which makes it possible to significantly improve the accuracy of determining the parameters of the periodic component [19].

In contrast to the classical methods of signal analysis used in modern speech recognition systems based on the definition of the mel-cepstral coefficients (MFCC) [32], [20] or linear speech prediction (LPC) [31], the method based on instantaneous harmonic analysis (IHA) makes it possible to obtain a high temporal and frequency resolution of the signal, as well as a more precise spectral picture of the localization of energy at the appropriate frequencies. In contrast to classical methods based on a short-time Fourier transform (STFT) or the definition of the autocorrelation function of a signal on a short fragment, the method in question does not impose strict limitations connected with observance of the stationary conditions of the signal parameters on the analysis frame 1. In this case, the parameters of the harmonic model, if necessary (for example, for describing the spectral envelope) can be relatively easily converted to other presentation methods, such as classical mel-cepstral or linear prediction coefficients.

Algorithms implementing the above-described method of signal processing based on instantaneous harmonic analysis have a standard implantation in the framework of analysis and synthesis of audio signals GUSLY [2], the individual components of which will be used for the demonstration example presented below.

As a technological basis for implementing the proposed approach, OSTIS [24] Technology will be used. Systems based on OSTIS technology are called ostis-systems, respectively, the module for understanding voice messages, the prototype of which is considered in this work, will be built as a reusable component, which in future will be integrated into various ostis-systems if necessary.

As a formal basis for encoding various information in the knowledge base, the SC-code [24] is used, the texts of which (sc-texts) are written in the form of semantic networks with a basic set-theoretical interpretation. Elements of such networks are called sc-elements (sc-nodes, sc-arcs).

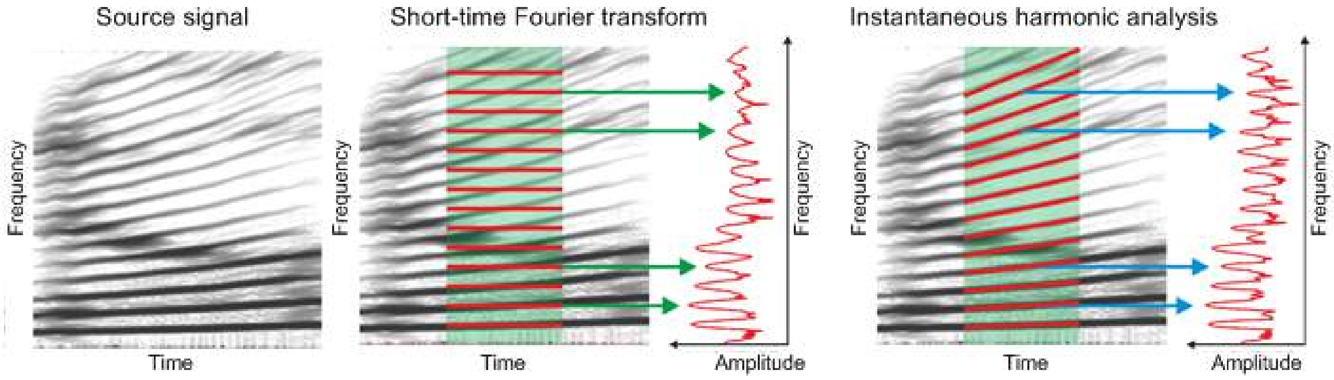


Figure 1. Comparison of signal analysis methods based on short-time Fourier transform vs instantaneous harmonic analysis.

The orientation of this work to OSTIS Technology is due to its following main advantages:

- within the framework of this technology, unified means of representing various types of knowledge, including meta-knowledge, are proposed, which allows to describe all the information necessary for analysis in one knowledge base in a unified way [6];
- used in the framework of technology formalism allows you to specify in the knowledge base not only concepts, but also any external files from the point of view of the knowledge base (for example, fragments of the speech signal), including the syntactic structure of such files;
- the approach to representation of various kinds of knowledge [6] and models of their processing [14] offered in the framework of technology provides modifiability of ostis-systems, i.e. allows to easily expand the functionality of the system, introducing new types of knowledge (new concepts systems) and new models of knowledge processing;
- the above-mentioned advantages allow to provide acoustic, syntactic and semantic analysis of messages in the same memory with the help of unified processing facilities, which in turn allows you to correct the analysis processes at any stage using different information from the knowledge base.

In its turn, the developed module for understanding voice messages is itself built as an ostis-system and has an appropriate architecture.

IV. Limitations of the proposed approach

In general, the task of understanding the meaning of a message in a language external to the ostis-system involves the following main steps:

1) Syntactic (semantic-syntactic) analysis of the message. The result of this stage is the sc-text describing the syntactic structure of the external text.
2) Semantic analysis (translation) of the message. The result of this stage is the sc-text semantically equivalent to the original message in the external language.
3) Immersion (integration) of the received sc-text into the knowledge base of the ostis-system. At this stage, there is a "glueing together" of synonymous sc-elements contained in the initial knowledge base of the system and the corresponding sc-elements included in the sc text, which is the result of the translation of the original message [25].
4) Alignment of concept systems. At this stage, the system of concepts used in the analyzed message is brought to the system of concepts accepted as the main one in the knowledge base of the ostis-system.
5) The analysis of the value of the information received, the logical conclusion. At this stage, new information is generated based on the logical inference mechanisms and existing in the system.

Some or all of the listed steps can be performed iteratively, taking into consideration the context, which is generally formed both using information from the knowledge base and information obtained from the external environment, including, explicitly requested from the user.

The focus of this work is on the first two of these stages of understanding, while the task of understanding speech messages is refined with several requirements that allow focusing attention on solving the problems listed above. Let us enumerate the specified requirements:

- voice messages do not contain noises, the words in the phrase are pronounced by the speaker, so that the pauses are clearly distinguishable in the speech signal, since the task does not set the construction of a robust algorithm for extracting words from an arbitrary speech signal;
- it is assumed that the analyzed speech message and the standards of the fragments of the speech signal contained in the knowledge base are recorded by the same speaker, e.g., those, in the speaker-dependent mode;
- only those speech messages that contain phrases of the form "subject-predicate-object" are analyzed. A predicate may be an action performed by a subject with respect to an object, or some relation connecting a subject and object. It is assumed that the components of the phrase follow exactly in this order, i.e. the first in order always

follows the subject, the last is the object, although in the natural language, in the general case, an inverse sequence is possible. It is believed that each of the components of the phrase is called with one word;
- it is assumed that all concepts used within the message are known to the system (specified in the knowledge base) and are the basic concepts, thus there is no need to align systems of concepts;
- the system of concepts used within the message does not change during the analysis, i.e. a priori, the information stored in the knowledge base is considered to be truth, and the meaning of the message is specified considering this information;
- it is assumed that there is a dictionary of lexemes corresponding to the words used in the analyzed phrase in the knowledge base, so all words are known to the system. In this case, each token corresponds to a set of fragments of the speech signal (standards) describing all the word forms of the corresponding word, thus morphological analysis is not carried out;
- it is assumed that the analyzed message is correct from the point of view of the current state of the knowledge base.

A. *Test cases*

In view of the problems and limitations discussed above, a number of test phrases has been selected and these phrases will be used to demonstrate the proposed approach to understanding voice messages (for phonetic transcription, the X-SAMPA notation will be used [17]):

1) «The goat ate cabbage» - «Коза съела капусту» - [kaza] [sj'Ela] [kapustu]
2) «Grozny captured Kazan» - «Грозный захватил Казань» - [grozn1j'] [zaxvat'il] [kazan']
3) «Kazan is located on the Kaban (Kaban - in this case the lake)» - «Казань расположена на Кабане (Кабан - в данном случае озеро)» - [kazan'] [raspaloz'Ena] [na] [kaban'E]
4) «Boar is the ancestor of a pig» - «Кабан – предок свиньи»- [kaban] [pr'Edak] [sv'in'i]
5) «Grozny is the capital of Chechnya» - «Грозный – столица Чечни» - [grozn1j'] [stal'itsa] [ts /Etsn/h'i]
6) «Shepherd bought a goat» - «Чабан купил козу» - [ts/aban] [kup'il] [kazu]
7) «Shepherd bought kazan» - «Чабан купил казан» - [ts/aban] [kup'il] [kazan]

Examples of correct constructions semantically equivalent to some of these examples are shown in the figures 2–4:

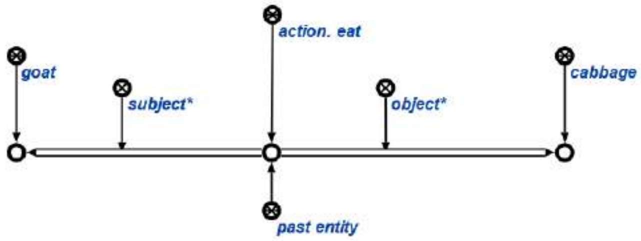


Figure 2. A construction equivalent to a phrase «The goat ate cabbage» - «Коза съела капусту».

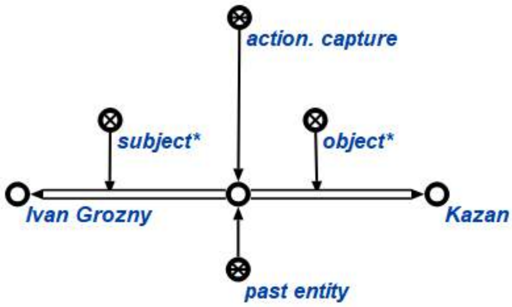


Figure 3. A construction equivalent to a phrase «Grozny captured Kazan» - «Грозный захватил Казань».

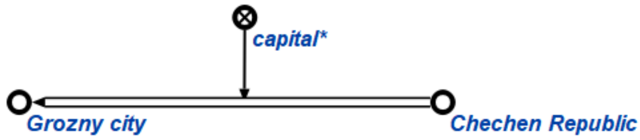


Figure 4. A construction equivalent to a phrase «Grozny - the capital of Chechnya» - «Грозный – столица Чечни».

V. Algorithm for analyzing the voice message

In general, the proposed algorithm for understanding speech messages with elimination of ambiguity includes the following steps:

1) Selection of the individual words (fragments of the speech signal between pauses) in the original speech message. The result of this step is the specification in the knowledge base of the syntactic structure of the message being analyzed, which, given the previously mentioned limitations in this work, is given by a set of words and the order of their following.
2) The received words are compared with the standards stored in the knowledge base and corresponding to some lexemes. For each of the standards, the degree of confidence (from 0 to 1) is calculated in that the analyzed fragment of the speech signal and the reference coincide. Coincidence, the degree of confidence for which below a certain threshold is discarded, the rest are fixed in the knowledge base.
3) Considering the received confidence levels, those pairs «signal fragment» - «reference» are chosen, for which

the confidence levels are maximal and their translation into semantically equivalent sc-text is carried out. Broadcasting in the general case can be carried out both with the help of a universal mechanism based on implicative rules, and with the use of specialized translation agents, each of which is oriented only to the construction of a certain type.

4) The received sc-text is verified by the existing means of the knowledge base verification in the system. In case of contradictions, return to step 3 is performed, and the following degree of confidence of the pair «signal fragment» - «reference» is chosen. In addition, it is considered in which fragments of the received sc-text a contradiction has arisen, and first of all those fragments of the speech signal that correspond to the specified fragments of the sc-text are taken into account.
   If contradictions are not revealed, then the conclusion is made that the received sc-text does not contradict the current system of concepts and is integrated into the knowledge base.

Below is the application of this algorithm to the example of a specific phrase.

VI. The architecture of the speech message comprehension software module

The architecture of the software module for understanding voice messages is shown on the figure 5.

The system consists of two main components: the acoustic component and the semantic analysis component. In its turn, the first component includes a signal analyzer, a module that implements the patter matching algorithm with a reference, a knowledge base, a configuration file, and initial parameters of the signal representation model (length and type of the analysis window, sampling and oversampling frequencies, etc. constants included in the configuration of the signal representation model).

The speech signal is fed to the input of the analysis module, where the procedures for dividing the signal into frames with a duration of 50 msec with 25% overlap are performed, the signal frames are weighted by multiplying the current signal fragment by the Hamming window, and the pitch frequency is searched. Next, the parameters of the signal model are estimated and a characteristic vector $\mathbf{x_m}$ is formed for the current frame, which is placed in a sequence of similar vectors $\mathbf{X}$, characterizing the entire word. Since the number of analysis frames, and accordingly the size of the sequence of characteristic vectors, will fluctuate depending on the duration of the word, it is necessary to perform the procedure for normalizing the number of vectors in the sequence. For all words, the size of the sequence is reduced to 10 vectors by applying the vector quantization procedure over the sequence [7].

Further in the comparison module, the obtained normalized sequence is compared with a number of standards stored in the database, which are a collection of the same sequences, but prepared and written to the database in advance. The comparison is performed using the dynamic time warping (DTW) procedure, based on the dynamic programming algorithm for multidimensional data [11]. For the three most appropriate standards, the degree of assurance of the pattern's compliance with the reference is calculated, the values obtained are recorded in the contents of nodes of the semantic network, which are appropriately specified in the knowledge base. Further, the semantic processing is carried out on the basis of information from the knowledge base, as mentioned above.

A. The knowledge base of the speech interface

To implement the semantically-syntactical analysis in the knowledge base, lexemes (sets of word forms) are specified that correspond to words that will be used in the framework of the speech message. Examples of the specification of lexemes (Figures 6-9):

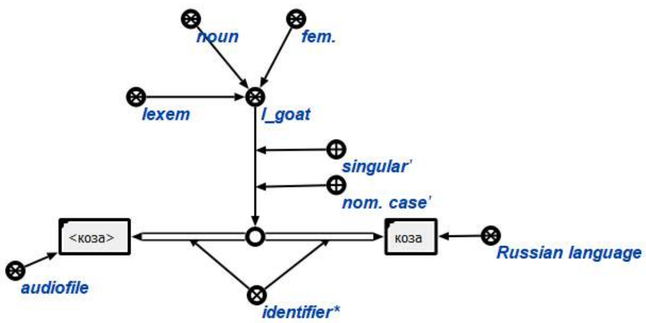


Figure 6. Lexeme «goat» - «коза».

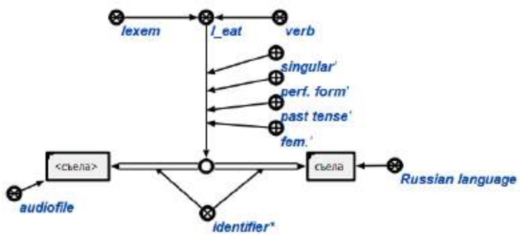


Figure 7. Lexeme «eat» - «есть».

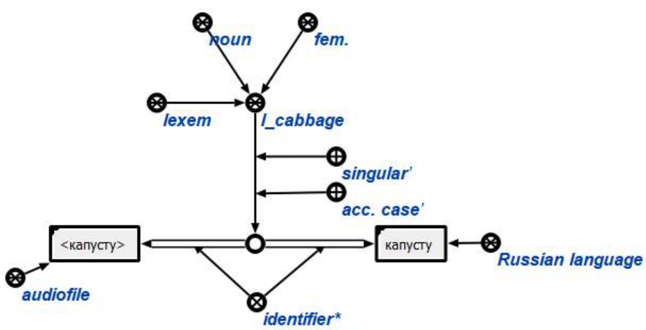


Figure 8. Lexeme «cabbage» - «капуста».

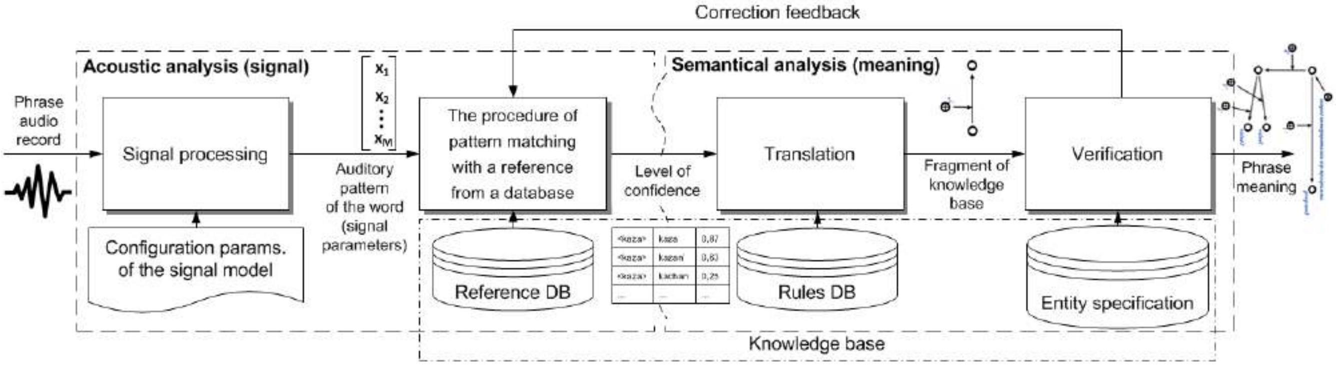


Figure 5. System architecture.

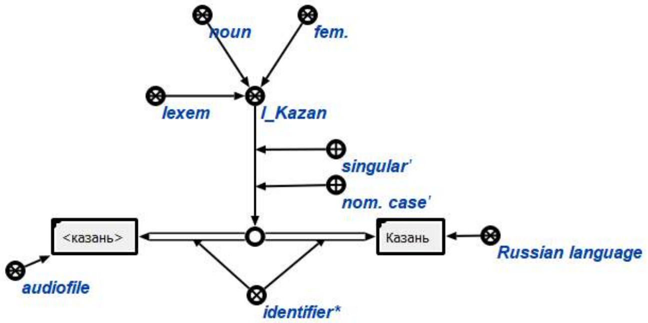


Figure 9. Lexeme «Kazan» - «Казань».

In the examples considered, the angular brackets (< goat> - <коза>) conditionally indicate fragments of the speech signal stored in the knowledge base as standards with which the comparison will be carried out at the analysis stage.

In addition, the knowledge base contains a specification of concepts corresponding to lexemes, which is used in the analysis process to verify the result of translation. The specified specification can include, for example, specifying the relationship domains (figure 10) or specifying the class of possible subjects and objects for the action class (figure 11).

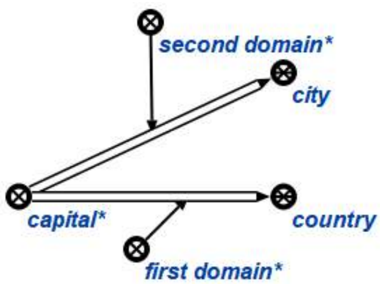


Figure 10. Relationship specification «capital*» - «столица*».

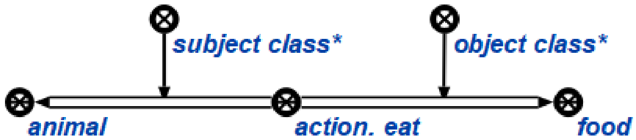


Figure 11. Activity class specification «eat» - «есть».

B. *The knowledge processing machine of the speech interface*

The knowledge processing machine for the speech interface according to the considered analysis algorithm currently has the following structure:

The speech interface knowledge processing machine
<=*abstract sc-agent decomposition**:
{
- *Abstract sc-agent for audio preprocessing*
- *Abstract sc-agent for recognizing audio fragments*
- *Abstract sc-agent for generating the translation task*
- *Abstract sc-agent for translating external files to the knowledge base*
- *Abstract sc-agent of knowledge base verification*
  <=*abstract sc-agent decomposition**:
  {
  - *Abstract sc-agent for checking the matching of bindings to its domains*
  - *Abstract sc-agent for checks the compliance of the activity specification with its class*

  }

}

VII. EXAMPLE OF THE ALGORITHM FOR UNDERSTANDING THE VOICE MESSAGE

As an example, consider the process of analyzing the phrase "the goat ate cabbage". In this example, double angular brackets (<<goat>> - <<коза>>) conditionally identify fragments of the analyzed speech signal. The above illustrations are recorded using one of the SC-code visualization variant, the SCg [29] language.

Step 1. In the first step, the original speech message is decomposed into separate words, the decomposition result is fixed in the knowledge base, as shown in the figure 12.

Step 2. The received fragments of the speech signal are compared with the standards, the result of the comparison is fixed in the knowledge base, as shown in the figure 13. In this example, we will assume that the analyzed word <<goat>> - <<коза>> coincides with the reference <goat> −<коза> with a confidence level of 0.55 and a reference <kazan> - <казань> with a confidence level of 0.65.

Step 3. Considering the confidence levels obtained, those pairs « signal fragment » - « reference » are chosen for which the confidence levels are maximal and their translation into semantically equivalent sc-text is carried out. In the example under consideration, the reference <kazan> - <казан> was first chosen, since the confidence level for it turned out to be larger (figure 14). Thus, the meaning of the phrase being analyzed is interpreted as «Kazan ate cabbage».

Step 4. For the obtained sc-text, the sc-agent for the verification of knowledge bases is initiated (figure 15), which is based on information from the knowledge base (figure 16) that Kazan is a city, not animal, and only animals can eat, so the agent concludes that the resulting structure is incorrect (figure 17).

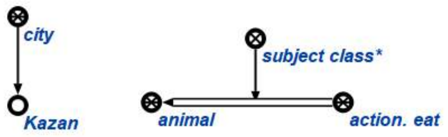


Figure 16. Verification context

Step 5. Since the received sc-text was incorrect, the translation result is deleted and a return to step 3 is made, where the next set of pairs «signal fragment»-«reference» is translated (figure 18) and verified (figure 19). In the example under consideration, the resulting sc-text turned out to be correct within the current state of the knowledge base, otherwise the translation and verification steps would be repeated.

VIII. Conclusion

An approach to the problem of eliminating ambiguities such as homonymy and paronymy in a speech signal with the use of semantically-acoustical analysis for speech understanding systems is proposed. This approach involves using the knowledge base to verify the results of message translation into the internal representation of an intelligent system, with subsequent correction of the results of recognizing fragments of the speech signal. Its main feature is the fact that it suggests the transition to semantic processing by passing the stage of processing textual information, the presence of which is a characteristic feature of all modern solutions to understanding speech. This implementation let to avoid the loss of part of the information and reduces the number of errors introduced at the text processing stage due to the imperfection of linguistic models. For the analysis and parametrization of the signal, a hybrid model is used. This model is based on the hybrid representation of the speech signal, which allows the most appropriate representation of any fragments of the speech signal of a different nature of sound formation, both vocalized and unvoiced. The method of instantaneous harmonic analysis is used to estimate the parameters of the model, which makes it possible to significantly improve the accuracy of determining the parameters of the periodic component.

For semantic analysis, OSTIS technology is used, which provides unified means for representing various types of knowledge, including meta-knowledge, which makes it possible to store and process all necessary information in one knowledge base in a unified way. In addition, technology allows you to specify in the knowledge base not only concepts, but also any external forms of knowledge representation, for example acoustic patterns of words, also allows you to easily expand the functionality of the system by introducing new types of knowledge and new models of knowledge processing. This technology provides for the modifiability of ostis-systems, i.e. allows you to easily expand the functionality of the system, introducing new types of knowledge and new models of knowledge processing.

The main differences of this work from the existing works in the field of understanding of voice messages and elimination of ambiguities in such messages include the following:

- the paper proposes an original approach that assumes consistent use of acoustic and semantic analysis, which allows to take into account the context at different stages of the speech message understanding and to adjust the results of each stage with use of the context;
- the proposed approach, unlike modern methods of speech recognition and understanding, excludes the need for an intermediate stage of the message presentation in text form, which allows to expand the context of the message analysis taking into account various parameters of the voice message (loudness, intonation, emotional coloring, etc.), and also allows in the future to analyze messages that do not have a unique text equivalent (sounds of the environment);
- the means of knowledge representation and processing used in the approach provide an ability to enhance the functionality of the system and the quality of understanding easily, including by specifying within the knowledge base of different types of context and their subsequent use in the analysis process, as well as expanding of intermediate information verification means at different stages analysis;
- the proposed approach to eliminating ambiguities in voice messages is part of the solution of a more general problem related to the learning and self-learning of intelligent systems by understanding of information obtained from various external sources, including speech and sound.

Further work will be focused on development, improving and expanding the proposed approach for a more general case, detailed comparative studies with existing systems of

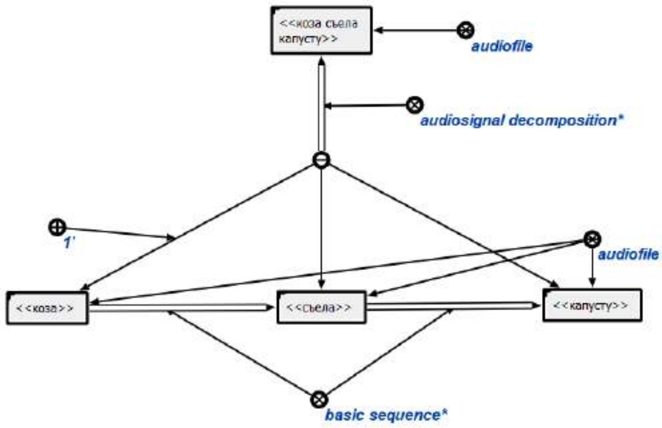


Figure 12. Analysis and parameterization of the speech signal

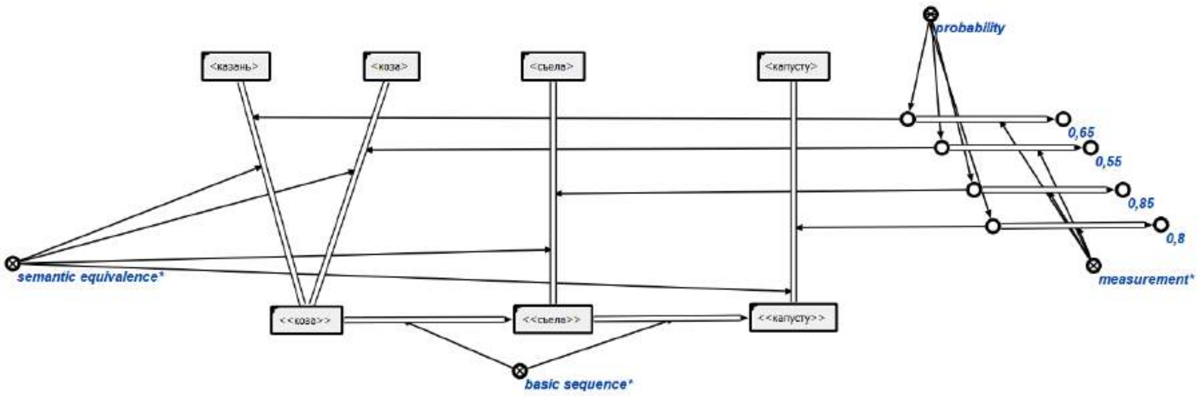


Figure 13. The result of matching the words of the original message with the standards from the knowledge base

speech recognition and understanding. Future improvements of the proposed approach should be in closer integration with approaches in the field of signal processing, psychoacoustics, psychosemantics and artificial intelligence to solve problems associated with pattern recognition and the formalization of meaning from information sources of any kind.

REFERENCES

[1] Audhkhasi, K. et al., Direct Acoustics-to-Word Models for English Conversational Speech Recognition. Proceedings Interspeech 2017, pp. 959–963.

[2] Azarov, E., Vashkevich, M., Petrovsky, A. Instantaneous harmonic representation of speech using multicomponent sinusoidal excitation. INTERSPEECH 2013: proceedings of 12th Annual Conference of the International Speech, Lyon, France, 2013. P. 1697–1701.

[3] Barr A. Natural language understanding. AIMagazine. 1980. vol. 1. N. 1. P. 5.

[4] Bellegarda, J. R. Spoken language understanding for natural interaction: The siri experience. Natural Interaction with Robots, Knowbots and Smartphones. Springer, New York, 2014. P. 3–14.

[5] Corona, R., Thomason, J., Mooney, R. Improving Black-box Speech Recognition using Semantic Parsing. Proceedings of the Eighth International Joint Conference on Natural Language Processing. 2017. vol. 2. P. 122–127.

[6] Davydenko, I. T. Ontology-based knowledge base design. Open semantic technologies for intelligent systems. Minsk: BSUIR, 2017. P. 57–72.

[7] Gersho, A., Gray, R.M. Vector quantization and signal compression. Springer Science & Business Media, 2012. vol. 159. 732 p.

[8] Lin, K. An Examination of Natural Language Processing as an Alternative to Traditional UI's for User Interaction with Applications. – 2017.

[9] MacTear, M., Callejas, Z., Griol, D. The Conversational Interface: Talking to Smart Devices. Springer, 2016. 422p.

[10] Pearl, C. Designing Voice User Interfaces: Principles of Conversational Experiences. O'Reilly Media, 2016. 287 p.

[11] Rabiner, L. R., Rosenberg, A. E., Levinson, S. E. Considerations in Dynamic Time Warping Algorithms for Discrete Word Recognition. IEEE Trans. Acoust. Speech Signal Processing, 1978. Vol. ASSP-26, N. 6. P. 575–582.

[12] Serra, X. A system for sound analysis / transformation / synthesis based on deterministic plus stochastic decomposition. PhD thesis. Stanford, 1989. 178 p.

[13] Schmitt, A., Wolfgang, M. Towards Adaptive Spoken Dialog Systems. Springer, 2012. 251 p.

[14] Shunkevich, D. V. Ontology-based design of knowledge processing machines. Open semantic technologies for intelligent systems. Minsk: BSUIR, 2017. P. 73–94.

[15] Stylinou, Y. Applying harmonic plus noise model in concatenative speech synthesis. IEEE Trans. on Speech and Audio Processing, 2001. Vol. 9, N. 1. P. 21–29.

[16] Tang, H. et al. End-to-End Neural Segmental Models for Speech Recognition. IEEE Journal of Selected Topics in Signal Processing, 2017. vol. 11. N. 8. P. 1254–1264.

[17] Wells, J.C. Computer-coding the IPA: a proposed extension of SAMPA. Revised draft, 1995. vol. 4. N. 28. P.19-25.

[18] Winograd, T. Understanding natural language. Cognitive psychology. 1972. vol. 3, N. 1. pp. 1–191.

[19] Azarov, I.S., Petrovskii, A.A. Mgnovennyi garmonicheskii analiz: obrabotka zvukovykh i rechevykh signalov v sistemakh mul'timedia

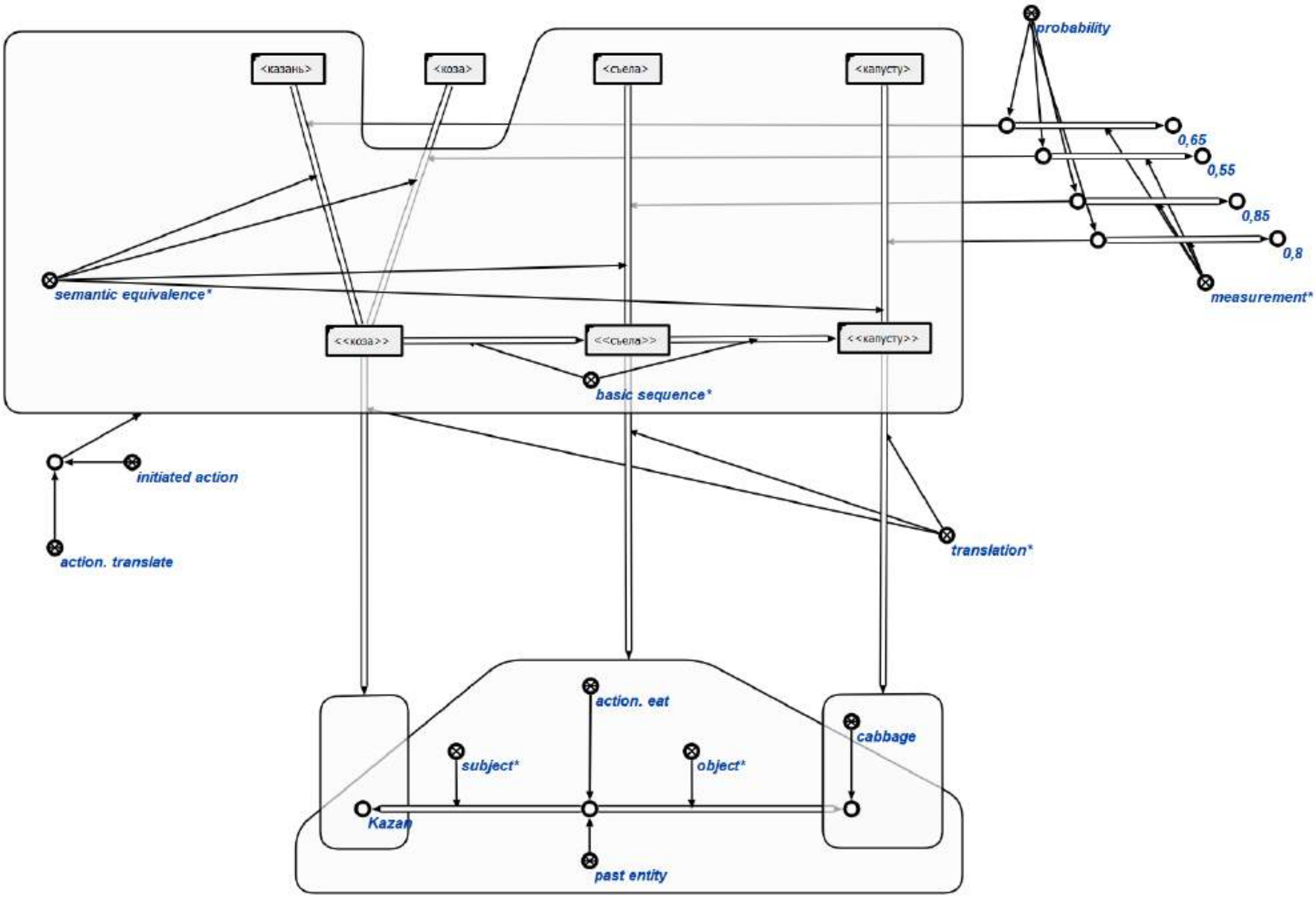


Figure 14. Translation result

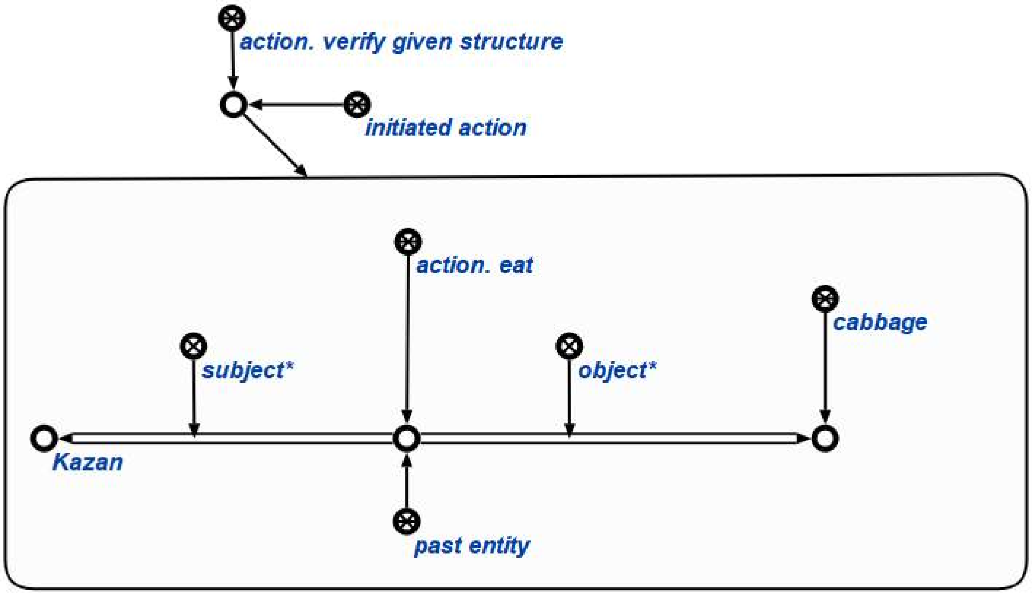


Figure 15. Assignment for verification

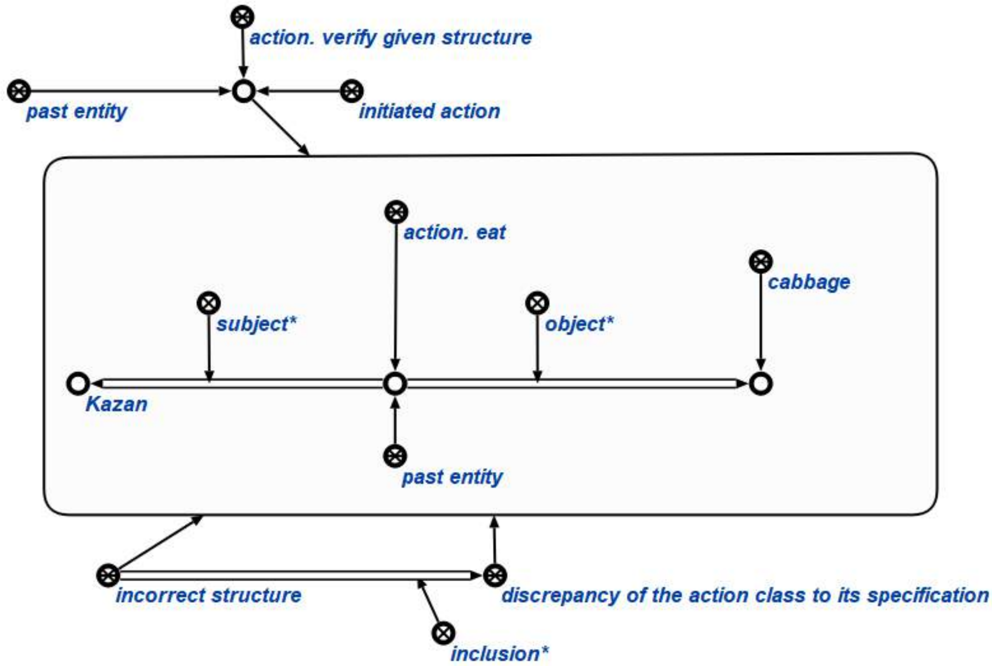


Figure 17. Verification result

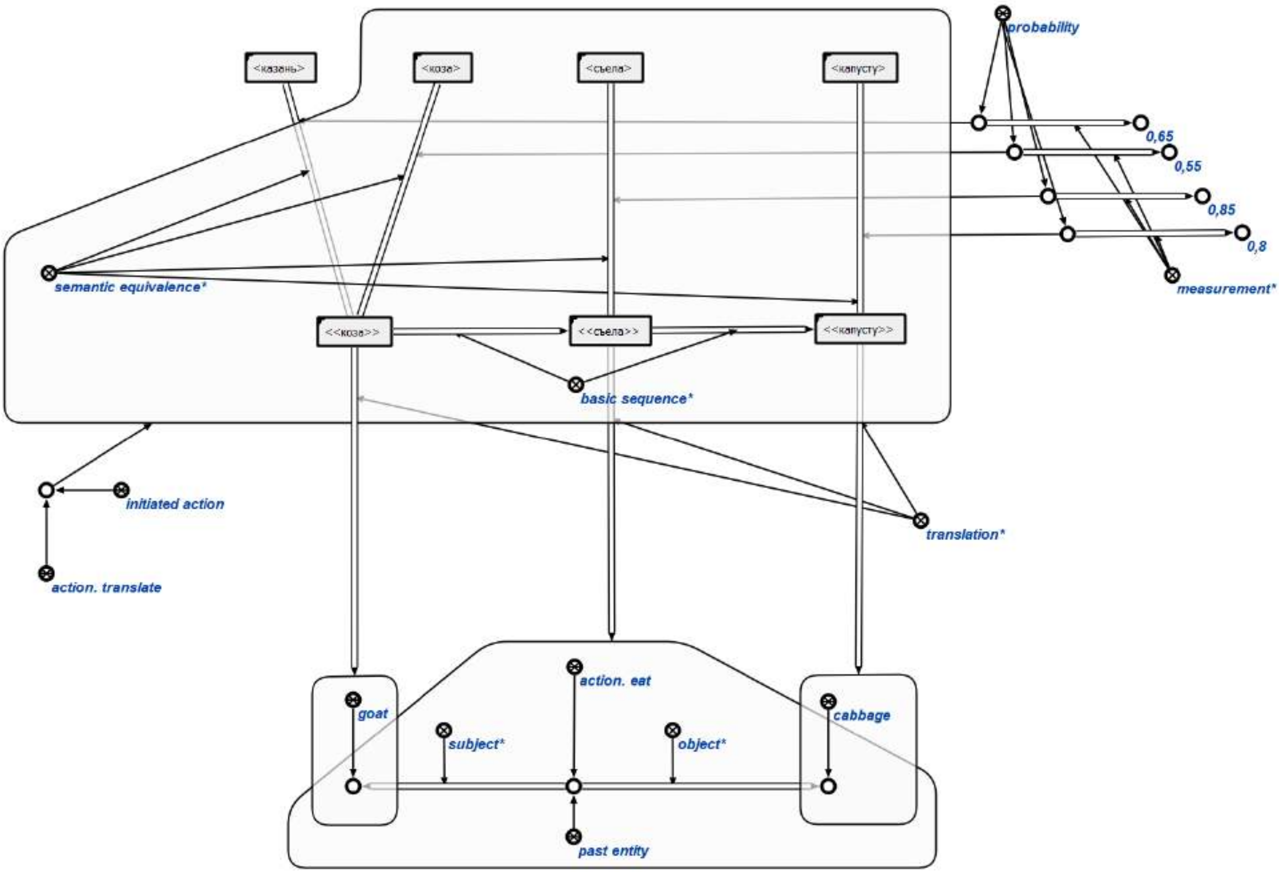


Figure 18. Repeated translation result

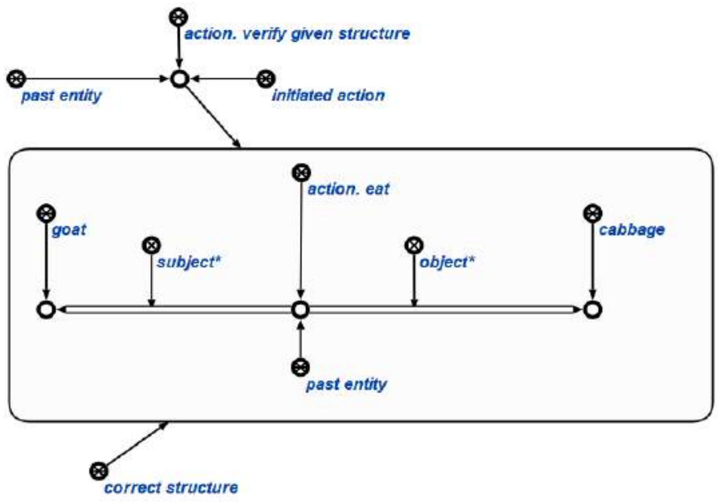


Figure 19. Repeated verification result

[Instant harmonic analysis: processing of sound and speech signals in multimedia systems]. LAP Lambert Academic Publishing, Saarbrucken. 2011. 163 p. (in Russian)

[20] Aificher, E., Dzhervis, B. Tsifrovaya obrabotka signalov: prakticheskii podkhod. 2-e izd. [Digital signal processing: a practical approach. 2nd ed.]. M.: Williams, 2004. 992 p. (in Russian)

[21] Vygotskii, L. C. Myshlenie i rech'. Psikhologicheskie issledovaniya. [Thinking and speaking. Psychological research]. Natsional'noe obrazovanie, 2015. 368 p. (in Russian)

[22] Goikhman, O. Ya., Nadeina, T. M. Rechevaya kommunikatsiya [Speech communication]. M.: Infra -m., 2007. 207 p. (in Russian)

[23] Zin'kina, Yu. V., Pyatkin N. V., Nevzorova O. A. Razreshenie funktsional'noi omonimii v russkom yazyke na osnove kontekstnykh pravil [Permission for functional homonymy in the Russian language on the basis of contextual rules]. Trudy mezhd. konf. Dialog. 2005. pp. 198–202. (in Russian)

[24] Golenkov, V.V., Gulyakina N.A. Semanticheskaya tekhnologiya komponentnogo proektirovaniya sistem, upravlyaemykh znaniyami [Semantic technology of component design of knowledge-driven systems]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Minsk: BSUIR, 2015. pp. 57–78. (in Russian)

[25] Ivashenko, V.P. Modeli i algoritmy integratsii znanii na osnove odnorodnykh semanticheskikh setei [Models and algorithms of knowledge integration based on homogeneous semantic networks]: PhD thesis: 05.13.17. Minsk, 2015. 151p. (in Russian)

[26] Kodzasov, S. V. Razmeshchenie tonal'nykh aktsentov v russkom slove. In G.E. Kedrovoi and V.V. Potapov (Eds.). Yazyk i rech': problemy i resheniya: sb. nauch. trudov k yubileyu professora L.V. Zlatousovoi. [Placement of tonal accents in the Russian word. Language and Speech: Problems and Solutions: Sat. sci. works for the jubilee of Professor L.V. Zlatausova]. M.:Maks-Press, 2004. pp.62–76. (in Russian)

[27] Leont'ev, A. A. Osnovy psikholingvistiki [Fundamentals of psycholinguistics]. Smysl, NPF Smysl, 2005. 310p. (in Russian)

[28] Lobanov, B. M., Tsirul'nik L. I. Komp'yuternyi sintez i raspoznavanie rechi [Computer synthesis and recognition of speech]. Minsk: Belorusskaya nauka, 2008. 344 p. (in Russian)

[29] (2017, Jun.) IMS metasystem. [Online]. Available: http://ims.ostis.net/

[30] Popov, V. E. Obshhenie s EVM na estestvennom jazyke [Communication with a computer in natural language]. M.:Nauka, 1982. 360 p. (in Russian)

[31] Rabiner, L., Gould, B. Teorija i primenenie cifrovoj obrabotki signalov [Theory and application of digital signal processing]. Perevod s angl. Zajceva A. L.; pod red. Aleksandrova J. N. M.: Mir, 1978. 848 p. (in Russian)

[32] Rabiner, L. R., Schafer R. W. Cifrovaja obrabotka rechevyh signalov [Digital processing of speech signals]. Perevod s angl.; pod red. M. V. Nazarova; Ju. N. Prohorova. M.: Radio i svjaz', 1981. 496 p. (in Russian)

[33] Tarasov, V. B. Problema ponimanija: nastojashhee i budushhee iskusstvennogo intellekta [The Problem of Understanding: The present and future of artificial intelligence]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems]. Minsk: BSUIR, 2015. pp. 25–42. (in Russian)

[34] Tolpegin, P. V., Vetrov D. P. and Kropotov D. A. Algoritm avtomatizirovannogo razreshenija anafory mestoimenij tret'ego lica na osnove metodov mashinnogo obuchenija [Algorithm for the automated resolution of the anaphora of third person pronouns based on machine learning methods]. Trudy mezhd. konf. Dialog. 2006. pp. 504–507. (in Russian)

ПОДХОД К УСТРАНЕНИЮ РЕЧЕВЫХ НЕОДНОЗНАЧНОСТЕЙ НА ОСНОВЕ СЕМАНТИКО-АКУСТИЧЕСКОГО АНАЛИЗА

Захарьев В.А. (БГУИР), Азаров И.С. (БГУИР), Русецкий К.В. (БГУИР).

В работе рассмотрен подход к проблеме устранения неоднозначностей в речевых сообщениях путем применения семантико-акустического анализа. Предлагается архитектура интеллектуальной системы, в которой, с использованием методов цифровой обработки речевого сигнала, а также формализации и обработки знаний с помощью семантических сетей, осуществляется непосредственный переход от обработки сообщения в речевой форме к анализу смыслового его содержимого (семантико-акустический анализ). На основе инструментов, предоставляемых в рамках технологии OSTIS и фреймворка обработки сигналов GUSLY, реализован прототип интеллектуальной системы для разрешения речевых неоднозначностей определенного типа: омонимов и паронимов. Показаны основные достоинства предлагаемого подхода по сравнению со стандартными системами автоматического распознавания речи, а также возможные пути дальнейшего развития предлагаемого подхода для решения задачи понимания речи.

An Intelligent System of Speech Intonation Analysis and Training

Boris Lobanov
The United Institute of Informatics Problems of National Academy of Sciences of Belarus
Minsk, Belarus
lobanov@newman.bas-net.by

Vladimir Zhitko
The United Institute of Informatics Problems of National Academy of Sciences of Belarus
Minsk, Belarus
zhitko.vladimir@gmail.com

***Abstract*—Presented in the paper is a software system designed to train learners in producing a variety of recurring intonation patterns of speech. The system is based on comparing the melodic (tonal) portraits of a reference phrase and a phrase spoken by the learner and involves active learner-system interaction. The main algorithms used in the training system proposed for analyzing and comparing intonation features are considered. A set of reference sentences is given which represents the basic intonation patterns of Russian, British English, American English, German and Chinese speech and their main varieties.**

***Keywords*—Speech intonation, melodic/tonal portrait, intonation analysis, intonation training, computer system for language learning (CALL), semantic of intonation.**

I. Introduction

Intonation plays a significant role in speech communication. It shows the general aim of an utterance and points out its information center (nucleus) as well as giving prominence to the nonnuclear **semantically relevant elements** and deaccenting those lacking in **novelty or semantic weight**; it splits an utterance into phrases (clauses) and intonation-units (groups), each presenting a syntactically organized parcel of information, and integrates these parts into an utterance, distinguishing thereby between more and less closely connected "chunks" of the speech flow. Intonation is widely recognized as an important aspect of speech that provides both linguistic and socio-cultural information. Therefore, prosodic aspects of speech should be explicitly introduced to language learners to help them communicate effectively in a foreign language.

A current linguistic idea is that a foreign accent is more evident and stable in intonation than in segmental sounds. A foreign accent in intonation emerges mainly as a result of prosodic interference, an inevitable "by-product" of bilingualism and, particularly, under the influence of the prosodic patterns of the learner's native language on those of the target language. Considering the variety of functions of intonation in speech and its potential socio-cultural effects, deviations in this area **can lead to serious semantic losses in communication**. It is a well-known fact that it is incorrect intonation that is often the cause of the wrong impression a non-native language speaker might produce [1]. Obviously, many Russian speakers fail to capture the language-specific phonetic-phonological features of American/British English intonation and, moreover, are unaware of the drastic socio-cultural effects of the deviations from the prosodic form of an utterance. Helping nonnative learners eliminate such errors presupposes ensuring their familiarity and acquisition of the prosodic patterns of the foreign language being studied.

Accuracy of reproducing the foreign intonation patterns in the process of speaking as well as adequacy of identifying the patterns on the level of perception present considerable difficulty for the learners, particularly related to their ability to control their performance and perception (especially for those who have no ear for music). The lingaphone courses and equipment available at present provide only "a hearing" feedback for intonation accuracy control, which is obviously insufficient.

The present paper is concerned with the progress achieved in developing a computer system of speech intonation analysis and training providing an additional visual feedback as well as a quantitative assessment of the learners' intonation accuracy in the foreign language teaching process.

II. Basic problems to be solve

In the course of creating the speech intonation training systems we faced a number of difficulties connected with the necessity of solving a number of technical problems, namely:

1. An adequate comparison of the pattern signal and a spoken one which is usually characterized by a non-linear time deformation and its beginning and end are not known beforehand. The solution of this problem has become possible thanks to the application of the modified method of a continuous dynamic time warping (CDTW) of two signals, developed by the author earlier [2]. The use of this method ensures automatic recognition of the end and beginning of a phrase being uttered simultaneously with its comparison with the pattern phrase.

2. Automatic segmentation of the signal being analyzed into areas for which the notion of F0 is relevant as far as the formation of the tonal contour of the phrase is concerned (the segments of vowels and most of the sonorants). This problem is being solved by means of a non-linear transfer of segment markers from the preliminarily marked pattern-phrase onto the phrase being uttered with the help of the author's earlier suggested technology of cloning the prosodic characteristics of speech [3].

3. Precise calculation of F0 of the pattern speech signal and of that produced by the learner within a very wide voice range {30 – 1000 Hz}, for male and female voices pooled. The task is solved by using the traditional methods of singling F0 out of a speech signal. Seeking a solution to the given problem has been the subject matter of a large number of publications (see e.g. [4]).

4. Automatic interpolation of current values F0 on the segments for which measuring F0 is invalid, i.e. on most of the consonants. This task is solved by using well-known interpolation mathematical formulas determining the way of finding intermediate values on the basis of an available discrete set of given values.

5. An adequate calculation of a similarity measure between the pattern signal and the uttered one under the condition of their differences in duration and F0 voice-ranges. This task is solved by using a representation of an intonation curve in the form of a unified melodic portrait (UMP) described below in the next section of the paper. Calculation of the similarity measure of two UMPs is carried out with the help of traditional formulas either by means of calculating a samples correlation coefficient or through determining the vector distance between the curves. In dealing with these problems, we relied on the results of earlier research in the field of developing automatic intonation assessment systems for computer aided language learning [5]–[8] as well as the results of our earlier research in the area of speech intonation analysis and synthesis [9]–[11].

Multi-lingual intelligent system of speech intonation analysis and training is presented here as a software package called "IntonTrainer". The software package "IntonTrainer" (hereinafter, "Application") is intended for analysis and representation on the screen of the pattern and spoken phrases intonation (F0 - basic tone trajectory), as well as for their comparison and estimation of intonation similarity. Estimation of intonation similarity is carried out on the basis of representation of intonation in the form of universal melodic portraits (UMP) [10].

III. Graphical User Interface (on example of British English learning)

The initial **Application** window that opens after the program is started is shown in "Fig. 1".

Before you start, you can preview the **Application**: settings (top right corner of the initial window) and correct them. In this window, the user can select the recording type of the signal from the microphone: (1) recording for N seconds, (2) manual control, (3) automatic, or (4) recording for N seconds + template length. In the 4th mode the choice of N = 1s is recommended. In addition, it is possible to specify the number of recorded phrases (files) stored in the **Records** folder. The "About" button opens a window with information about the developers.

After clicking the "Start" button, the main window opens, containing a structured list of reference phrases ("Fig. 2")

Figure 1. The initial window of the Application.

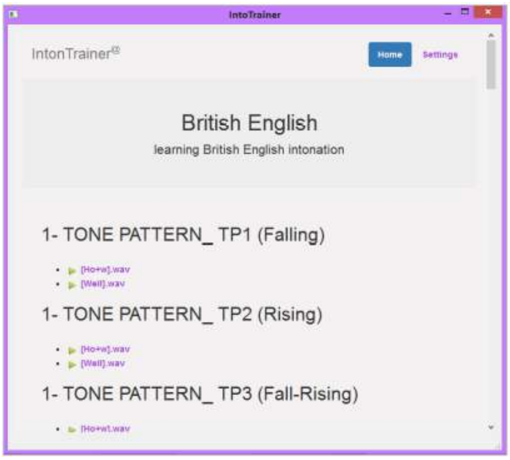


Figure 2. Main window.

IV. Study of intonation constructions of speech

By scrolling the page of the window from top to bottom, the user is given the opportunity to see the examples for the main tone patterns (TP1-TP3) of English speech. Each example provides audio and visual representation of UMP, pairwise comparison of different TPs, explains peculiarities of TP usage, as well as TP implementation in dialogues, prose and verse.

For example, clicking the mouse cursor at directory:

{TONE PATTERN_ (TP1) (Falling) [We+ll].wav},

will be open the "Graph" window in which the results of the intonation analysis of this phrase are displayed graphically ("Fig. 3").

In "Fig. 3" the red column on the left shows the range of the melody change, i.e. frequency of the pitch (F0), expressed in octaves. On the right, a linear graph of the UMP is displayed in red, the core of which is marked with frequent vertical lines. Below the graphs, the minimum and maximum values of F0 for the selected phrase are listed, as well as the text of the phrase in which the nuclear vowel is indicated by the "+" sign. Listening to the selected phrase is carried out by pressing

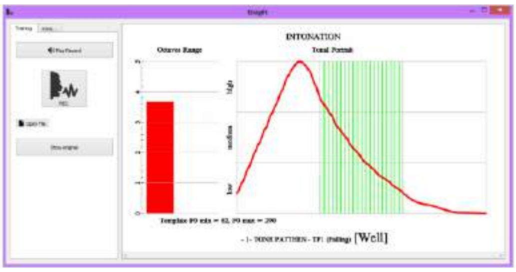


Figure 3. Analysis results window.

the **"Play Record"** button.

When the **"Show original"** button is pressed, an additional window opens ("Fig. 4"), at the top of which a waveform of the phrase signal with pre-nucleus marks (red line), nucleus (black line) and post-nucleus (blue line) is displayed. In the middle part of the window the real curve of the change F0 is depicted showing the parts of the pre-nucleus, the core and the behind-nucleus, from which the UMP is formed, shown in "Fig. 3".

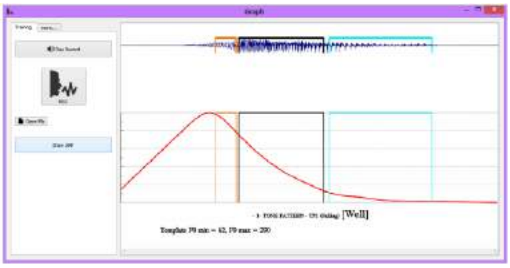


Figure 4. Additional window displaying F0 in real time.

Looking through the structured list of reference phrases in the manner described above (see "Fig. 2"), the user is introduced to and studies main intonation patterns (TP1 – TP3) of English speech, their pairing, proper of use and implementation in dialogue, prose and verse.

V. Individual Intonation Training

When using the Application for individual intonation training in the study of Russian as a foreign language, as well as for improving oral-speech intonation skills in such professions as call center operators, radio speakers, etc., the user must use an external or built-in microphone. In this case, the user should press the **"Rec"** button, wait for a short "beep-signal" and pronounce the phrase in the microphone, the text of which is indicated in the lower part of the window in "Fig. 3". After recording to the **"RECORDS"** folder and processing of the entered speech signal, the user will hear the 2nd "beep-signal", and the image in the graphics window ("Fig. 3") will be replaced with the image shown in "Fig. 5". In the upper part of the window the results of comparison of the reference and pronounced phrases are shown: Pr - proximity in % on the variation range F0 and Ps - proximity in % in the form of the trajectory F0.

In "Fig. 5" the red column on the left shows the range of change F0 of the reference phrase, and the brown one - the spoken phrase. On the right, the linear graph of the UMP of the reference phrase is displayed in red, and the brown one of the spoken phrase. Below the graphs, the minimum and maximum values of F0 of the reference and spoken phrases are given.

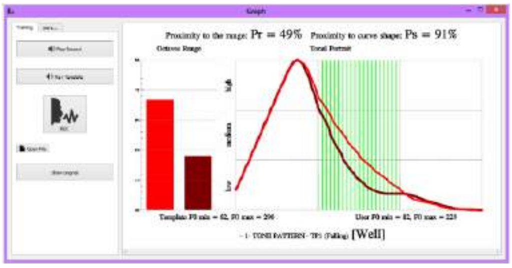


Figure 5. The window displaying the results of analysis and comparison.

Listening to the selected reference phrase is carried out by pressing the **"Play Template"** button and the pronounced phrase **"Play Record"**.

Pressing the "Show original" button opens an additional window ("Fig. 6"), at the top of which waveforms of signals of both phrases are displayed with labels of pre-nucleus (red lines), nucleus (black lines) and post-nucleus (blue lines). In the middle part of the window, the trajectories of the change F0 of the reference (red) and the spoken (brown) phrases in real time are depicted showing the sections of the pre-nucleus, nucleus and post-nucleus, from which the UMPs shown in "Fig. 5". The information contained in this window can be useful for controlling the correctness of the transfer of pre-nucleus, nucleus, and post-nucleus labels from the reference phrase to the spoken phrase. Errors in the transfer of labels can greatly distort the actual form of the UMP phrase.

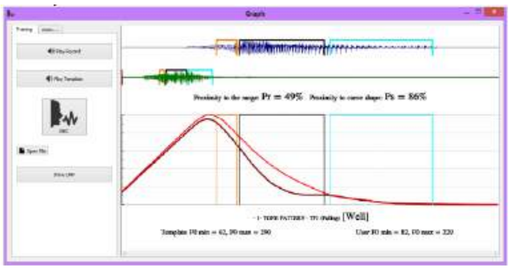


Figure 6. Additional window displaying F0 in real time.

VI. Additional Options for Using the Application

The software package "IntonTrainer" is an open system and allows its various modifications. First of all, this refers to the used set of reference data, which can optionally be supplemented or formed anew in accordance with the task at hand. For example, if you configure the Application DB for the task of learning the intonation of American English.

An important factor in the formation of the acoustic database of the studied phrases is their prosodic marking on the areas (regions) of pre-nuclear, nuclei and post-nuclei. Currently, this operation is performed manually using the standard application **"Sound Forge"**, but in the future it will be automated. The speech signal of the phrase is recorded in a "wav" format with a sampling of 8 kHz, 16 bits and is labeled into regions P1 (pre-nuclear), N1 (nuclear) T1 (per-nuclear) as shown in "Fig. 7" for a single-nuclear (one-accented) phrase: **"We+ll"**, pronounced by a male voice.

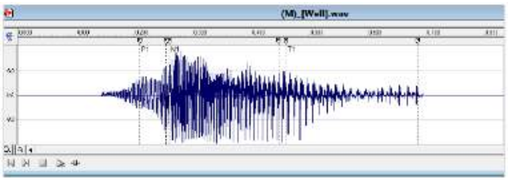


Figure 7. Example of markup of one-accented reference phrase.

Note that if the phrase is one accent, then the index is assigned to all regions. If the phrase contains 2 or more accent units (accent groups), then the P, N, T regions are assigned indices corresponding to the numbers of the accent units in the phrase. In "Fig. 8" shows an example of marking 2 accent phrases: **"Befo+r you open the do+or, ..."**.

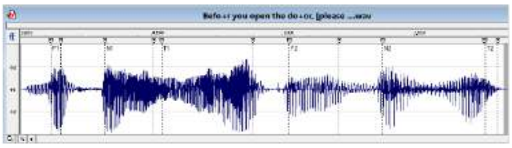


Figure 8. Example of marking a two-accent phrase.

As already mentioned, the software package **"Inton-Trainer"** can be used not only for personal purposes of linguistic education and learning intonation of oral speech, but also in a number of general scientific and practical studies. For example, the **Application** can be successfully used in experimental-phonetic or forensic studies, during which it becomes necessary to compare the standard intonation with the intonation of the phrases studied from various sources. In this case, instead of using an external or built-in microphone, acoustic implementations of these phrases are triggered by pressing the "Open File" button (see "Fig. 3"-"Fig. 6") from the specially created "TEST" folder (see "Fig. **??**").

VII. Conclusions

To date, there are demo versions of the "IntonTrainer" system, focused on learning the intonation of Russian, British English, American English, German and Chinese (see site https://intontrainer.by). The software package is recommended for use in the following current fields:

- In linguistic education (Used as a means of visualizing intonation). Primary introduction and study of the basic tone patterns (TP) of oral speech, their pairwise comparisons, application features, as well as their implementation in dialogue, prose and verse.
- In self-learning of intonation of oral speech (Used as a means of intonational training). Individual training for correct pronunciation of TP when studying a foreign language or improving intonation skills of native language in some professions: call center operators, radio and TV announcers, etc.
- In scientific and practical research (Used as a means of comparing intonation from different sources). Experimental phonetics, medical diagnostics, psychological testing, criminalistics, etc..

References

[1] Chun, D. M.: The neglected role of intonation in communicative competence and proficiency. Modern Language Journal, 72, (1988), pp. 295-303.

[2] Lobanov B.M., Levkovskaya T.V.: Continuous Speech Recognizer for Aircraft Application // Proceedings of the 2nd International Workshop "Speech and Computer" – SPECOM'97 - Cluj-Napoca, (1997), pp. 97-102.

[3] Lobanov, B.M., Tsirulnik L.I., Sizonov O.N.: ≪IntoClonator≫ – Computer system of cloning prosodic characteristics of speech (in Russian) // Proceedings of the International Conference "Dialogue 2008", Moscow, (2008), pp. 330-338.

[4] Shimamura T. and Kobayashi H.: Weighted Autocorrelation for Pitch Extraction of Noisy Speech // IEEE Transactions on Speech and Audio-Processing, Vol. 9, (2001), pp. 727–730

[5] Anne Bonneau, Vincent Colotte.: Automatic Feedback for L2 Prosody Learning. Speech and Language Technologies, (2011), pp.55-70.

[6] Yi Xu: ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. In Proc. of the TRASP'2013, Aix-en-Provence, France (2013), pp. 7-10.

[7] Juan Arias, Nestor Yoma, Hiram Vivanco.: Automatic intonation assessment for computer aided language learning. Speech Communication 52 (2010), pp. 254–267.

[8] Dávid Sztahó, Gábor Kiss, László Czap, Klára Vicsi.: A Computer-Assisted Prosody Pronunciation Teaching System. In Proc. of the WOCCI 2014, Singapore, (2014), pp. 121-124.

[9] Lobanov B.M. et al: Language- and speaker specific implementation of intonation contours in multilingual TTS synthesis // Speech Prosody: Proceedings of the 3-rd International conference, Dresden, Germany, (2006), pp. 553-556.

[10] Lobanov B., Okrut T.: Universal Melodic Portraits of Intonation Patterns of Russian Speech (in Russian) // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue". Issue 13 (20). — Moscow, (2014). pp. 330-339.

[11] Lobanov B.M.: Comparison of Melodic Portraits of English and Russian Dialogic Phrases // Computational Linguistics and Intellectual Technologies: Papers from the Annual Inernational Conference "Dialogue". Issue 15 (22). – Moscow, (2016), pp. 382-392.

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗАТОР И ТРЕНАЖЁР РЕЧЕВОЙ ИНТОНАЦИИ

Лобанов Б. М., Житко В. А.

Объединённый институт проблем информатики НАН Беларуси, Минск, Беларусь

Данная работа посвящена описанию разработанной компьютерной системы, ориентированной на начальное обучение РКИ в рамках освоения учащимися интонационных конструкций русской речи. Понятие интонационных конструкций (ИК1 — ИК7), предложено в 1960-х гг. и эффективно используется во многих современных методических пособиях по обучению РКИ.

Convolutional Neural Network with Semantically Meaningful Activations for Speech Analysis

Ryhor Vashkevich, Elias Azarov
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
ryhorv@gmail.com, azarov@bsuir.by

***Abstract*—Semantic analysis of speech is more prospective compared to analysis of text since speech contains more information that is important for understanding. The most important distinguishing feature of speech is intonation, which is inaccessible in the text analysis. For successful semantic analysis of speech it is necessary to transform the speech signal into features with semantic interpretation. The mathematical apparatus of convolutional neural networks (CNN) seems suitable to implement this kind of transformation. However there is a scalability problem that makes it hard to combine many CNN's in a single solution. To overcome this we propose to develop a CNN model with semantically meaningful activations i.e. the model that is capable of semantic interpretation of its internal states. The ultimate goal of the transform is to extract all semantically meaningful information, however the present work is confined to voice activity detection (VAD) and intonation extraction. Unlike other VADs based on artificial neural networks, the proposed model does not require a lot of computing resources and has a comparable or even better performance.**

***Keywords*—semantic speech analysis, voice activity detection, convolution neural network, VAD, CNN.**

I. Introduction

Known speech analysis and processing solutions based on neural networks can hardly be embedded into semantic systems because their internal states cannot be interpreted in semantic terms. In this paper we propose a CNN model that extracts speech intonation and voice activity using semantically meaningful activations.

A voice activity detector per se is one of the most important modules in many speech processing applications, such as audio coding, speech recognition, speaker identification, etc. The problem of voice detection in an audio signal has not yet been solved, especially in the presence of noises, which often present in the audio signal in the real world.

Significant development of machine learning in other speech processing tasks led to attempts to apply machine learning methods to VAD. In [1] the authors used a deep belief network (DBN) as the main tool for building their own VAD system. In [2, 3, 4] the authors used restricted Boltzmann machines (RBM) and networks with fully connected layers. A support vector machine (SVM) is used in [5, 6, 7] to classify features of a speech signal as one of the most computationally simple methods of classification. In [8, 9], the authors used the fact that a sound signal is a time series, and they used recurrent neural networks (RNN) for building VAD systems.

Another disadvantage of deep neural networks is computational complexity. Considering that the voice detector is often only an auxiliary module of a speech processing system, it is necessary to be sure that the VAD module consumes as little computing resources as possible.

The model proposed in this work provides a high accuracy of VAD comparable to existing solutions based on neural networks but uses much fewer (by several orders) parameters. A useful property of the obtained solution is the possibility of estimating a basic pitch of a speech signal. This estimation is generated by network activations.

II. Proposed method

A. Features extraction

The choice of characteristic features of a sound signal is one of the most important part of a VAD system building process. We propose to use a fact that a speech signal has harmonic components, which our model tries to detect. As basic features of a speech signal, most works use mel-frequency cepstral coefficients (MFCC) [1, 4, 5, 6, 7, 8, 10]. Instead, we propose to use a spectrogram of an audio signal, and show that our model is able to efficiently detect the harmonic components of the signal, which are the main criterion for the presence of a voice in an audio signal. As shown in Fig. 1, the harmonic components of the sound signal are clearly visible in those parts of the spectrogram where the voice is present. In this case, the speech signal can be described using a fundamental frequency ($F0$), an amplitude, and number of harmonic components.

The fundamental frequency of the harmonic signal is the frequency corresponding to the first harmonic. The frequencies of all the harmonics of the speech signal are multiples of $F0$. To determine whether the signal is harmonic, we select the amplitudes of only those frequency components from the spectrogram that correspond to harmonics for a given $F0$ and feed them to the CNN model. In our experiments, we assume that $F0$ takes values from the range from $F0_{min} = 70$Hz and to $F0_{max} = 350Hz$. Also we introduce the notation for number of harmonics M and number of possible fundamental frequencies N. These variables are hyperparameters and may have different values.

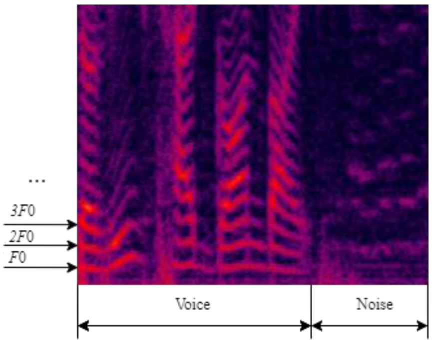


Figure 1. Spectrogram of sound signal.

The possible F0 range $[F0_{min}, F0_{max}]$ is uniformly covered by frequency grid G $[F0_0, \dots, F0_{N-1}]$ where each point can be calculated as:

$$F0_i = Fo_{min} + i * \frac{F0_{max} - F0_{min}}{N-1}, i = \overline{0, (N-1)}$$

The indices of all the M-harmonics for given $F0_i$ are calculated as:

$$index_j = round\left(\frac{(1+j) * F0_i * N_{fft}}{2 * f_s}\right), j = \overline{0, (M-1)}$$

where f_s is audio sample rate, N_{fft} is the format of the fast Fourier transform (FFT).

Decision is frame-based (one frame one decision). The log amplitude spectrum of signal frame s is calculated as:

$$S = log_{10}|FFT(s)|$$

which is transformed into features vector X:

$$X(i,j) = S(index_j)$$

Thus for each signal frame we form a matrix of features X with shape $N \times M$, consisting of N points for M components. The basic idea is shown in Fig. 2.

We can consider each point of the frequency grid G as a pitch candidate. The task of the neural network model is to determine whether among the selected candidates there is one, which clearly represents the harmonic structure of the signal. If the model can detect a candidate that describes a harmonic signal, then the current input example corresponds to a speech signal, otherwise this example is classified as noise.

B. CNN architecture

In this paper we propose to use a simple model of a convolutional neural network consisting of only two 2D convolutional layers, followed by a global max-pooling layer. Fig. 3 shows the architecture of the proposed model.

The input of the model is a matrix of features X with shape $N \times M$. The first convolutional layer has K filters of size $1 \times M$ with a ReLu activation function. The second convolutional layer has only 1 filter with size 1×1. It aggregates the features selected by the convolution filters on the previous

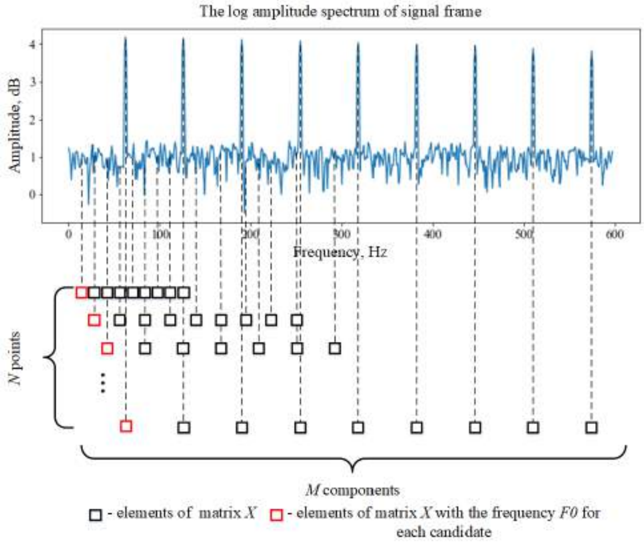


Figure 2. The basic idea of features selection.

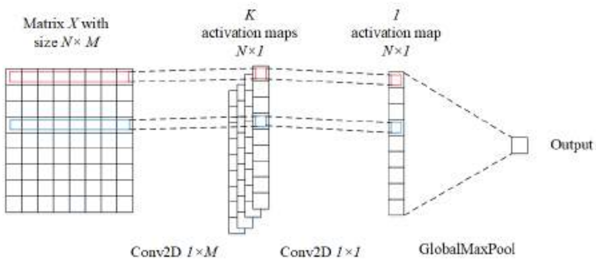


Figure 3. Architecture of CNN for VAD.

layer for each of the N candidates separately. This layer does not have an activation function. After than, the global max-pool layer selects the candidate with the maximum activation, thereby assuming that the candidate contains the harmonic component of the input signal. Further, the candidate selected passes through a sigmoidal activation function, as a result of which the output of the network will represent the probability that the input signal is a voice.

C. Pitch extraction

The second Conv2D layer forms an estimate of the pitch frequency. A high value of any activation of the second Conv2D layer gives a high degree of confidence that an input sample is a periodic signal.

III. Experiments

A. Training

The network was trained using SGD with learning rate η and momentum 0.9. We used a binary cross-entropy (1) as a loss function.

$$L_{ce} = \frac{1}{N_X} \sum_{i=0}^{N_x - 1} -y_i * log(t_i) - (1-y_i) * log(1-t_i) \quad (1)$$

where N_X is number of training samples, y_i is network output for i-th training sample, t_i is target value of the class label for the i-th training sample.

B. Dataset preparation

To train the model, we used our own dataset, consisting of 50101 examples. This dataset was divided into training and testing subsets in a ratio of 3 to 1. Additive and multiplicative components from white noise were added to the dataset.

C. Hyperparameters tuning

To determine optimal values for hyperparameters, we divided the dataset into training and testing subsets in a ratio of 3 to 1 and trained several model configurations with the different hyperparameter values. We experimented with different values of N, M and K. The hyperparameters tuning results are presented in Table I.

Table I
HYPERPARAMETERS TUNING RESULTS

Model	K	M	N	Training L_{ce}	Testing L_{ce}
M-3-7-100	3	7	100	0.61215	0.44838
M-3-7-50	3	7	50	0.59604	0.52501
M-3-7-200	3	7	200	0.63900	0.48520
M-3-14-100	3	14	100	0.63713	0.70615
M-3-14-50	3	14	50	0.59120	0.52208
M-3-14-200	3	14	200	0.56230	0.59663
M-3-20-100	3	20	100	0.69536	0.80170
M-3-20-50	3	20	50	0.57937	0.46226
M-3-20-200	3	20	200	0.61532	0.87776
M-10-7-100	**10**	**7**	**100**	**0.57937**	**0.37481**
M-10-7-50	10	7	50	0.60293	0.40301
M-10-7-200	10	7	200	0.89240	0.45877
M-10-14-100	10	14	100	0.59618	0.49430
M-10-14-50	10	14	50	0.65694	0.44893
M-10-14-200	10	14	200	0.92659	0.65541
M-10-20-100	10	20	100	0.62169	0.44377
M-10-20-50	10	20	50	0.61024	0.44842
M-10-20-200	10	20	200	0.88882	0.80149

As shown in Table I, the best results on the testing dataset has a model named "M-10-7-100". The optimal values of the all hyperparameters are presented in Table II. We used the hyperparameters of this model in the remaining experiments.

Table II
OPTIMAL HYPERPARAMETERS VALUES

Hyperparameter	Value
Number of candidates, N	100
Number of harmonics, M	7
Number of filters in the first Conv2D layer, K	10
Frame size, FFT size, N_{fft}	4096
Frame step	224
Signal sample rate, f_s	44100Hz
Learning rate, η	0.01

D. Comparison with other models

As a baseline model for comparison, we took the VAD model similar to that proposed in [11]. This model is a 4-layer neural network, consisting only of dense layers and uses MFCC, delta-MFCC, and delta-delta-MFCC as basic features. The input vector consists of $13 * 3 = 351$ nodes. Here, 13 is the number of MFCC coefficients, and 3 is the total number of different features.

The network was trained using SGD with learning rate 0.01, momentum 0.9, and loss function (1).

The training results of the proposed and basic models are presented in Table III. Both models were tested using the cross-validation technique.

Table III
RESULTS OF EXPERIMENTS

Criterion	Proposed model	Baseline model
Number of parameters	90	969,218
Training L_{ce}	0.33469	0.69316
Training accuracy, %	86.27	50.486
Testing L_{ce}	0.45104	0.69159
Testing accuracy, %	78.99	57.668

As show in Table III, our model has extremely less trainable parameters and it takes a lot less time to perform a forward pass.

E. Pitch extraction

As shown in Fig 4 our model can estimate the pitch frequency. The pitch frequency corresponds to the activation number of the second Conv2D layer. The model generates a hight output value if it can detect the pitch frequency of input sample.

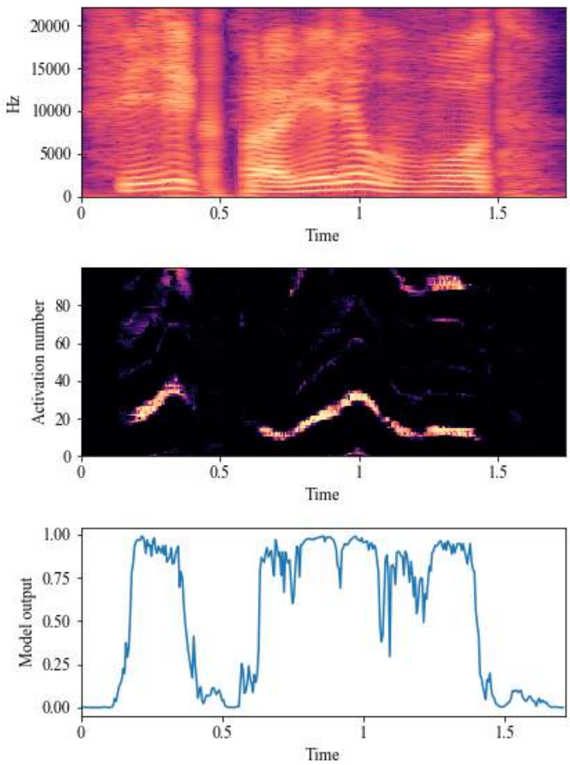


Figure 4. The pitch frequency estimation.

IV. Conclusion and future work

In this paper, we proposed a method for voice activity detection in a sound signal based on a simple convolutional neural network. A key feature of the method is the selection of characteristic features of a speech signal. Using the fact that a voiced speech has a harmonic structure as characteristic features we proposed to use sound signal spectrogram coefficients which are multiples of the specified fundamental frequency. Due to the fact that each voice has its own pitch frequency - we use 100 variants of values for the fundamental frequency. These values are uniformly located in the range from $70Hz$ to $350Hz$. The obtained features thus feed to the input of a 2-layer convolutional neural network, which classifies the input example into two classes - a voice or a noise.

Performance of the proposed model is comparable to state-of-the-art models based on neural networks, but our model contains significantly fewer trainable parameters. Therefore much less data is needed to train the model, and much less time is taken to perform a forward pass, and it increases the performance of the entire system.

Further improvement of the method will be aimed at 1) extending internal states of CNN to represent additional semantically important information; 2) increasing a model's inference quality using examples containing a harmonic signal which is not a voice (for example, musical instruments).

References

[1] Xiao-Lei Zhang, Ji Wu, Deep Belief Networks Based Voice Activity Detection, IEEE Transactions on Audio, Speech, and Language Processing, 2013, vol. 21, no. 4, pp. 697-710.

[2] Qing Wang1, Jun Du1, Xiao Bao, A universal VAD based on jointly trained deep neural networks, INTERSPEECH, 2015.

[3] T. G. Kang, K. H. Lee, W. H. Kang, DNN-based voice activity detection with local feature shift technique, Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016, pp. 1-4.

[4] Ryant, N., Liberman, M., Yuan, J., Speech activity detection on youtube using deep neural networks, INTERSPEECH, 2013, pp. 728-731.

[5] Tomi Kinnunen, Evgenia Chernenko, Marko Tuononen, Voice Activity Detection Using MFCC Features and Support Vector Machine, 2007.

[6] Y. X. Zou, W. Q. Zheng, W. Shi, Improved Voice Activity Detection based on support vector machine with high separable speech feature vectors, 19th International Conference on Digital Signal Processing, 2014, pp. 763-767.

[7] X. L. Zhang, J. Wu, Denoising deep neural networks based voice activity detection, IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 853-857.

[8] F. Vesperini, P. Vecchiotti, E. Principi, Deep neural networks for Multi-Room Voice Activity Detection: Advancements and comparative evaluation, International Joint Conference on Neural Networks (IJCNN), 2016, pp. 3391-3398.

[9] F. Eyben, F. Weninger, S. Squartini, Real-life voice activity detection with LSTM Recurrent Neural Networks and an application to Hollywood movies, IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 483-487.

[10] S. M. R. Nahar, A. Kai, Robust Voice Activity Detector by combining sequentially trained Deep Neural Networks, International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA), 2016, pp. 1-5.

СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ С СЕМАНТИЧЕСКИ-ЗНАЧИМЫМИ АКТИВАЦИЯМИ ДЛЯ АНАЛИЗА РЕЧИ

Г. Вашкевич, И. Азаров
БГУИР
Минск, Беларусь

Семантический анализ речи является более перспективным по сравнению с анализом текста, поскольку речь содержит больше информации, которая важна для понимания. Самой важной отличительным признаком речи, недоступным текстовом анализе, является интонация. Для успешного семантического анализа речи необходимо из речевого сигнала выделить характеристические признаки с семантической интерпретацией. Математический аппарат свёрточных нейронных сетей (CNN) представляется подходящим для реализации такого рода преобразований. Однако существует проблема масштабируемости, которая затрудняет объединение нескольких CNN в одном решении. Чтобы преодолеть это, мы предлагаем разработать модель CNN с семантически значимыми активациями, то есть модель, внутрениие состояния которой можно интерпретировать с семантической точки зрения. Конечная цель преобразования состоит в том, чтобы извлечь из речи всю семантически значимую информацию, однако настоящая работа ограничивается детектированием голосовой активности и выделением интонации.

Благодатя предложенному в работе методу выделения характеристических признаков звукового сигнала и выбранной архитектуре нейронной сети стало возможным оценить частоту основного тона гармонического сигнала. Сильная активация какого-либо выхода второго слоя нейронной сети позволяет судить о гармонической природе входного сигнала. Если при этом сопоставить данный выход со шкалой частот, то можно будет получить численное значение частоты основого тона гармонического сигнала.

Предложенная модель по производительности сопоставима с другими современными моделями на основе нейронных сетей, однако содержит значительно меньше обучаемых параметров. Из этого следует, что для ее обучения необходимо гораздо меньше данных. При этом простота архитектуры нейронной сети позволяет использовать ее в мобильных платформах или встраиваемых системах.

Дальнейшее совершенствование метода будет направлено на повышение качества работы модели на примерах, содержащих гармонический сигнал, но при этом не относящийся к голосу (например звук музыкальных инструментов).

Analysis of verbal disorders with the help of neural networks

Menshakou P.A.
Gomel State Technical University named after P.O. Suhoy
Gomel, Belarus
pmenshakov@gmail.com

***Abstract*—At present, voice identification and voice analysis are one of the foremost directions, both in the field of information security systems and in the definition of voice parameters [1]. It is proposed to consider a method for analysing the state of the voice and identifying its problems, such as fatigue of the vocal cords, damage or inflammation of the vocal tract with the help of the previously recorded voice of the speaker and neural network analysis of voice changes.**

***Keywords*—Voice, identification, neural networks.**

I. Introduction

Evaluation of voice impairment involves the use of the evaluation of multiple speech characteristics (degree of respiration, tension, roughness). One of the problems associated with the use of multidimensional data is their comparison. To perform the comparison and classification it is proposed to use the self-organizing map of Kohonen. In view of the possibility of learning without a teacher, it does not need a target vector for outputs and, consequently, does not require comparison with predetermined ideal answers, and the training set consists only of input vectors. The learning process, therefore, highlights the statistical properties of the learning set and groups similar vectors into classes. The input of a vector from a given class will give a certain output vector.

II. Hardware implementation

To implement the device you need a microphone, filter and analog-to-digital converter, for further work with digital voice recording. The circuit of the device is shown in Figure 1.

From the output of the microphone, the signal is fed to the input of the filtration unit. The next step is the passage of the ADC [2]. Further, the digitized signal enters the digital processing unit. In the digital processing block, the signal is filtered and converted into a vector, with which the microprocessor and the neural network processor will continue to operate. For the subsequent comparison with the previously saved vector of chalk-cepstral coefficients, the obtained vector is stored in non-volatile memory. After comparing the vector in memory with the resulting vector, the micro controller instructs the control unit of the external device, for example, on the magnetic door lock. The process of voice identification is not demanding of resources, and consists of two stages. The first stage is the receipt of the speech characteristics of the announcer and the conversion to a form in which it can be compared with others. The second step is to compare them with a trained neural network [3].

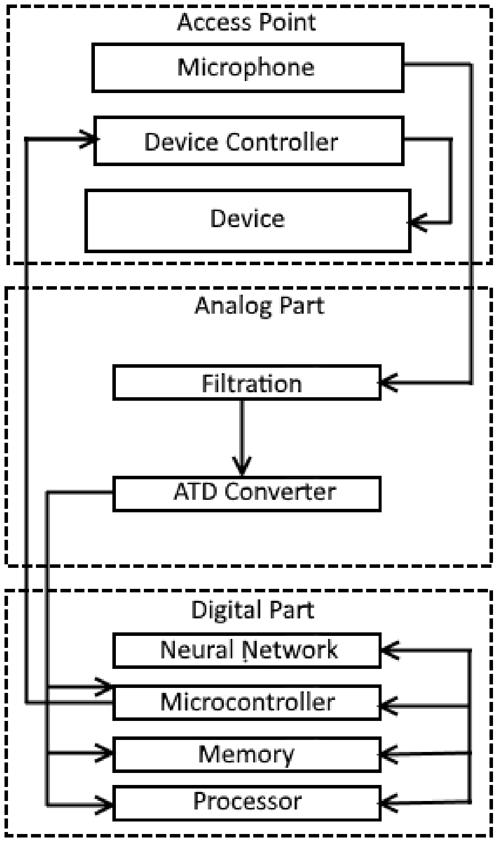


Figure 1. Hardware scheme

III. Dysphonia

Dysphonia is clearly defined as the underlying disorder of phonation, a consequence of diseases or pathology of the vocal cords. Since this deviation of the work of the vocal tract is accompanied by an audible change in the voice, it can be recorded and processed in relation to the sample of the voice of the announcer, before the appearance of deviations. There are two types of dysphonia - functional and damaging. Damaging dysphonia is divided into two types: congenital and acquired dysphonia. Damaging dysphonia in some cases may appear after functional dysphonia. Functional dysphonias are possible:

- hoarseness;

- laryngitis;
- inflammation of the larynx;
- hyperkinesis;
- mucous cyst or cyst shell;
- hypokinesis;
- a throat infection;
- glossoplegia;
- rhinopharyngitis.

These diseases have different degrees of severity, which significantly affects the quality of life. Thus, diagnosing fade disorders will help determine the level of quality of life. The scale of voice disorders is shown in Table 1.

Table I
SCALE OF VOICE DISORDERS

Gradation of violations	Degree of Violations	Description	Recommendations
No violations	0	-	-
Minor violations	1	Linguistic disorder is hardly felt or felt by the patient alone	Speech therapy is recommended
Moderate abnormalities	2	Decreased ease and speed of speaking	Speech therapy needed
Heavy Violations	3	The Talker needs the help of a listener. The patient often can not be understood, but understands himself	Speech therapy and help from the listener are necessary
Deep damage	4	Speaking with fragmentary expressions. The listener has to guess a lot. Information is small, and the listener	Need speech therapy and the study of sign language, consultation or synthesis of voice.

IV. ANALYSIS OF SPEECH IMPAIRMENT

In a study conducted by Leinonen Et Al.[4], To replace direct listening to the voice, a scale of assessments of various degrees and forms of dysphonia was created. To compare the criteria, a neural network without a teacher was used, the training of which was conducted using a perceptual map of estimates of the normal and dysphonic voice [3]. The results of the experiment are shown in Fig. 2.

This approach has several disadvantages:

- Absence of comparative characteristics with the previous state of the voice. In view of this shortcoming, it is impossible to separate the congenital disorders from acquired ones;
- Lack of diagnosis of several disorders at the same time.

Classification of various forms and the degree of dysphonia can be made by using not only perceptual assessments of pathology, roughness, respiration, tension and asthenia, but also by comparing the estimates with the previous value, by including a voiceprint with a simulation of some degree of disease as input parameters. Thus, the voice print can be used not only for access control equipment, but also for assessing the speaker's voice deviations. The obvious advantage of this approach is the absence of direct contact with the speaker and the possibility of remote diagnostics, including the use of electronic means of communication. Also, this procedure has the possibility of full automation and undemanding resources.

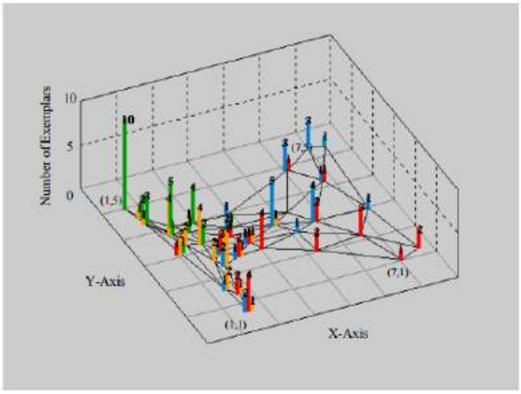


Figure 2. The results of determining the voice dysphonia (green - normal voice, yellow - hypotonic, red - hypertonic, blue - spasmodic)

V. NEURAL NETWORK COMPARISON

The self-organizing map (SOM) (Kohonen, 1982) is one of the most important neural network architecture. Since its invention it has been applied to so many areas of Science and Engineering that it is virtually impossible to list all the applications available to date (van Hulle, 2010; Yin, 2008). In most of these applications, such as image compression (Amerijckx et al., 1998), time series prediction (Guillen et al., 2010; Lendasse et al., 2002), control systems (Cho et al., 2006; Barreto and Araújo, 2004), novelty detection (Frota et al., 2007), speech recognition and modeling (Gas et al., 2005), robotics (Barreto et al., 2003) and bioinformatics (Martin et al., 2008), the SOM is designed to be used by systems whose computational resources (e.g. memory space and CPU speed) are fully available. However, in applications where such resources are limited (e.g. embedded software systems, such as mobile phones), the SOM is rarely used, especially due to the cost of the best-matching unit (BMU) search (Sagheer et al., 2006). Essentially, the process of developing automatic speech recognition (ASR) systems is a challenging tasks due to many factors, such as variability of speaker accents, level of background noise, and large quantity of phonemes or words to deal with, voice coding and parameterization, among others. Concerning the development of ASR applications to mobile phones, to all the aforementioned problems, others are added, such as battery consumption requirements and low microphone quality. Despite those difficulties, with the significant growth of the information processing capacity of mobile phones, they are being used to perform tasks previously carried out only on personal computers. However, the standard user interface still limits their usability, since conventional keyboards are becoming smaller and smaller. A natural way to handle this new demand of embedded applications is through speech/voice commands. Since the neural phonetic typewriter (Kohonen, 1988), the SOM has been used in a standalone fashion for speech coding and recognition (see Kohonen, 2001, pp. 360-362). Hybrid architectures, such as SOM with MultiLayer Perceptrons (SOM-MLP) and SOM with Hidden Markov Models (SOM-HMM), have also been proposed (Gas et al., 2005; Somervuo, 2000). More specifically, studies involving speech recognition in mobile devices systems include those by Olsen et al. (2008); Alhonen et al. (2007) and Varga and Kiss (2008). It is worth noticing that Portuguese is the eighth, perhaps, the seventh most spoken language worldwide and the third among the Western countries, after English and Spanish. Despite that, few automatic speech recognition (ASR) systems, specially commercially available ones, have beendevelopedand it isavailable worldwidefor the Portugueselanguage. This scenario is particularly true for the Brazilian variant of the Portuguese language, due its large amount of accent variation within the country. Scanzio et al. (2010), for example, report experiments with a neural network based speech recognition system and include tests with the Brazilian Portuguese language. Their work is focused on a hardware-oriented implementation of the MLP network. In this context, the current paper addresses the application of self-organizing maps to the Brazilian Portuguese isolated spoken word recognition in embedded systems. For this purpose, we are particularly interested in evaluating several software strategies to speedup SOM computations in order to foster its use in real-time applications. The end-user application is a speaker-independent voice-driven software calculator which is embedded in smartphones.

We used learning without a teacher, because it is much more plausible model of learning in the biological system. Kohonen developed and many others, it does not need to output the target vector and therefore, does not require comparison with predetermined ideal responses, and learning set consists only of the input vectors. The training algorithm adjusts network weights so as to produce consistent output vectors, ie, to sufficiently close the presentation of input vectors produce the same outputs. The learning process, therefore, highlights the statistical properties of the training set and groups similar vectors in the classes. Presentation of the input vector of this class will give a certain output vector. The spread signal in such a network is as follows: input vector is normalized to 1.0 and applied to the input, which distributes it on through the matrix of weights W. Each neuron in layer Kohonen calculates the sum at its input and depending on the condition of the surrounding neurons becoming active layer or inactive (1.0 and 0.0). Neurons in this layer operate on the principle of competition, ie. E. As a result of a certain number of iterations is still an active one neuron or a small group. This mechanism is called lateral. Since testing of this mechanism requires significant computing resources, in my model it replaced by finding the maximum neuron activity and awarding him the activity 1.0, and 0.0 all other neurons. Thus, the neuron is activated for which the input vector closest to the vector of the weights. As a sigmoid activation function is used, which is as follows:

$$f(x) = 1/(1 + e^{-a*x}) \quad (1)$$

where a – slope parameter.
Geometrically, this rule shows next picture:

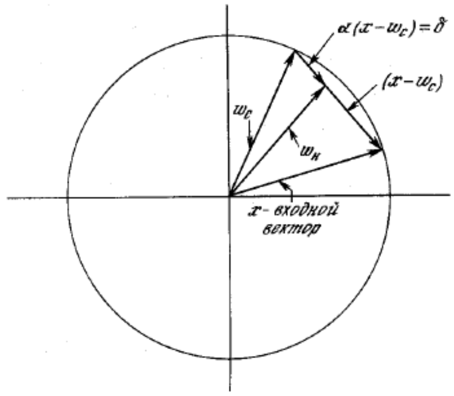


Figure 3. Correction weights of Kohonen neuron

Since the input vector x is normalized, ie. E. Is on a hypersphere of unit radius in the space of weights, then the correction weights on this rule is rotated vector weights toward the input that allows to produce statistical averaging of input vectors, which reacts active neuron. Thus, the study was replaced lateral approach leading to the activation of neurons.

VI. Conclusion

The result of this study was to modular application performing voice user authentication with analysis of user voice disorders. The program consists of three main parts. The first carries the addition of users, the second and the third carries the identification sends information to identify the user.

References

[1] GOST R 51241-2008. Sredstva i sistemy kontrolya i upravleniya dostupom. Klassifikatsiya. Obshchie tekhnicheskie trebovaniya. Metody ispytanii [State Standart R 51241-2008. Means and systems of control and access control. Classification. General technical requirements. Test methods]. Moscow, Standartinform, 2009. 10p.

[2] Raho, G.I. [et al.] Universal Developing of Persons Identification Based on RFID/ In Journal of Emerging Trends in Computing and Information Sciences, 2015. Vol. 6, No. 10. P. 592-597.

[3] Adeyemo, Z.K. Development of Hybrid Radio Frequency Identification and Biometric Security Attendance System. In International Journal of Applied Science and Technology, 2014. Vol. 4, No. 5. P. 190-197.

[4] Leinonen, L. Dysphonia detected by pattern recognition of spectral composition. In Journal of Speech, Language, and Hearing Research. April 1992, Vol. 35, 287-295. doi:10.1044/jshr.3502.287

[5] Biometrics, Computer Security Systems and Artificial Intelligence Applications / Ed. K. Saeed, J. Pejas, R. Mosdorf. Springer, 2006. 345 p.

[6] You, Y. Audio Coding: Theory and Applications. NY: Springer, 2010. 349 p.

[7] Herbig, T. [et al.] Self-Learning Speaker Identification: A System for Enhanced Speech Recognition. Berlin:Springer Verlag Gmbh, 2011.

[8] Al-Shayea Q.K., Speaker Identification: A Novel Fusion Samples Approach / Qeethara Kadhim Al-Shayea, Muzhir Shaban Al-Ani. In International Journal of Computer Science and Information Security (IJCSIS), 2016. Vol. 14, No 7. P.423-427.
[9] Ding Ing-Jr., Chih-Ta Yen, Yen-Ming Hsu. Developments of Machine Learning Schemes for Dynamic Time-Wrapping- Based Speech Recognition. In Mathematical Problems in Engineering, 2013. P. 5.
[10] Kohonen, T. Self-Organizing Maps. Berlin: Springer, 2001.
[11] Bosi, M., Goldberg, R.E. Introduction to digital audio coding and standards. Springer Science+Business, Media USA, 2010. 434 p.
[12] Menshakou, P.A. Voice user identification in access control systems / P.A. Menshakov, I.A. Murashko // Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2016, pp. 411-414
[13] Sergienko, A. B. Cifrovaya obrabotka signalov. 2-e izdanie. SPb: Piter, 2006. 751 p.
[14] Harris, F.J. On the use of windows for harmonic analysis with the discrete Fourier transforms. In Proceedings of the IEEE. Vol. 66, Jan 1978. P. 51-83.
[15] Azarov, I.S., Petrovskii, A.A. Primenenie mgnovennogo garmonicheskogo analiza dlya antropomorficheskoi obrabotki rechevih signalov. In Vestnik MGTU im. N.E. Baumana. Ser. "Priborostroenie", 2012. p. 51.
[16] Using self-organizing maps to cluster music files based on lyrics and audio features. In Research Congress 2013 De La Salle University, Manila, March 7-9, 2013.
[17] Ghitza, O. Auditory Models and Human Performance in Tasks Related to Speech Coding and Speech Recognition. In IEEE Transactions on Speech and Audio Processing, 1994. Vol. 2, No 1. P. 115–132.
[18] Azarov, I.S., Petrovskii, A.A. Primenenie mgnovennogo garmonicheskogo analiza dlya antropomorficheskoi obrabotki rechevih signalov. In Dokladi BGUIR, 2011. No 4. PP. 59–70.
[19] Mwasiagi, I. Self Organizing Maps – Applications and Novel Algorithm Design. In Josphat Igadwa Mwasiagi InTech. January 21, 2011. P. 91.
[20] Alejandro, C. Analysis of Kohonen's Neural Network with application to speech recognition. In Mexican International Conference on Artificial Intelligence. Mexico: Guanajuato, 2009. P. 8.
[21] Koikkalainen, P. Self-Organizing hierarchical feature maps/ P. Koikkalainen, E. Oja// Proceedings of 1990 International Joint Conference on Neural Networks vol. II, San Diego, 1990. pp. 279–284.
[22] Zagumenov, A. P. Komputernaya obrabotka zvuka. M.: DMK, 1999. 384 pp.

ИССЛЕДОВАНИЕ РЕЧЕВЫХ РАССТРОЙСТВ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ

Меньшаков П. А.
Кафедра «Информационные технологии»,
Гомельский государственный технический университет имени П.О. Сухого
Гомель, Республика Беларусь

В настоящее время, голосовая идентификация и анализ голоса являются одним из передовых направлений, как в области систем защиты информации, так и в определении голосовых параметров [1]. Предлагается рассмотреть способ анализа состояния голоса и выявления его проблем, таких как усталость голосовых связок, повреждения или воспаления речевого тракта при помощи ранее записанного голоса диктора и нейросетевого анализа голосовых изменений.

Traffic sign detection and problems in the field of computer vision

Marina Lukashevich
Department of Electronic Computing Machines
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
lukashevich@bsuir.by

***Abstract*—Object detection is a typical task of computer vision. This paper presents some results of implementation of the traffic sign recognition. We use R-CNN for traffic sign detection system. We focus on speed limit superclasses of traffic sign. R-CNN deep learning detector is a simple and suitable model for the traffic sign recognition. This approach combines multiple low-level image features with high-level context from object detectors and scene classifiers. Despite the existing advances in computer vision, the article considers the problems that exist and which need to be solved in the future in the field of computer vision system design.**

***Keywords*—object detection, deep learning, R-CNN, intelligent system, computer vision**

I. Introduction

In recent years image processing and analysis, pattern detection are the most exciting and fastest-growing research areas in the computer vision. Recent computer vision technologies and algorithms are support efficient semantic image segmentation and classifications. Intelligent driver assistance systems are systems to help the driver in the driving process. Traffic sign recognition is a technology by which a vehicle is able to recognize the traffic signs put on the road. This is part of the Intelligent driver assistance systems [1]. The goal of our research is to detect traffic signs into speed limit superclasses as shown in Fig. 1.

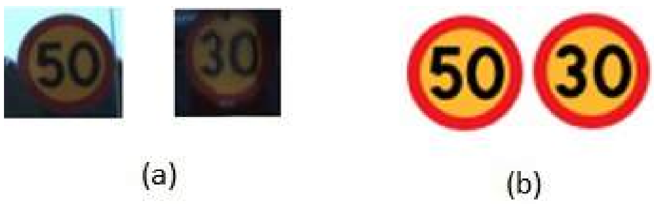


Figure 1. Examples of real-life traffic signs superclasses (a) and there synthetic examples (b).

Object detection is the task of finding the different objects in an image and classifying them. In our research, instead of developing of traffic sign detection system by applying R-CNN detector we try to formulate problems and goals that exist in the field of computer vision. Computer vision is one of the fields of Artificial Intelligence that has grown the most in the last 15 years. But we can speak only about private decisions of individual narrow tasks of computer vision.

This paper is structured as follows. Session II-IV describes image processing stages and used algorithm. In session V we formulate problems and goals in the field of computer vision.

II. Related research studies

A comprehensive review on the recent achievements of traffic sign recognition was presented in [2]. Many different approaches have been used for traffic sign recognition. Typical sign detection algorithms consist of three stages: segmentation, detection and classification, Fig. 2.

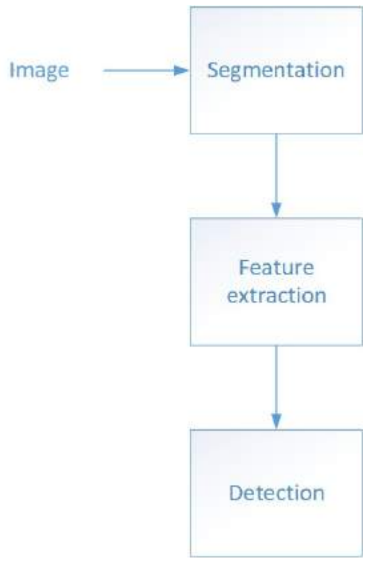


Figure 2. Typical sign detection algorithms

Traditional approach for segmentation is to threshold in chosen color space (HSV, HSI or CBH) [3-5], in order to obtain a binary image, Fig. 3.

Feature extraction is the second part of this process. In [6] authors have calculated Fourier descriptors. Other approach is using various HOG-features or edges [2]. The last stage is detection and classification using suitable classifier (SVM, neural network, Cascaded classifier, Fuzzy templates, Hough shape detection, etc.) [2], Fig. 4.

Figure 3. Example of segmentation process

Figure 4. Example of traffic sign detection

III. R-CNN OBJECT RECOGNITION

This typical scheme does not always provide the required accuracy of the detection. Traditional machine learning makes it possible to obtain good results with small datasets. It is possible to train a model quickly. But there is an accuracy plateaus in this case. We need to try different features and classifiers to achieve best results.

Recently, deep learning have significantly improved image classification and object detection accuracy [7]. Traditional neural networks contain only 2 or 3 layers, while deep networks can have hundreds. Deep learning is a type of machine learning in which a model learns to perform classification tasks directly from images. The network increases the complexity and detail of what it is learning from layer to layer. Deep learning requires very large data sets and computationally intensive. However, such approach allows us to learn features and classifiers automatically. The accuracy is unlimited in this case. The network learns directly from the data. We have no influence over what features are being learned [8].

At present, a large number of models of deep neural networks are proposed. A convolutional neural network (CNN) is one of the most popular algorithms for deep learning with images. The first successful applications of Convolutional Networks were developed by Yann LeCun. Its named LeNet [9], Fig. 5.

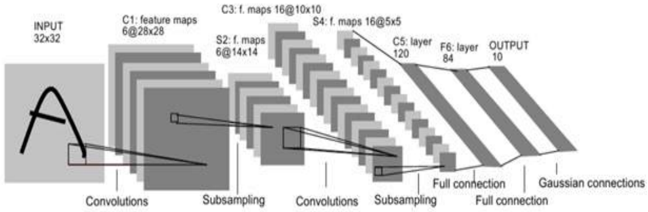


Figure 5. Architecture of LeNet-5, a Convolution Neural Network

The AlexNet was submitted to the ImageNet ILSVRC challenge [10] and significantly outperformed the second runner-up [11], Fig. 6.

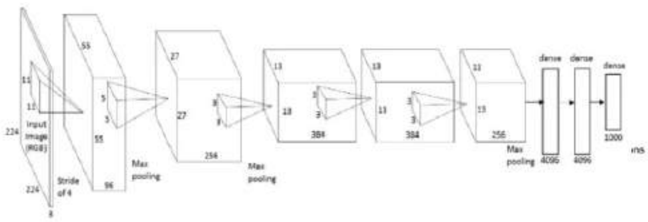


Figure 6. Architecture of AlexNet, a Convolution Neural Network

The ILSVRC 2014 winner was a Convolutional Network from Szegedy et al. from Google (GoogLeNet) [12]. This is not a complete list of models.

The goal of R-CNN (Regional CNN) is to take in an image, and identify where the main objects (via a bounding box) in the image. The first generates category-independent region proposals. These proposals define the set of candidate detections available to detector. The second module is a large convolutional neural network that extracts a fixed-length feature vector from each region. The third module is a set of class-specific linear SVMs [13].

Object detection system is presented in Fig. 7.

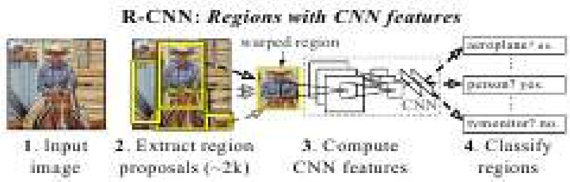


Figure 7. Object detection system [13]

IV. EXPERIMENTAL EVALUATIONS

A few publicly available traffic sign data sets exist:

- German TSR Benchmark (GTSRB) [14,15];
- KUL Belgium Traffic Signs Data set (KUL Data set) [16];
- Swedish Traffic Signs Data set (STS Data set) [6];
- RUG Traffic Sign Image Database (RUG Data set) [17];
- Stereopolis Database [18].

The evaluation of traffic sign detection is based on Swedish Traffic Signs Data set [6, 19]. A dataset has been created by recording sequences from over 350 km of Swedish highways and city roads. A 1.3 mega-pixel color camera at the resolution of 1280×960, a Point-Grey Chameleon, was placed inside a car on the dashboard looking out of the front window. The camera was pointing slightly to the right, in order to cover as many relevant signs as possible. The lens had a focal length of 6.5 mm, resulting in approximately 41 degrees field of view. Typical speed signs on motorways are about 90 cm wide, which corresponds to a size of about 50 pixel if they are to be detected at a distance of about 30 m.

In total, in over 20 000 frames have been recorded of which every fifth frame has then been manually labeled. The label for each sign contains sign type (pedestrian crossing, designated lane right, no standing or parking, priority road, give way, 50 kph, or 30 kph), visibility status (occluded, blurred, or visible) and road status (whether the signs is on the road being traveled or on a side road), see figure 1 [6, 19].

Dataset consists of two subsets. We divided subset1 into training and validation sets consisting of 1970 (20%) images and 7903 (80%), respectively.

In Table II, the results of the implementation of the R-CNN detector for detecting traffic signs for speed limits are presented.

Table I
THE RESULTS OF THE IMPLEMENTATION OF THE R-CNN DETECTOR

Traffic sign	Precision	Recall
Speed limits	0.833± 0.01	0.908

In the future, it is planned to conduct experimental evaluations for all sign classes from the used database and expand the test database of images.

In the process of working on a specific task of building intelligent computer vision systems have been obtained acceptable results. However, identified problems and constraints of all tasks in computer vision. Here they are:

- the lack of a unified algorithmic approaches, a large number of individual solutions of a problem;
- testing and analysis of test results on narrow bases;
- the success of the solution depends on the researcher's experience that solves it.

Summarise the foregoing and formulate the problems arising in the process of building computer vision systems in particular, and in the construction of intelligent systems.

V. Computer vision systems: problems and goals

Let us generalize what was said above to the general case of constructing an intelligent system of computer vision. We have successfully solved a particular narrow problem.

Until now, research teams have only implemented a method that they believe has potential or perhaps tested a few solutions. This statement is true both for the task of traffic sign detection and for computer vision systems in whole. Without a way to compare performance with other systems, it is not clear which approaches work best. There are a large number of disparate databases, but there is **no knowledge base**.

At present time computer vision systems can perform many tasks and has many real-life applications in a wide range of areas including optical character recognition, face detection, emotion detection, object recognition, vision-based biometrics, identity verification through Iris code, login with fingerprint or face, 3D modeling, special effects, etc. This is a very active research area, and rapidly changing. There are many examples of current computer vision systems. Many software applications have been developed in the last five years.

At present, there are fenced borders, which do not allow speaking about the complex solution of computer vision task. Computer vision system is a particular case intelligent systems in general. The list of limitations in the construction of intelligent systems is large. Let's name only a few of them:

1) There is **no general approach for choosing a method, methodology or algorithm** for solving any computer vision problem. The **experience and knowledge of computer vision specialists are crucial**.
2) Many methods have been proposed for solving problems of computer vision. More sophisticated approaches consist of several computer vision stages. These methods are based on traditional algorithms. Recent works suggest combining different approaches for increasing performance. More often, an increase in system efficiency is achieved by modifying existing methods/algorithms or a combination thereof. Fundamentally new and breakthrough solutions are rarely offered. There are **no universal methods/algorithm for solving wide class of problems**. It follows that there is **no software tools for solving a wide class of computer vision problems**.
3) **Practical vision systems** need to be compact and low cost, but given the previous comments, they are **large-scale and costly**.

The difficulty of choosing an approach for building computer vision system is shown in Table III.

Today there are already a lot of modern technologies of designing intelligent systems. But they can solve not all problems mentioned above. The following statements, formulated for intelligent systems, are also valid for computer vision systems. These are problems to be solved.

1. **The development of intelligent system theory** in general and the theory of computer vision systems in particular is an extremely important task.

2. It is necessary to have a general (complex, integrated, holistic) **technology of designing intelligent systems** for improving effectiveness of designing intelligent systems. Compatibility of such design solutions means compatibility of different kinds of intelligent system components which, in general, can be the products of developing by different and independent developer teams. Intelligent systems based on such technology should be flexible, easy modified, reconfigurable.

3. Development of **different kinds of intelligent system**

Table II
THE DIFFICULTY OF CHOOSING AN APPROACH FOR BUILDING COMPUTER VISION SYSTEM

Component/stage of computer vision system	**Technical/algorithmic solution**
Visual sensors	**Sensors** Surveillance camera Satellite UAV **Registration method** Monocular Stereo Infra-red Thermal **Video or Image** Video **static or moving** **etc.**
Information extraction	Color features Shape features Temporal features Texture features, **etc.**
Decision-making	Artificial neural networks Support Vector Machine Fuzzy logic Decision trees Bayes classifier Wavelets Hidden Markov Models, **etc.**

components which, in general, can be the products of developing by different and independent developer teams [20].

VI. CONCLUDING REMARKS

In this paper, R-CNN detector is studied as alternative to the typical algorithms of traffic sign detection. Deep learning have improved object detection accuracy. Public traffic signs database was used for experiments.

But this is a special case of solving a narrow problem. Despite the existing advances the problems that exist and which need to be solved in the future in the field of computer vision system design were formulated.

REFERENCES

[1] F. Lu, S. Lee and R. Satzoda, "Embedded Computing Framework for Vision-Based Real-Time Surround Threat Analysis and Driver Assistance," IEEE CVPR - Embedded Vision Workshop, (Best Demo Award), 2016.

[2] A. MØgelmose, M. Trivedi, and T. Moeslund, "Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 4, 2012.

[3] H. Fleyeh, M. Dougherty, D. Aenugula, S. Baddam, "Invariant Road Sign Recognition with Fuzzy ARTMAP and Zernike Moments," Intelligent Vehicles Symposium, 2007 IEEE. pp. 31–36, 2007.

[4] Y.Y. Nguwi, A.Z. Kouzani, "Detection and classification of road signs in natural environments," Neural Comput. Appl. 17, 265–289, April 2008.

[5] Q. Zhang, S.i. Kamata, "Automatic road sign detection method based on Color Barycenters Hexagon model," ICPR 2008.

[6] F. Larsson and M. Felsberg, "Using Fourier descriptors and spatial models for traffic sign recognition," Proc. Image Anal., 2011, pp. 238–249.

[7] Yuan Li, Andreas Møgelmose, and Mohan M. Trivedi, "Pushing the "Speed Limit": High-Accuracy U.S. Traffic Sign Recognition with Convolutional Neural Networks," IEEE Transactions on Intelligent Vehicles, 2016.

[8] MathWorks website [Online]. Available: https://www.mathworks.com

[9] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, 86(11):2278–2324, November 1998.

[10] Large Scale Visual Recognition Challenge (ILSVRC) website [Online]. Available: http://www.image-net.org/challenges/LSVRC/

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in neural information processing systems, 2012.

[12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, "Going deeper with convolutions". [Online]. Available: https://arxiv.org/pdf/1409.4842.pdf

[13] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation Tech report (v5)". [Online]. Available: https://arxiv.org/pdf/1311.2524.pdf.

[14] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The german traffic sign recognition benchmark: A multi-class classification competition," Proc. IJCNN, 2011, pp. 1453–1460. [Online]. Available: http:// benchmark.ini.rub.de/?section=gtsrb

[15] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," Neural Netw., vol. 32, pp. 323–332, Aug. 2012. [Online]. Available: http:// www.sciencedirect.com/science/article/pii/S0893608012000457

[16] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," Machine Vision and Applications. New York: Springer-Verlag, Dec. 2011, pp. 1–15. [Online]. Available: http://dx.doi.org/10.1007/s00138-011-0391-3.

[17] C. Grigorescu, and N. Petkov, "Distance sets for shape filters and shape recognition," IEEE Trans. Image Process., vol. 12, no. 10, pp. 1274–1286, Oct. 2003.

[18] R. Belaroussi, P. Foucher, J. Tarel, B. Soheilian, P. Charbonnier, and N. Paparoditis, "Road sign detection in images: A case study," Proc. ICPR, Istanbul, Turkey, 2010, pp. 484–488.

[19] Computer Vision Laboratory (CVL) website [Online]. Available: http://www.cvl.isy.liu.se/research/datasets/traffic-signs-dataset/

[20] V. V. Golenkov, "Ontology-based Design of Intelligent Systems" Proc. of the Intenational Conference "Open Semantic Technologies for Intelligent Systems (OSTIS-2017)", Minsk, 16 - 18 February 2017, BSUIR – pp. 37–56, 2017.

ДЕТЕКЦИЯ ДОРОЖНЫХ ЗНАКОВ И ПРОБЛЕМЫ В ОБЛАСТИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Лукашевич М. БГУИР, Минск, Беларусь

Детекция объектов является типовой задачей компьютерного зрения. В данной статье представлены некоторые результаты реализации системы распознавания дорожных знаков. Для детекции использован R-CNN детектор. Ключевое внимание было уделено детекции дорожных знаков ограничения скорости. R-CNN детектор, основанный на глубинном обучении, - это простая и подходящая модель для распознавания дорожных знаков. Этот подход комбинирует низкоуровневые признаки изображения с высокоуровевым контекстом, включающим детекцию объектов и классификатор. Несмотря на существующие подвижки в компьютерном зрении в статье рассмотрены существующие и требующие решения проблемы построения систем компьютерного зрения.

Automatic recognition of consultants on video records

Rozaliev V.L.
CAD Department
Volgograd State Technical University
Volgograd, Russia
vladimir.rozaliev@gmail.com

Alekseev A.V.
CAD Department
Volgograd State Technical University
Volgograd, Russia
alekseev.yeskela@gmail.com

Ulyev A.D.
CAD Department
Volgograd State Technical University
Volgograd, Russia
ulyev-ad@yandex.ru

Orlova Y.A.
SAS Department
Volgograd State Technical University
Volgograd, Russia
yulia.orlova@gmail.com

Petrovsky A.B.
Laboratory of methods and decision support systems Federal Research Center Computer Sciences and Control, RAS
Moscow, Russia
pab@isa.ru

Zaboleeva-Zotova A.V.
Laboratory of methods and decision support systems Federal Research Center Computer Sciences and Control, RAS
Moscow, Russia
zabzot@gmail.com

***Abstract*—This article presents method for recognizing consultants in showroom based on cascade of neural networks. The cascade consists of two networks - high-performance detector and refining module with recognition of pose. A brief review of the analog systems is given. The description of the proposed method is presented, the obtained results and ways of improvement are shown.**

***Keywords*—neural network, artificial intelligence, recognition of human pose, analysis of video stream**

I. Introduction

The modern era is characterized by a transition from the economy of producers to the economy of consumers. In the conditions of toughening competition in the sphere of trade and rendering services, client-oriented services acquire special importance.

The main problem of introducing such services is the human factor, control of which is problematic due to the lack of ready-made software products.

Ensuring the proper quality of service delivery becomes the main objective of the market strategy for business development.

To improve the quality of service, it is proposed to develop and implement a software product to monitor the activities of consultant salesmen through the analysis of their work with the use of equipment for video fixing [11].

The basic principle of the software product is based on a neural network for detecting a person on a frame from a video stream, and also on the algorithm "Pose Estimation" [3], whose main function is to recognize the human pose.

II. The proposed methodology

To solve the problem, we propose to use a cascade of two neural networks:

- fast detector Yolo [1];
- Neural network for recognition of the pose [2] of the consultant, the general scheme is shown in figure 1.

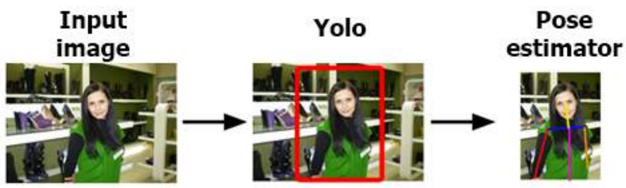


Figure 1. The proposed method for recognizing consultants.

III. The first stage, the use of the "YOLO" neural network

The first stage of the cascade is the neural network [1]. The main advantage of Yolo is the speed of work, it allows you to achieve a speed of 60-90 frames per second, while maintaining a sufficiently high quality of work. The input data of the algorithm "Yolo" is the image, and the output data - rectangles (wireframe figures) that limit the found objects (this network can find other objects, but we are only interested in people). These values (regions) are transmitted to the second neural network for clarification.

The main task solved by the Yolo network in our approach is a quick determination of the fact of having people in the frame, so as not to start a slower stage on empty frames. Advantage in the speed of work is achieved due to the approach classifiers that are different from the classical networks. In the case of

Yolo, a picture is placed at the entrance, divided into small regions with the probabilities of finding objects in the region. The result of the network is shown in figure 2.

Figure 2. The result of the work of the Yolo neural network.

IV. The second stage, the use of the neural network "Pose Estimator".

The main task of the neural network is to establish a person's pose through a nonparametric representation called the Part Affinity Fields (PAFs) by developers, to further determine the location of the seller's uniform of the consultant (branded T-shirt, cap, etc.) [4] [5] [6].

The main advantage of the neural network is the high quality of the work. The main drawback is the demanding nature of the neural network "PoseEstimator" for the technical characteristics of the equipment used. In the absence of an appropriate technical base for this network, there is a sharp increase in the processing time.

Input data for the algorithm "PoseEstimator" is a graphic image of the sales consultant, on the output - an image with the selected parts of the human body. The result of this network can be seen in figure 3.

Figure 3. The result of the Pose Estimator neural network.

V. The third stage, the determination of the dominant colour in the uniform section

The main task of this stage is to establish a dominant colour in the area of the uniform of a person to determine it in the group of sales consultants. Within the framework of this algorithm, an image from the "Pose Estimator" with the tops of the human body parts is input. On the basis of which there is a selection of the necessary clothing of a person. An example is shown at figure 4.

Figure 4. Allocation of the seller's clothing segment.

To implement the definition of dominant colour in an established area, there are several methods: determining the ratio of a pixel to a given set of colours and clustering by the k-means method.

In the first method, the image is converted to HSV colour space, after which all pixels of the image are analyzed and based on the Hue, Saturation, Value data, the colour is set.

The idea of the k-means method is to minimize the total quadratic deviation of the cluster points from the centers. At the first stage, you select points (three-dimensional RGB space) and determine whether each point belongs to this or that center. Then at each stage, the centers are redefined until a single center is found. An example of clustering is shown at figure 5 [7].

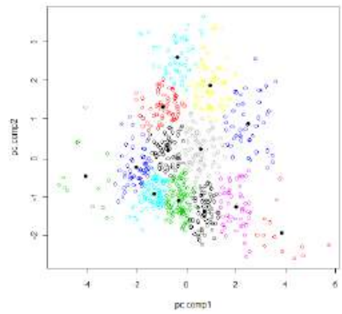


Figure 5. Example of clustering.

VI. Training the appearance of sales consultants

To determine only the colour of a person's clothing is not enough to classify him as a sales consultant group. It is necessary to take into account the conditions of the difference in the illumination of the room at different times of the day, as well as the likelihood that there may be clarified and dark areas in the room. Thus, the recognized colour of the shape can vary [10].

To solve this problem, it is necessary to teach the system all possible colors that can be "read" from the clothing of the seller-consultant. The operator of the software product at the start of work with the program must manually select on the frame of the seller-consultant, which will arbitrarily move around the room.

In addition, to determine the degree of colour deviation from the established colour of the seller-consultant, the colour difference formula is used, shown at formula 1, which will allow numerically to express the difference between the two colours in colorimetry. On the basis of the data obtained, it can be concluded on how much the colour is "close" to the colour established for the seller-consultant.

$$\Delta E^*_{ab} = \sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2} \qquad (1)$$

VII. Overview of analogues

It should be noted that the finished software products that allow to solve the problem discussed in this article are not present. Similar software products perform only part of tasks [12].

The simplest example of intelligent video surveillance is motion detection. One detector can replace several video surveillance operators. And in the 2000s, the first video analytics systems began to appear, capable of recognizing objects and events in the frame. Most of the solutions work with face recognition technologies. Solutions in this area include Apple, Facebook, Google, Intel, Microsoft and other technology giants. Surveillance systems with automatic passenger identification are installed in 22 US airports. In Australia, they are developing a biometric system of face recognition and fingerprinting within a program designed to automate passport and customs control. An interesting project of NTechLab company showed a system capable of real-time recognition of sex, age and emotions using the image from a video camera. The system is able to evaluate the audience's reaction in real time, so you can identify the emotions that visitors experience during presentations or broadcasts of advertising messages. All NTechLab projects are built on self-learning neural networks. In our system, we do not yet use data on a person's face. We plan to process this information at the next stages of the project development.

In other systems, the object tracking function is used - tracking. The operation of the tracking modules is related to the operation of the motion detector. To construct the trajectories of the movement, a sequential analysis of each frame is carried out, on which moving objects are present. In the general case, several moving objects can be present in one frame, so the program needs not only to construct trajectories, but also to distinguish objects and their movements. The simplest implementation of tracking considers two frames and builds trajectories along them. First, the movements on the current and previous frame are marked, then, by analyzing the speed, the direction of movement of objects, and also their sizes, the probabilities of the transition of objects from one point of the trajectory of the previous frame to another point of the current are calculated. The most probable movements are assigned to each object and added to the trajectory. Objects in the frame can move in different ways: their trajectories may intersect, they can disappear and arise again. To improve the accuracy of tracking, some manufacturers use the technology of sequence analysis and continuous post-processing of the results obtained. The program builds graphs - it analyzes the transitions of objects from one state to another. In order to understand which object the movement corresponds to, the speeds and directions of motion, position, color characteristics are also analyzed. As a result, a set of the most probable displacements of the object is formed, forming a trajectory. We have planned to use this approach in our system.

Another analogue of our system - GPS-trackers. These systems work based on the definition of geolocation. To implement this solution, each employee must be equipped with a separate GPS tracker, the data from which will be sent to the server at some interval. However, this solution has a number of drawbacks:

1) The solution is not cost-effective, since it is necessary to purchase GPS trackers for all personnel.
2) We can't exclude the situation in which the seller can give his GPS-tracker to a partner to deceive the system.
3) Such a solution is not universal. When identifying sales consultants through the camera, it is possible to expand the functionality, determine the level and time of interaction of the seller with the buyer, and much more.

Also, analogs include systems for counting the number of visitors on a video stream. These systems also have a number of shortcomings, the main one of which is the impossibility of identifying sales consultants and the quality of their services. An example of the work of such products is shown at figure 6.

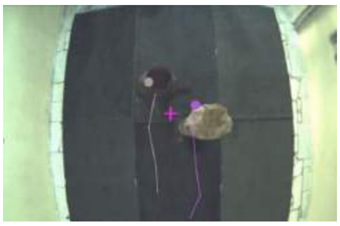


Figure 6. Example of a program for counting the number of visitors.

VIII. Conclusion.

In order to improve the technological process of detecting the seller's consultant, it is possible to develop additional functionality.

To more accurately determine the seller's consultant, it is possible to analyze several elements of the uniform at once (for example, a yellow T-shirt and black pants).

In addition, it is possible to search for the company logo on the uniform, the location of which will allow us to identify with confidence the person as the seller-consultant.

Another factor that allows to detect the seller, can serve as a definition of behavior, characteristic for the seller-consultant. To solve this problem, you will need to create another neural network.

Thus, the developed software product, consisting of a cascade of neural networks YOLO and Pose Estimator will allow to qualitatively improve the work of the seller-consultant and, as a result, improve the client-oriented business.

Figure 7 shows the input image for the system, in figure 8 - the obtained result.

This work is a contin of the work [8] [9], where the features and possibilities of determining the post-sense of its semantic distinctive feature were considered.

This work was partially supported by RFBR (grants 16-07-00407, 16-47-340320, 16-07-00453, 18-07-00220).

Figure 7. Image to be input.

Figure 8. The image obtained as a result of the software product.

REFERENCES

[1] Redmon, J., Farhadi A. YOLO9000: luchshe, bystree, sil'nee [YOLO9000: Better, Faster, Stronger]. Retrieved from http://arxiv.org/abs/1612.08242

[2] A.A. Shpirko, V.L. Rozaliev, J.A. Orlova, A.V. Alekseev Avtomatizatsiya postroeniya vektornoi modeli tela cheloveka [Automation of constructing a vector model of a human body.]. Izvestiya VolgGTU. Seriya "Aktual'nye problemy upravleniya, vychislitel'noi tekhniki i informatiki v tekhnicheskikh sistemakh" [Izvestiya VSTU. A series "Actual problems of control, computer technology and informatics in technical systems"], 2013, vol. 7, pp. 67-71/

[3] Cao Z., Simon T., Wei S.-E., Sheikh Y. Otsenka pozy v real'nom vremeni Multi-persony 2D s ispol'zovaniem polei blizosti. [Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.]. 2016, Retrieved from http://arxiv.org/abs/1611.08050

[4] Zhe Cao Otsenka 2D-otsenki v real'nom vremeni s ispol'zovaniem otdel'nykh polei. [Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.]. 2017, Retrieved from http://www.ri.cmu.edu/wp-content/uploads/2017/04/thesis.pdf

[5] U. Iqbal, J. Gall Otsenka individual'nosti cheloveka s uchastiem mestnykh assotsiatsii. [Multi-person pose estimation with local joint-to-person associations.]. 2016, Retrieved from https://arxiv.org/pdf/1608.08526.pdf

[6] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, B. Schiele Bolee glubokii srez: bolee glubokaya, bolee sil'naya i bolee bystraya model' otsenki pozy dlya neskol'kikh chelovek. [Deepercut: A deeper, stronger, and faster multi-person pose estimation model.]. 2016, Retrieved from https://arxiv.org/pdf/1605.03170.pdf

[7] M.S.Shmakov, A.A. Tovmasyan Analiz tsvetovoi palitry izobrazhenii dlya opredeleniya preobladayushchikh tsvetovykh tonov [Analysis of the color palette of images to determine the prevailing color tones]. Raboty Belorusskogo gosudarstvennogo tekhnicheskogo universiteta. [Works of the Belarusian State Technical University], 2016, vol. 9

[8] V.L.Rozaliev, Y.A.Orlova Opredelenie dvizhenii i pozy dlya identifikatsii emotsional'nykh reaktsii cheloveka. [Recognition of gesture and poses for the definition of human emotions]. 11-ya Mezhdunarodnaya konferentsiya po raspoznavaniyu obrazov i analizu izobrazhenii: novye informatsionnye tekhnologii (PRIA-11-2013), Samara, 23-28 sentyabrya 2013 g.: Trudy konferentsii [11th International Conference of Pattern Recognition and Image Analysis: New Information Technologies (PRIA-11-2013), Samara, September 23-28, 2013 : Conference Proceedings], 2013, vol. 2, pp. 713-716

[9] A.S.Bobkov, V.L.Rozaliev Fazzifikatsiya dannykh, opisyvayushchikh dvizhenie cheloveka. [Fuzzification of data describing the movement of a person]. Otkrytye semanticheskie tekhnologii dlya proektirovaniya intellektual'nykh sistem (OSTIS-2011): mater. stazher. nauchno-tekhnich. konf. (Minsk, 10-12 fevralya 2011 g.) [Open semantic technologies for the design of intelligent systems (OSTIS-2011) : mater. intern. scientific-techn. conf. (Minsk, Feb. 10-12. 2011)], 2011, pp. 483-486/

[10] M.D.Khorunsjiy Metod kolichestvennoi otsenki tsvetovykh razlichii v vospriyatii tsifrovykh izobrazhenii. [The method of quantitative estimation of color differences in the perception of digital images.]. Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki i optiki. [Scientific and technical herald of information technologies, mechanics and optics], 2008

[11] Angjoo Kanazawa, Michael J. Black, David W. Jacobs, Jitendra Malik Kompleksnoe vosstanovlenie formy cheloveka i pozy [End-to-end Recovery of Human Shape and Pose]. Retrieved from https://www.researchgate.net/publication/321902575_End-to-end_Recovery_of_Human_Shape_and_Pose?discoverMore=1

[12] Korobkov A. Otslezhivanie ob'ektov v videopotoke. Metody postroeniya traektorii. [Tracking objects in the video stream. Methods for plotting trajectorie]. Sistemy bezopasnosti. [Security], 2014, val. 3

АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ КОНСУЛЬТАНТОВ НА ВИДЕОЗАПИСИ

Розалиев В.Л., Алексеев А.В.,
Ульев А.Д.,Орлова Ю.А.,
Петровский А.Б., Заболеева-Зотова А.В.

В данной статье представлен метод распознавания консультантов в торговом зале на основе каскада нейронных сетей. Каскад состоит из двух сетей - высокопроизводительного детектора и уточняющего модуля с распознаванием позы. Представлено описание предлагаемой методики, показаны полученные результаты и пути улучшения.

Recognition of human emotional state using the SVM and ANN algorithms

Astakhov Dmitriy
Volgograd State Technical University
Volgogorad, Russia
astakhov.d.a@gmail.com

Kataev Alexander
Volgograd State Technical University
Volgogorad, Russia
alexander.kataev@gmail.com

***Abstract*—The work compares the results of the work of the machine learning algorithm SVM, ANN on a set of data from the normalized key points and the distances between them. The process of preparing the training base and normalization algorithms is described.**

***Keywords*—system of recognition of emotions, key points, machine learning, face recognition, ann, svm, emotions**

I. Introduction

Long before speech arose, interaction between people was realized through non-verbal communication. Change in facial expression has become one of the most significant means of conveying the emotions and intentions of a person.

In recent years, the task of automatically analyzing the emotional state of a person has attracted a large number of researchers.

In particular, in the work led by Zachelova-Zotova A.V. the questions of definition of emotional reactions of the person on mimicry, nonverbal movements and a voice are considered [1]. The task of recognizing dynamic gestures of a person is considered in the works of Devyatkov V.V. and Alfimtseva A.N. [2]. In his doctoral dissertation Dementienko V.V. the system of automatic prediction of the driver's falling asleep due to his blinking and movement of the eyes is considered [3]. In the work of Zeifeng Shan, an algorithm for detecting emotions based on local binary patterns is considered [4].

Recently, machine learning methods such as SVM and ANN have often been used to define emotions, as can be seen from such works as [5]. The purpose of this study is to compare the work of these methods with different sets of data.

II. Determination of the emotional state of a person

It is known that the expression of emotions can be very diverse and vary depending on individual characteristics, as well as on the situational context. The cultural context plays an important role when it comes to emotions that do not belong to the category of basic ones, since the methods of expressing complex emotional states adopted in a certain community are different. If you observe a person, you can see that most emotions are rapidly changing each other and making it difficult to recognize [6]. In addition, manifestations that correspond to "mixed" emotions are difficult to recognize, because behind them lies the whole complex of feelings experienced by a person. Another factor that affects the effectiveness of the recognition of emotions is the possibility of faking emotional behavior. Cultural requirements, personal perceptions of admissible or any considerations arising from a person's perception of the actual situation (the desire to hide something or demonstrate a feeling that is not really there) - these and other factors affect the expression of a person's emotions. Automatic identification of images (text, sound, face, person, objects, etc.) with the help of a computer is one of the most important directions in the development of artificial intelligence technologies, which makes it possible to give the key to understanding the features of the work of the human intellect. Research methods of automatic recognition of emotions allows you to give the computer the ability to assess the mood of a person, for this purpose, the algorithm uses the recognition of emotions, presented in "Fig. 1".

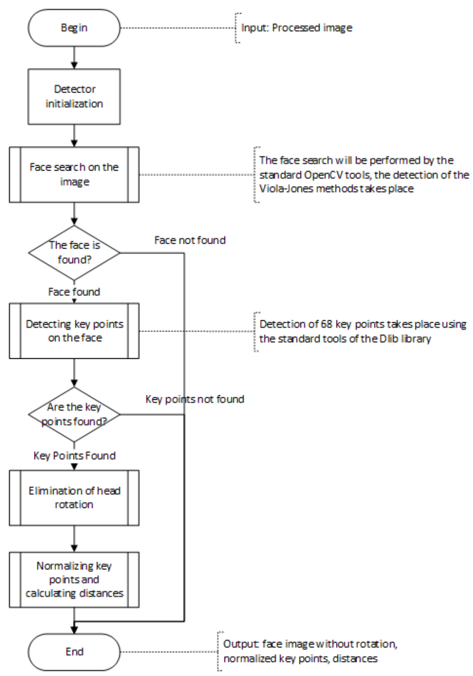


Figure 1. The main algorithm for recognizing emotions.

III. Preparing a sample for training

For learning machine learning algorithms, the Extended Cohn-Kanade Database (CK +) sample [10] was used, consisting of 11,061 photographs, with a resolution of 640 × 490 pixels, in * .png format. In the CK + database 623 structures with emotions are marked out. The sample is divided into 8 classes of emotions: 1 - anger, 2 - contempt, 3 - disgust, 4 - fear, 5 - happiness, 6 - sadness, 7 - surprise, 8 - normal. In the future, for brevity, the classes of emotions will be denoted by the corresponding numbers.

The original database was reworked in such a way as to minimize recognition errors, for this purpose, an image with the most "expressed" emotion was selected for each emotion and a person and a mirror image was made for each image to increase the sample size. The resulting volume of the processed database produced 2016 photographs. “Fig. 2” shows the distribution of the volume of images for each class.

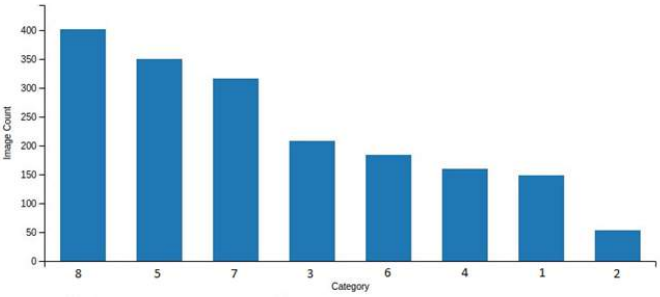


Figure 2. The diagram of the distribution of the data volume by classes.

Uneven distribution of images by classes is associated with the limited sample size for each class in the original CK + database, classes 5, 7, 8 have more data in the source database than classes 1, 2, 3, 4, 6. Example of images used in the database is shown in “Fig. 3”.

Figure 3. An example of images from the database, the classes are located on the left to the right (the leftmost image corresponds to class 1, the rightmost to the 8th class) (The Extended Cohn-Kanade Database).

For further work, the sample was normalized. By normalization is meant - an image in which the person is located without turns, inclinations and 95% of the image is the person itself. An example of normalized images for each emotion class is shown in “Fig. 4”. The image size after normalization is not fixed, it ranges from 201 x 199 pixels to 306 x 275 pixels, type of images * .png.

A. Facial coding system

For the encoding of facial movements, the systems of FACS and EmFACS are used. Facial Action Coding System (FACS) - is a system for the taxonomy of human facial expressions [7]. This standard is generally accepted for the systematic classification of the physical expression of emotions [8].

Figure 4. An example of images from a normalized base, classes are located on the left to the right (the leftmost image corresponds to class 1, the rightmost to the 8th class).

Emotional Facial Action Coding System (EmFACS) - this system is considered only the coding of facial movements associated with emotions [9].

These systems were developed by Paul Ekman and Wallace Friesen in 1978.

With the use of FACS, it is possible to manually code any practical, anatomically possible facial expression, constructing it from the actions of specific units of action and the time required by them to reproduce a particular facial expression. The FACS determines the units of action that cut or relax one, or more muscles.

This system allows you to align which key points are involved in facial movements. An example of the encoding of motion movements is shown in “Fig. 5”.

Figure 5. Encoding facial movements in the FACS system.

B. Key Points

For the process of highlighting the key points on the face, a third-party Dlib library was used, which allows 68 face key points to be found on the image “Fig. 6”.

In accordance with the CLRD, for the training from 68 key points, 44 points describing the facial movements were selected, for this purpose points describing the contour of the face (points 0 - 16) were excluded, which do not affect the description of facial movements in the emFACS system. The next step for each image from the sample is two sets of data: normalized key points (“Fig. **??**”) and distances between the key points (“Fig. **??**”).

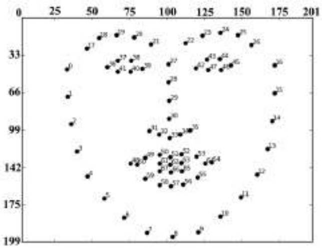
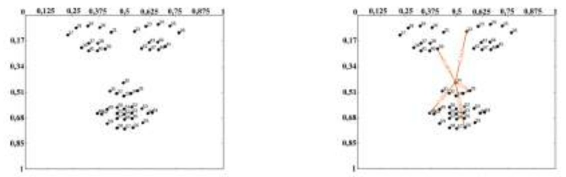


Figure 6. Example of detecting key points by the dlib library.



a) b)

Figure 7. Example of data used, figure (a) - normalized key points, figure (b) - distance between key points

To the normalized face image, methods for detecting key points are applied, resulting in the output of 68 key points in the coordinate system of the normalized image. For normalization of data, all coordinates are given to the range [0, 1].

Calculation of normalized key points was carried out according to the formula:

$$x_i' = \frac{x_i}{\text{width}}; y_i' = \frac{y_i}{\text{height}} \quad (1)$$

Calculation of distances between the key points was carried out according to the formula:

$$d_i = \sqrt{(x_c - x_i') + (y_c - y_i')} \quad (2)$$

C. System Training

The training was done using the built-in machine learning in the OpenCV library.

To train the algorithms, the prepared database was divided into two parts: training and test. For testing, 15% of the total sample size was selected, which is 272 images, the training base consists of 1744 images.

For the training of SVM, the type C_SVC was used, which makes it possible to classify into 8 classes using a multiplier C with a value of 0.5. The type of the kernel was chosen to be polynomial with parameters: a level equal to 0.001, coef0 equal to 0.1, gamma equal to 0.00015.

The network architecture for learning the ANN method on normalized key points consists of 4 layers: an input layer of 44 neurons, two intermediate layers of 88 neurons each, and an output layer of 8 neurons. The sigmoidal activation function is used, the network is trained by the method of back propagation of the error with the scale parameter equal to 0.1. The maximum number of iterations was set to 1000. The selection of parameters for both methods of machine learning is empirical.

IV. Preparing a sample for training

The work of the machine learning methods SVM and ANN was carried out on two sets of data: normalized key points and distances between the key points. The results of the experiments carried out for each method of machine learning with different data sets.

Table I
THE RESULT OF TESTING THE SVM METHOD ON NORMALIZED KEY POINTS

Expected emotion	Recognized emotion							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	**30.62**	0.00	22.50	0.00	0.00	31.25	0.00	15.62
2	3.70	**0.00**	3.70	0.00	16.66	11.11	0.00	64.81
3	0.00	0.00	**31.77**	0.00	17.77	14.77	0.00	35.69
4	0.00	0.00	0.00	**37.07**	17.53	0.00	12.69	32.69
5	0.00	0.00	3.33	0.00	**93.33**	0.00	0.00	3.33
6	0.00	0.00	9.37	6.13	9.37	**48.88**	0.00	26.25
7	0.00	0.00	0.00	4.16	0.00	4.17	**91.67**	0.00
8	0.00	0.00	0.56	1.11	4.33	1.67	0.00	**90.94**

Analyzing the data obtained during the testing ("Tab. I"), we can conclude that the SVM method tested on a set of key points showed low accuracy for this classification problem. This is due to the fact that the training sample does not have a uniform distribution of the number of images by classes, it can be observed that in most classes the percentage of recognized emotions refers to grades 5, 7 and 8, and the parameters chosen in training could not be found to be the most optimal.

Table II
THE RESULT OF TESTING THE SVM METHOD ON DISTANCE DATA

Expected emotion	Recognized emotion							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	**28.13**	0.00	15.63	0.00	3.13	31.25	0.0	21.88
2	1.85	**0.00**	12.96	1.85	1.85	1.11	0.00	70.37
3	0.00	0.00	**50.00**	0.00	50.00	0.00	0.00	100
4	0.00	0.00	0.00	**69.23**	7.69	19.23	3.85	0.00
5	3.33	0.00	6.67	3.33	**76.67**	3.33	0.00	6.67
6	3.13	0.00	15.63	6.25	3.13	**56.25**	0.00	15.63
7	0.00	0.00	0.00	12.50	12.50	4.17	**70.83**	0.00
8	1.72	0.00	8.33	1.39	2.94	1.33	0.00	**86.28**

The SVM method tested on a set of distances between key points ("Tab. II"), showed itself the same way as when testing at key points, low accuracy for the task of classifying emotions. This is due to the fact that the training sample does not have a uniform distribution of the number of images by classes, it can be observed that in most classes the percentage of recognized emotions refers to 4, 5, 7 and 8 classes. But, nevertheless, the result of training at these distances proved to be more effective than at these key points by 2.17

The ANN method tested on a set of key points ("Tab. III") showed the accuracy of the recognition of emotions equal to 53.5%, therefore, this approach may be applicable when working with real data. The inaccuracy of the classification

Table III
THE RESULT OF ANN TESTING ON NORMALIZED COORDINATES

Expected emotion	Recognized emotion							
	1	2	3	4	5	6	7	8
1	**37.50**	0.00	3.12	0.00	9.37	21.87	28.12	0.00
2	0.00	**27.78**	0.00	1.85	7.40	33.30	29.62	0.00
3	0.00	0.00	**21.30**	0.00	0.00	34.80	21.70	22.20
4	0.00	3.84	0.00	**54.10**	3.84	5.380	23.80	9.04
5	0.00	0.00	0.00	0.00	**50.00**	16.40	14.50	19.10
6	0.00	0.00	0.00	0.00	0.00	**71.87**	8.12	20.00
7	0.00	0.00	0.00	3.33	0.00	4.20	**77.91**	14.36
8	0.00	1.38	0.00	4.17	2.70	0.00	3.61	**88.14**

is due to the fact that the training sample does not have a uniform distribution of the number of images by classes, and the parameters chosen in training could not be found to be the most optimal.

Table IV
THE RESULT OF TESTING THE ANN METHOD ON DISTANCE DATA

Expected emotion	Recognized emotion							
	1	2	3	4	5	6	7	8
1	**41.87**	0.00	0.00	0.00	6.25	12.5	18.75	20.62
2	0.00	**38.51**	0.00	1.85	1.85	14.81	11.11	21.85
3	0.00	0.00	**50.00**	0.00	0.00	0.00	0.00	50.00
4	0.00	0.00	0.00	**58.46**	0.00	11.53	16.92	13.07
5	0.00	0.00	0.00	0.00	**60.00**	0.00	13.33	26.67
6	0.00	0.00	0.00	0.00	3.12	**60.62**	9.37	26.87
7	0.00	0.00	0.00	0.00	0.00	0.00	**87.50**	12.50
8	0.00	1.38	0.00	0.00	0.00	4.16	16.66	**77.77**

The ANN method tested on a sample from the distances between the key points (“Tab. IV”) showed a classification accuracy of 53.5%, which is 6.5% better than on data with key points. Therefore, this approach may be applicable when working with real data. The inaccuracy of the classification is due to the fact that the training sample does not have a uniform distribution of the number of images by classes, and the parameters chosen during training could not be found to be the most optimal.

Table V presents a comparative analysis of the results obtained.

Table V
SUMMARY TABLE OF TEST RESULTS

Method	Dataset	Accuracy of recognition,%
SVM	Key Points	52.50
	Distance	54.67
ANN	Key Points	53.50
	Distance	60.00

Thus, analyzing the data obtained for each test example, it can be concluded that SVM and ANN methods showed very close results in the accuracy of the classification of emotions. But the ANN method showed greater recognition accuracy than the SVM method at a distance of 5.33%, and at these key points by 1%.

As you can see, the accuracy of the classification increases, if you use the data based on the distance between the key points for learning. For the SVM method, the accuracy increased by 2.67%, and for the ANN method the accuracy increased by 6.50%.

CONCLUSION

In the paper, the machine learning methods ANN and SVM were tested on a set of data from normalized key points and the distances between them. Both methods have an acceptable detection level. However, during the research it was revealed that the ANN method on the set of distances showed the best result.

Both of these methods can be used in the interface of an intelligent system to create conditions for natural and intuitive human-machine interaction.

REFERENCES

[1] The task of creating a system of automated recognition of emotions / A. V. Zachelova-Zotova [and others] // Open semantic technologies for the design of intelligent systems: materials of the International. scientific-techn. Conf. OSTIS-2012 (Minsk, February 16-18, 2012) / BSUIR. - Minsk, 2012. - P. 347-350.

[2] Devyatkov, V. V. Recognition of manipulative gestures / V. V. Devyutkin, A. N. Alfimtsev // Bulletin of the Bauman Moscow State Technical University. N. E. Bauman. - 2007. - vol. 68, i. 3. - P. 56-75.

[3] Dementienko, V. V. Physical principles of constructing systems for the safe monitoring of the human operator status: the author’s abstract. dis. ... Dr. techn. Sciences / V. V. Dementienko. - Moscow, 2010. - 38 p.

[4] Caifeng Shan, Shaogang Gong, Peter W. McOwan .: Facial expression recognition based on Local Binary Patterns: A comprehensive study. Image and Vision Computing (2008).

[5] Enikilopov S. N., Kusnetsova Yu. M. .: The Task of Recognition of Violent Situations Using the Automatic Systems and Methods of Artificial Intelligence. Journal of Psychology and Law, vol. 2, pp. 1- 16. Moscow psychological and pedagogical University, Moscow (2011).

[6] Rozaliev, V. L. Modeling the emotional state of a person on the basis of hybrid methods / V. L. Rozaliev // Software products and systems. - 2010. - No. 2. - P. 141-146.

[7] LeCun, Y. Convolutional Networks for Images, Speech, and Time Series / Y. LeCun, Y. Bengio // The Handbook of Brain Theory and Neural Networks. - USA: MIT Press, 2002. - P. 276-279.

[8] Jizheng Yi, Xia Mao, Lijiang Chen, Yuli Xue, Angelo Compare facial expression by identifying individual differences in facial structure and texture. IET Computer Vision 2014, vol. 8, pp. 429-440. The Institution of Engineering and Technology (2014).

[9] Panagiotis Perakis, Georgios Passalis, Theoharis Theoharis, Ioannis A. Kakadiaris .: 3D Facial Landmark Detection & Face Registration. Department of Informatics and Telecommunications University of Athens, Greece (2010).

[10] The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression / P. Lucey [et. al.] // Proceedings of IEEE CVPR Workshop on Biometrics, 13-18 Jun 2010, San Francisco, CA, USA / IEEE Computer Society. - San Francisco, 2010. - P. 94-101.

РАСПОЗНАВАНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ SVM И ANN

Астахов Д.А., Катаев А.В.
Волгоградский Государственный Технический Университет

В работе сравниваются результаты работы алгоритмом машинного обучения SVM, ANN на наборе данных из нормализованных ключевых точек и дистанций между ними. Проводится описание процесса подготовки базы для обучения и алгоритмов нормализации.

Realtime computer-aided object detection in endoscopic screening

Katsiaryna Halavataya
Intelligent systems dept.
Faculty of Radiophysics and Computer Technologies
Belarussian State University
Minsk, Belarus
katerina-golovataya@yandex.ru

Aliaksandr Kurachkin
Intelligent systems dept.
Faculty of Radiophysics and Computer Technologies
Belarussian State University
Minsk, Belarus
lawliet29@gmail.com

***Abstract*—The project focuses on using computer vision techniques to provide visual support and highlighting when performing endoscopic screening. The system is meant to provide visual clues to the physician performing the screening, highlighting regions of interest and polyps in real time, in order to increase the evaluation accuracy of esophageal, gastric and colon cancers.**

***Keywords*—medical image processing, computer vision, realtime object detection, machine learning**

I. Introduction

Endoscopy is a widespread approach in medical diagnostics that involves a series of procedures for getting visual information from inside the human body in order to examine internal organs. The defining triat of all endoscopic research is the acquisition of colored realistic representative image, image sequence or video as part of the procedure, which greatly helps in early detection, diagnosis and treatment of a wide array of diseases. Most of the modern videoendoscopic system presume the acquisition of high-definition high-framerate video in realtime from the endoscopic camera inside the human body, equipped with additional light source, and allow to control the examination in order to acquire more information about specific areas that might require the physician's attention.

The effectiveness of endoscopic examination is determined by a multitude of factors - patient's preparation for the procedure, equipment quality, physician's skills in working the equipment, ability to spot areas that require further attention and ability to make educated decisions based on the data acquired, sometimes during the procedure in order to control its further flow. Additional equipment features like retroflexion, second view, water/air inflation, etc. also affect the effectiveness of the procedure. However, modern research indicates that physician's own skills play a great role, which led to several medical care institutions of the USA, including U.S. Multisociety Task Force on Colorectal Cancer, American College of Gastroenterology, American Society for Gastrointestinal Endoscopy Task Force on Quality, to propose in 2002 a number of research quality indicators to evaluate physician's effectiveness during the research. In following years, most of these indicators were adopted in other countries, and are nowadays considered a worldwide standard by various health care institutions throught the world.

One of the most important quality indicators of endoscopy screening quality is so-called personal adenoma detection rate, or ADR. A physician's ADR is the proportion of individuals 50 or more years of age undergoing a complete screening colonoscopy who have one or more adenomas, or polyps, detected. A typical value for good ADR is at least 15% for women and 25% for men. It should be noted that while not all detected adenomas may be potentially dangerous, it's still important to spot them during the screening in order to obtain enough information to make sure it poses no threat to the patient. [1]

The paper focuses on development of automated object detection system to aid with screening process. The main constraint on the system is the ability to work in real time – object detection must be performed during the procedure in order to give the physician necessary time to decide if the detected object should be examined further and to minimize overall examination time.

II. Realtime object detection techniques

The problem of detecting and localizing objects on the image is a well-known one in the field of computer vision. The basic definition of the problem can be postulated as follows: given an image that can potentially contain an object of interest, detect said object by making sure that it conforms to specific kinds of visual, spatial, pattern and brightness criteria; produce, if possible, a bounding area (usually bounding rectangle) containing that object; and, finally, produce, if possible, a collection of pixels that belong to the detected object. [2], [3]

The simplest form of object detection may be defined as a binary classifier with image dimensions as bounding rectangle. Algorithm implementation based on this definition provides no information about the actual location of detected object on the image, since the bounding rectangle covers every point that belongs to the image itself. The resulting classes of this binary classification can be treated as "object present on the image" and "object absent on the image" for positive and negative classification, respectively. Most of the times, binary classification output is smoothed, and classification result x is represented as a continious value over the range of $x \in [0; 1]$. With this definition, classification result can be treated as a confidence score of an object being present on the image.

Binary classifiers can be used as a base for creating multi-class classifiers – given n independent binary classifiers that produce $x_i \in [0; 1]$ confidence score of an object of i-th type being present on the image, it's possible to construct a multi-class n-dimensional vector $\vec{x} = (x_1, x_2, ..., x_n)$ with confidence scores for each of the n classes. Choosing a "winning" class or classes can be then implemented as choosing classes with maximum confidence scores, or choosing classes based on a certain confidence score threshold. [2], [4], [5]

One step further is the implementation of actual localization of detected objects on the image. Usually localization presumes finding a rectangular region on the image that encompasses the target object of the classification. It's possible to implement basic localization using a given binary image classifier with a simple brute force approach by iterating over a fixed number of blocks and their adjacency combinations and running classifier with each block to find the region that produces the highest confidence value.

After the actual localization, sometimes it's also necessary to produce the actual uneven boundaries of detected object, i.e. perform classification on a pixel-per-pixel basis in order to find the exact pixels that comprise the classified object on the image. [6], [7]

The most common approach to real-time object detection presumes the usage of features. Features of the image are usually defined as points of the image that represent its characteristic visual triat in a form of some well-known visual abstraction. There is no general criteria that makes any given pixel or any given region of the image a feature – these are usually defined depending on the specific problem and application; however, most commonly feature points are chosen based on sharp brightness difference, thus corresponding to edges and corners of a given image.

The algorithms for computing a set of feature points (or keypoints) of the image are called feature or keypoint detectors. These include most common edge detectors, corner detectors and blob detectors. Some higher-order feature detectors aim to produce keypoints that are scale- and transform-invariant, i.e. don't depend on a relative scale, rotation and skew of their spatial surrounding area. [2], [4], [8]

An actual classifier based on features also requires some notion of feature correspondence or feature comparison. The main idea is that the sought object on an image posesses some features that are similar – to a certain degree – to the set of reference features of a model object that is being detected. Comparison of features is usually performed in a feature metric vector space, and projections from keypoint to feature space are performed by algorithms called feature descriptors.

A feature descriptor f is a projection of any point p^I_{ij} of any image I to an n-dimensional metric vector space F:

$$f(p^I_{ij}) = \vec{v}^I_{ij} \in F \quad (1)$$

Since vector space F is also a metric space, an appropriate metric is defined on it:

$$m : F \times F \to \mathbb{R} \quad (2)$$

Two arbitary points, p^{I_1} of image I_1 and p^{I_2} of image I_2, are considered similar by a feature descriptor (1) if their feature vectors $\vec{v}_1 = f(p^{I_1})$ and $\vec{v}_2 = f(p^{I_2})$ are similar by measure of metric (2), or $m(\vec{v}_1, \vec{v}_2) < t$. The threshold t is selected based on a specific descriptor implementation.

Histogram-based descriptors use histogram analysis methods to produce feature vectors – the subregions of the immediate keypoint surroundings are aggregated to magnitude and orientation values and sampled across a fixed-size grid into histogram bins, and the descriptor itself is defined as all values of those histograms. Feature vectors of histogram-based methods usually have a very high number of dimensions (e.g. 128 for SIFT), and the most common distance metric (euclidean multidimensional vector distance) can be computationally complex to calculate.

Subfeature-based descriptors consist of multiple subfeature detectors, each producing its single distinct output value based on sampling pattern, which is then usually normalized. The most typical sampling patterns produce pair-by-pair comparison between specific points of surrounding region. More complex subfeature detectors are also coupled with orientation compensation pattern to reduce the effect of affine transform on pair-by-pair comparison of sampling pattern. Each subfeature value is then stored in the resulting feature vector. Since each subfeature is distinct, they can only be compared by their corresponding values, so the most common distance metric is the sum of differences for the same features.

The most repetative and computationally expensive task when working with feature descriptors is usually the actual feature vector comparison, i.e. the calculation of a metric m (2). Most of the time the algorithmic complexity of feature descriptors is high enough as it is; moreover, common recognition tasks presume comparison of two sets of feature points by performing comparison for every possible pair. In this context, of particular interest are subfeature-based feature extractors for which each of the subfeatures can be normalized and then binarized with an appropriate threshold. That way, each subfeature output becomes a binary value, and feature vector space becomes an n-dimensional boolean $\mathbb{B}^n$. The most common metric for these descriptors becomes the Hamming distance – a count of non-matching subfeatures. Descriptors that produce binary feature vectors are called binary descriptors, and are most widely used in real-time processing tasks, because calculating Hamming distance between to feature vectors is extremely fast.

The classical approach to object detection was proposed by P. Viola in 2001 paper "Rapid Object Detection using a Boosted Cascade of Simple Features". It is a machine learning approach that uses Haar-like feature extractors and appropriate feature descriptors to match around 6000 features in a 24x24 pixel windows, combined in a cascade of classifiers. This approach is still widely used for face recognition problems and is computationally effective enough to be used in real-time. Various modifications of this approach were proposed in the following years, making it more suitable for specific pattern recognition tasks. [9]

Another, more recent approach to object detection uses region-based active contour models for image segmentation. The boundaries are computed and adjusted iteratively using edge-based and region-based terms of particular feature extractor and descriptor combination, defined as an optimization problem of a local cost function of boundary inconsistency accros feature space.

A more modern approach to object detection and localization is the usage of convolutional neural networks. Convolutional neural networks differ from more traditional object detection methods because the network itself is trained to produce and accentuate features it requires for a particular object classification. By examining the output of the deeper layer neurons it's also possible to evaluate the boundaries of a given object. While training a convolutional neural network usually takes a significant amount of time, fully trained networks are actually performant enough to perform object classification and detection in real-time.

III. Real-time polyp detection

There are multiple specifics that must be taken into account when working with videoendoscopic images and image sequences.

First, most of the videoendoscopes come equipped with a wide-angle lens camera. The non-linear transformation caused by a wide-angle lens produces the effect most commonly known as fisheye distortion. Fisheye lenses capture the light of not only immediate forward frame area, but also of objects reflected from around its vicinity. Because of the nature of such a distortion, objects closer to the edges of the image appear much larger than they actually are, and objects closer to the center appear smaller. Also, straight lines moving away from the screen towards the center gain a curvature that may be recalculated given known lens parameters. [8]

For most practical applications of digital image processing distortion correction is a necessary pre-processing step. It allows to eliminate the inconsistencies of scale and curvature which are determinal when using keypoint-based feature extractors and object detection algorithms.

Of particular note is the narrow-band imaging (NBI) technique implemented in most of the modern videoendoscopic systems. The main idea of NBI is applying a set of color filters on wavelengths that correspond to typical color chroma of blood, blood vessels, background tissue and other objects most likely present on endoscopic image, and then normalizing the remaining wavelength range non-linearly. The result is usually much more visually contrast image that makes expert evaluation of the region much easies.

All of the above produce a set of requirements that an automated real-time polyp detection system must meet:

- Detection must be invariant against image rotation. This requirement is based on the fact that a certain region may be observed under different angles, and its detection and classification must keep working in those cases.
- Detection must be distortion-invariant. A specific pattern of polyp should be detected regardless of the way it's distorted, i.e. on distorted image detector must produce accurate results on the edges and towards the center of the image.
- Detection should work regardless of relative contrast and absolute color values of the objects. Since regular image, post-processed image and NBI image produce different color representations of the same region, detector must use spatial and differential features in order to classify objects.
- Detection visualization should not interfere with the procedure. Detected object boundaries should not obstruct important parts of the image, while at the same time location of detected object should be clear and distinct enough to be able to spot it easily.
- Detection, as mentioned several times in this paper, should be performed in real time. This requirement means that detection should work on a separate framerate than the main camera in order to prevent input lag, and it should be fast enough to have a detection performance of 15-20 frames per second to be able to keep up with the main screen.
- Detection should be precise. False-negative detections (i.e. failing to detect an object) mean that a potentially dangerous polyp can be overlooked, while false-positive detections (i.e. detecting an object where there is none) may unnecesarily divert physician's attention, thus complicating and prolonging the procedure.

Image rotation invariance can be achieved by using scale-invariant feature extractors. Most of the modern feature extractors already provide scale-invariance i.e. descriptors of features remain similar when the image is subjected to simple affine transformations. For the actual implementation of the system, it was decided to use ORB (Oriented FAST and Rotated BRIEF) feature detector and extractor. It provides numerous significant advantages over more traditional Scale-Invariant Feature Transform based approaches in a significant increase of efficiency; moreover, descriptors produced by ORB are binary, which means that keypoint matching is extremely fast. [4]

Distortion invariance is implemented using a simple distortion correction algorithm that uses standard fisheye distortion model with adjusted parameters acquired using endoscopic camera calibration. While the actual transformation of the entire image can be computationally expensive, actual per-pixel transformation can be postponed until its evaluation is required. [8], [10] Moreover, keypoint detection for some of the video frames can be skipped entirely and instead localized to areas around keypoints detected on the previous frame, thus the evaluation of spatial brightness and distortion correction transformation will only be required in these areas. The system re-calculates a new set of keypoints on each 20th frame.

Color invariance is achieved by simplifying input frame colorspace. There are 2 modes supported – using normalized blue color component of RGB image representation and using normalized brightness component of HSB image representation. Normalization makes sure that relative contrast doesn't affect the results as much.

Unobtrusiveness of detection result visualization is hard to achieve. For initial implementation, it was decided to use a rectangle selection to outline the bounding frame of detected object with the ability to hide detection markers at will if they happen to interfere with the observation.Example frame with outlined detected object is presented on Figure 1.

Figure 1. Example of a figure caption.

Finally, in order to optimize the precision of detection it is possible to adjust the sensitivity of the resulting detection. In order to match the sought objects, a comprehensive feature vector sets were built by using a mapped image sets built from several colonoscopy screening videos containing known types of polyps that should be detected by the system, evaluated by an expert. While the precise sensitivity can only be determined on a case-by-case basis, a reasonable default value is provided. Also, the system optionally provides suggestions to enable or disable certain endoscopy system built-in image enhancements when output confidence score is not high enough to qualify as exact match.

IV. Conclusion

Implemented software complex can be used as a decision support system for endoscopic examination. The system is able to detect suspicious objects in real-time during the screening, make them visually clear to the physician performing the procedure in order to optimize the procedure time, emphasise the attention on certain automatically detected areas during the screening procedure which, in turn, should serve to increase the effective ADR of the physician.

Future system enhancements include the implementation of multi-class detection in order to not only spot suspicious objects on the image, but to also provide insights about the exact type of the object. This can further be expanded to propose suggestions about the optimal diagnosis and even further treatment. Also, these tasks can be also performed after the initial examination by analyzing the resulting video, thus lifting the real-time processing constraint. Since object detection relies heavily on calculating and evaluating keypoints using feature detectors and descriptors, it's also possible to include different types of processing that makes use of feature extraction. One of those is the problem of 3-dimensional spatial reconstruction of the scenes present on the video. This can prove benificial in a more detailed analysis of areas of interest.

References

[1] D. A. Corley et al, *Adenoma Detection Rate and Risk of Colorectal Cancer and Death*. The New England Journal of Medicine; vol. 370, pp. 1298-1306; 2014

[2] L. G. Shapiro, G .C. Stockman, *Computer Vision*. New Jersey, Prentice-Hall, 608 p.; 2001

[3] C. Barata, J. S. Marques, J. Rozeira, *The role of keypoint sampling on the classification of melanomas in dermoscopy images using bag-of-features*. Iberian Conference on Pattern Recognition and Image Analysis, pp. 715-723. Springer, Berlin, Heidelberg, 2013.

[4] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, *ORB: an efficient alternative to SIFT or SURF*. 2011 IEEE International Conference on Computer Vision (ICCV), pp. 2564-2571, IEEE, 2011

[5] K. M. Yi, E. Trulls, V. Lepetit, P. Fua, *Lift: Learned invariant feature transform*. European Conference on Computer Vision, pp. 467-483. Springer International Publishing, 2016.

[6] W. Wu, A. Y. C. Chen, L. Zhao, J. J. Corso, *Brain Tumor detection and segmentation in a CRF framework with pixel-pairwise affinity and super pixel-level features*. International Journal of Computer Aided Radiology and Surgery; vol. 9, pp. 241-253; 2014

[7] M. Gabryel, M. Korytkowski, R. Scherer, L. Rutkowski, *Object detection by simple fuzzy classifiers generated by boosting*. International Conference on Artificial Intelligence and Soft Computing, pp. 540-547. Springer, Berlin, Heidelberg, 2013.

[8] A. Sotiras, C. Davatzikos, N. Paragios, *Deformable Medical Image Registration: A Survey*. IEEE Transactions on Medical Imaging, vol. 32, issue 7, pp.1153-1190, IEEE, 2013

[9] P. Viola, M. Jones, *Rapid object detection using a boosted cascade of simple features*. IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1., pp. 511-519, IEEE, 2001

[10] A. Foncubierta-Rodriguez, H. Müller, A. Depeursinge, *Region based volumetric medical image retrieval*. SPIE Proceedings, vol. 8674, p. 10. 2013.

АВТОМАТИЗИРОВАННОЕ ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ В РЕАЛЬНОМ ВРЕМЕНИ ПРИ ПРОВЕДЕНИИ ЭНДОСКОПИЧЕСКОГО СКРИНИНГА

Головатая Е.А., Белорусский Государственный университет

Курочкин А.В., Белорусский Государственный университет

В данной работе рассматриваются методы компьютерного зрения для предоставления визуальной поддержки и выделения объектов при проведении процедуры эндоскопического скрининга. Разработанная авторами система предоставляет визуальные подсказки специалисту, осуществляющему скрининг, и выделяет образования и другие области интереса в реальном времени, с целью повышения точности диагностики различных видов рака.

Metadata for object-oriented television

Olena Osharovska
Dept. Radio and Television Broadcasting
O.S. Popov Odessa National Academy of Telecommunications
osharovskaya@gmail.com

***Abstract*—In broadcast multimedia metadata covers both the description of services and multimedia control. Each media object has a set of descriptors associated with it. Object descriptor identifies all streams associated with the media object. TV mesh objects are combining several structures, in particular: polygonal mesh cage; a set of matching texture; a set of normal vectors to faces; the color histogram about the content. The model of the content-oriented metadata being generated must be smaller than the original data of mesh object. In the article, it is proposed to supplement Language's Web Ontology Dictionary with terms on multimedia object-oriented broadcasting.**

***Keywords*—TV, mesh-object, metadata, broadcasting service, semantic relation, ontology**

I. Introduction

For almost a century, the broadcasting television was based on the concept of simultaneous transmission of socially significant information to all viewers. With the introduction of digital broadcasting, there is a rapid transition to new types of consumption of personalized video content at a convenient time for the subscriber. In addition, television becomes interactive with the possibility of active participation of the viewer in the formation of programs. The active position of viewers demanded not only to change the system of delivery and storage of media information, but also a new approach to the process of producing video content. One of the innovative approaches to creating programs is object-oriented creation of scenes, considered in the Recommendations MPEG-4, MPEG-7 and MPEG-21 [1, 2, 3]. The main purpose of the MPEG-4 standard in working documents of the MPEG group is formulated as follows: it sets the principles of working with content (digital representation of media data) for three areas: actually, interactive multimedia (including products distributed on optical disks and via the Network), graphics applications (synthetic content) and digital television – DTV. MPEG-4 is not only a standard, in fact it sets the rules for the organization of the environment, and the environment is object-oriented. It deals not just with flows and arrays of media data, but also with media objects this is the key concept of the standard. Objects can be audio, video, audiovisual, graphic (flat and three-dimensional), text. They can be either "natural" (recorded, filmed, scanned, etc.), or synthetic (i.e., artificially generated). Examples of objects can serve as a fixed background, video characters separate from the background, text-synthesized speech, musical fragments, and a three-dimensional model that can be moved and rotated in the frame. Media objects can be streamed by any video-pass. Each media object has a set of descriptors associated with it, where all its properties are defined, the operations necessary to decode the associated streaming data, placement in the scene, as well as the behavior and permissible reactions to the user's actions. There are advanced tools for working with two-dimensional polygonal models, adapting them to existing video content for later animation. Using these tools allows you to perform many functions, for example, representation of the contours of objects using grid vertices (instead of bit ones masks), replacement in the scene of "live" video objects by synthetic ones, etc. Material of this work is a continuation of previous studies in this area [4].

II. Mesh object metadata

The use of metadata or data about data, provides answers to how to handgrip the diverse services and contents of the new digital TV platform efficiently and a consumer-friendly way. However, the use of metadata is broadcast multimedia should not be seen as just a tool to cope with the challenges inherent in a complex networked multimedia environment. Instead, it inputs up new possibilities for the development of new innovative services.

In broadcast multimedia metadata covers both the description of services and multimedia control [5]. Metadata integrates into the broadcasting value-chain with considerations for each step in the development of a digital TV broadcasting service. Any different metadata, standards related to the digital broadcasting life-cycle. The concept bears the idea of companying on insight into syntonic and semantically complex data by refining their essence into a set of simple descriptors. Metadata also helps to arrangement and accomplish information in varied setting. The unit of interchange is a structured digital object (digital item (DI) described by metadata and referring related multimedia content assets. Media objects may need a data stream that is converted into one or more elementary streams. This allows you to hierarchically process the encoded data, as well as the associated media information about the content (called "object content information").

Each stream is characterized by a set of descriptors for configuring information, for example, to determine the necessary resources of the recording device and the accuracy of the encoded time information. Moreover, the descriptors may contain hints regarding the quality of service (QoS) that is required for transmission (e.g., the maximum number of bits/s, BER, priority, etc.). Binary format for Scenes (BIFS) describes the space-time relationships of objects on the stage. Viewers can have the ability to interact with objects, for example, by moving them on the stage or by changing their position of the observation point in a 3D virtual environment. The scene description provides a wide range of nodes for compositional 2-D and 3-D operators and graphic primitives [6, 7].

TV mesh objects are combining several structures, in particular: polygonal mesh cage; a set of matching texture; a set of normals to landfills; the color information for each polygon.

Mesh data has spatial and temporal aspects, as well as arbitrarily high dimensionality, which aggravates the task of finding compact, accurate, and easy-to-compute data models. This approach leads to interesting challenges. The model (equivalently, the content-oriented metadata) being generated must be smaller than the original data by at least an order of magnitude. Second, the metadata representation must contain enough information to support a broad class of queries. Finally, the accuracy and speed of the queries must be within the tolerances required by users.

We approach this problem by building compact, approximate, multi-resolution models of the mesh data and then using the models to support high-fidelity ad hoc queries [8].

Statistical and mathematical data analysis and compression methods such as clustering [9, 10], spline-based fitting [11, 12] and wavelets [13, 14], are all suitable in the preprocessing phase.

Let us define mesh data as a discrete representation of continuous data, which can be defined as an ordered set of tuples as follows. (Mesh data represented in this manner is called a “point mesh” [8], which is just a collection of data points with no topological connection among them.)

$\{t, x_1, x_2, ..., x_n, v_1, v_2, ..., v_m\}$ where t denotes a temporal variable defining a time step, x_i, $i \in [1, ...n]$, denotes a spatial variable defining the geometrical coordinates in an n-dimensional space, and v_j, $j \in [1, ...m]$, denotes a field variable defined at each node (positioned at $(t, x_1, x_2, ..., x_n)$ >) or each zone in the mesh at time step t. A zone in a Regular mesh is an n-dimensional cubic bounded by the surrounding 2n mesh nodes, whereas an indeterminate number of nodes surrounds a zone in an irregular mesh. One organize the metadata in two levels. The lower level contains two elements. The first element is a multi-resolution model of mesh data, modeled per partition of mesh data, as well as a collection of summary information that is generated as the result of preprocessing. For example, the summary information includes, per variable, information on min, max, mean, median, standard deviation, and the first several moments for each variable. The actual models vary from cluster prototypes, to regression equations, to matrixes of equations, and so on. The second element of metadata is a set of indexes that define the structure that must be traversed to reach the first element data. More specifically, each node of an index contains the summary information and model information of the corresponding mesh partition. There are many challenges to overcome for this approach to work. Resolving them involves investigating and making complex trade-offs between preprocessing time, compression level, query speed, query accuracy, and the range of queries the metadata supports.

Scalable solutions: all aspects of the preprocessing phase must be scalable to terabyte mesh data, and the algorithms must be amenable to efficient parallelization on machines.

Complete models: little is known about what makes an effective model of mesh data hat is capable of supporting arbitrarily complicated ad hoc queries. There are many alternatives, which have varying degrees of accuracy and descriptive power. The models must be compact, yet contain enough information to answer a wide range of possible queries.

Compact, efficient, multi-resolution metadata: we allow the scientists to trade query response time for accuracy, allowing interactive ad hoc queries at one extreme, and slower, and more accurate responses at the other.

It can be assumed that the goal of preprocessing the mesh data is to generate a representation of model that is much smaller and yet can be used to support approximate ad hoc queries. Firstly, the entire mesh is decomposed into an appropriate partition, then, the mesh data within each partition is concisely approximated using an appropriate parametric model (e.g., splines, wavelets, or clusters). Algorithms to generate this metadata typically iterate through two phases: mesh partitioning and partition characterization. Iteration between these phases takes place to revise the actual partitioning based on a similarity (or error) metric that measures the difference between the characterization of the data in a partition and the actual mesh data itself.

III. THE QUERY BY IMAGE CONTENT

Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases. Content-based image retrieval is opposed to traditional concept-based approaches.

Researchers [15, 16] mentioned three levels of queries in CBIR. Level 1: Retrieval by primitive features such as color, texture, shape, or the spatial location of image elements. Level 2: Retrieval of objects of given type identified by derived features, with some degree of logical inference. Level 3: Retrieval by abstract attributes, involving a big amount of high-level reasoning about the aim of the objects or scenes depicted. A CBIR system should provide full support in bridging the semantic gap between numerical image features and the richness of human semantics [16] in order to support query by high-level concepts.

Current systems mostly perform retrieval at Level 2. There are three fundamental components in these systems: low-level image feature extraction; similarity measure; semantic gap reduction.

Low-level image feature extraction is the basis of CBIR systems. To perform CBIR, image features can be either extracted from the entire image or from regions (region-based image retrieval, RBIR).

To perform RBIR, the first step is to implement image segmentation. Then, low-level features such as color, texture, shape, or spatial location can be extracted from the segmented regions. Similarity between two images is defined based on region features.

IV. Ontology for multimedia applications

Humans tend to use high-level features (concepts) to interpret images and measure their similarity. In general, there is no direct link between the high-level concepts and the low-level features [16]. Though many complex algorithms have been designed to describe color, shape, and texture features, these algorithms cannot adequately model image semantics and have a lot of limitations while dealing with broad content image databases [8] have not been described in the literature so far. An essential part of this methodology is to develop in OWL (Web Ontology Language) an ontology that fully captures the semantic metadata model defined in the Semantic Part of the MPEG-7 MDS [3]. The need for an ontology that fully captures the MPEG-7 metadata model has been pointed out by several research groups [7, 16, 17, 19]. Some important work in this direction has been carried out in [18] for the case of Resource Description Framework (RDF) [20], but it has some limitations: Classes corresponding to the MPEG-7 complex types have been defined, but not all of the (simple and complex) attributes of the classes are represented. In addition, typed relationships among the metadata items are not represented, although MPEG-7 provides complete support for typed relationships. Furthermore, this work has been based on RDF [21, 22].

Our approach for the integration of OWL ontologies in the framework for the support of ontology-based semantic indexing and retrieval of audiovisual content, utilizes an ontology that captures the model provided by the Semantic Part of the MPEG-7 MDS for the representation of semantic metadata for audiovisual content description. This ontology captures the semantics of the first layer of the two-layered model for semantic metadata used in the DS-MIRF framework. The second layer of the model encapsulates domain-specific knowledge, which extends the audiovisual content description standards so that they integrate transparently domain-specific ontologies. The model for the semantic description of audiovisual content provided in the Semantic part of the MPEG-7 MDS is comprised of complex data types defined, using the XML Schema Language syntax [2], in a set of Description Schemes (DSs) rooted in the Semantic Base DS.

Semantic Base DS: Semantic Base Type is the base type extended by other description schemes according to the needs for the description of semantic entities of specific types. Semantic Base Type has a set of simple attributes (ID for instance identification, time Base, time Unit, media Time Base and media Time Unit for timing support) and the following complex attributes: Abstraction Level, which represents the abstraction existing in the current semantic entity. Label, corresponding to a term that describes in brief the semantic entity. Definition, which is a textual annotation that describes the semantic entity. Property, which is a term that associates a property with the semantic entity. Relation, which relates the semantic entity with other semantic entities. Media Occurrence, which relates the semantic entity to specific media items (e.g. video segments, images etc.).

Semantic Bag DS and Semantic DS: Description schemes used for the description of collections of semantic entities. Semantic Bag Type is an abstract type, defined in the Semantic Bag DS, which extends Semantic Base Type.

Object DS: The Object Type defined here extends Semantic Base Type and is used for the description of objects and object abstractions (e.g. a table).

Agent Object DS: The actors that appear in an audiovisual segment are related with the instances of the Agent Object Type that extends the Object Type. Actors in general are represented using the Agent Type, an abstract type extending Semantic Base Type defined in the Agent DS. Person Type, Organization Type and Person Group Type extend Agent Type, are defined respectively in the Person DS, the Organization DS and the Person Group DS and are used for the representation of persons (e.g. football players), organizations and groups of persons. Event DS: The Event Type defined here extends Semantic Base Type and is used for the description of events (e.g. a goal).

Semantic State DS: The Semantic State Type defined here extends Semantic Base Type and is used for the description of states described in an audiovisual segment and the parametric description of its features (e.g. the score in a soccer game before and after a goal).

Semantic Place DS: The Semantic Place Type defined here extends Semantic Base Type and is used for the description of places (e.g. Athens). Semantic Time DS: The Semantic Time Type defined here extends Semantic Base Type and is used for the description of semantic time (e.g. Christmas). Concept DS: The Concept Type defined here extends Semantic Base Type and is used for the description of concepts present in an audiovisual segment (e.g. cooperation).

Semantic Relation DS: The Semantic Relation Type defined here extends Semantic Base Type and is used for the description of relationships among semantic entities. The relationships may be typed, as described in the Semantic Relation DS. In addition to the attributes inherited from Semantic Base Type, Semantic Relation Type has the following attributes: Source, which is the ID of the semantic entity of the relationship. Target, which is the ID of the semantic entity that is the target of the relationship. Argument, which may be used as an alternate to source and target definition. Strength, which denotes the strength of the relationship. Name, which denotes the name of the relationship. Arity, which denotes the arity of the relationship. Properties, where the properties of the relationship are denoted. MPEG-7 Simple Datatype Representation: The simple datatypes needed are integrated in the core ontology, as OWL permits the integration in OWL ontologies of simple datatypes defined in the XML Schema Language [2]. MPEG-7 Complex Type Representation: MPEG-7 complex types correspond to OWL classes, which define groups of individuals that belong together because they share some properties. Thus, for every complex type defined in the MPEG-

7 MDS a respective OWL class is defined. Simple Attribute Representation: The simple attributes of the complex type of the MPEG-7 MDS are represented as OWL datatype properties, which relate class instances to datatype instances. Complex Attribute Representation: For the representation of the complex attributes of the complex type, OWL object properties are used, which relate class instances. An OWL class for the representation of the complex attribute instances is defined, if it does not already exist.

V. Conclusions

Metadata in the form of text and visual information constitute the most important part of digital information flows between different TV systems. The advent of information systems tools into television has led to the need for media data modeling. Within the framework of data modeling, the concepts of entities, their attributes and relationships between modeling entities are used. In TV production systems, incoming metadata allows you to find and select fragments of source materials to create new programs for use. The programs completed in production should be appropriately described in the form of an organized set of metadata. Metadata on the created programs includes a part of the input metadata and some new metadata reflecting the essence of the created programs. Thus, it can be stated that in the production systems of TV programs there are operations for the formation and conversion of metadata. Metadata in the form of text and visual information constitute the most important part of digital information flows between different TV systems. The advent of information systems tools into television has led to the need for media data modeling. Within the framework of data modeling, the concepts of entities, their attributes and relationships between modeling entities are used.

References

[1] ITU-T Recommendation H.264 *ITU. 2003-05-30. Retrieved 2013-04-18.*

[2] ISO/IEC 15938-5 *FDIS Information Technology - Multimedia Content Description Interface - Part 5: Multimedia Description Schemes* . [*July 2001, Sydney*]

[3] Jane Hunter *Reconciling MPEG-7 and MPEG-21 Semantics through a Common Event-Aware Metadata Model* . [*https://arxiv.org/html/cs.cs/0210021*]

[4] Gofaizen Oleg, Osharovska Olena, Patlayenko, Volodymyr Pyliavskyi *Associative Relations in the Assessment of TV Images* . [*Open Semantic Technologies for Intelligent Systems*]. 16-18 february 2017, report 52, 4 p.

[5] Cox M., Mulder E. and Tadic L *Descriptive metadata television: an end to end introduction* . [*Focal Pr*]. 2006 - 157 P.

[6] ISO/IEC TR 21000-1:2001(E) *Part 1: Vision, Technologies and Strategy* . [*MPEG, Document: ISO/IEC JTC1/SC29/WG11 N3939*]. <http://www.cselt.it/mpeg/public/mpeg-21_pdtr.zip>

[7] ITU-R Recommendation BT-1865 *Metadata for monitoring errors in standard definition television signals and high definition television in the broadcasting chain* . []. 2010, 46 P.

[8] Lee S., Snapp R., Musick R., Critchlow T. *Ad hoc Query Support for Very Large Simulation Mesh Data: the Metadata Approach* . [ ]. 14 P.

[9] Hartigan John A. *Clustering Algorithms* . [*Johns Vviley & Sons*]. New York 1975

[10] Zupan J. *Clustering of Large Data Sets* . [*Johns Vviley & Sons*]. 1982

[11] C. de Boor *A Practical Guide to Splines* . [*Springer Verlag*]. 1978

[12] Dierckx P. *Curve and Surface Fitting With Splines Monographs on Numerical Analysis* . [*Oxford University Press*]. 1993.

[13] Antoniadis A. Oppenheim G. *Wavelets and Statistics* . [*Springer*]. Verlag 1998

[14] Härdle W., et all *Wavelets, Approximation and Statistical Applications* . [*Springer Verlag*]. 1998.

[15] Rui Yong, Huang Thomas S., Chang Shih-Fu *Image Retrieval: Current Techniques, Promising Directions, and Open Issues* . [*Journal of Visual Communication and Image Representation*]. 10: P. 39–62. doi:10.1006/jvci.1999.0413

[16] Nuno Vasconcelos *From Pixels to Semantic Spaces: Advances in Content-Based Image Retrieval* . [*Computer*]. 2007, pp.20-26. http://www.svcl.ucsd.edu/publications/journal/2007/computer /computer07.pdf

[17] EBU *Operating Eurovision and Euroradio* . [*TECH 3351 EBU Class Conceptual Data Model*]. Geneva 2016

[18] EBU *Operating Eurovision and Euroradio* . [*TECH 3293 EBU Core Metadata Set*]. Geneva 2016

[19] Seung-Jun Yang, Jung Won Kang, Dong- San Jun, Min Je Kim, Han-Kyu Lee *TV-Anytime Metadata Authoring Tool for Personalized Broadcasting Services* . [*Consumer Electronics, ISCE 2007. IEEE International Symposium*]. 2007. 6 p. DOI: 10.1109/ISCE.2007.4382191

[20] B. Makni et al *Semantic integration of TV data and services. A survey on challenges and approaches* . [*Web Intelligence and Agent Systems An International Journal 1*]. 2010 pp. 1-24.

[21] Tsinaraki C., Polydoros P., Christodoulakis S. *Integration of OWL Ontologies in MPEG-7 and TV-Anytime compliant Semantic Indexing* . [*International Conference on Advanced Information Systems Engineering CAiSE: Advanced Information Systems Engineering*]. 2004 pp 398-413

[22] Migual-Angel Sicilia *Handbook of Metadata, Semantics and Ontologies* . [*World Scientific Publish*]. 2014. 523 p

МЕТАДАННЫЕ ДЛЯ ОБЪЕКТНО-ОРИЕНТИРОВАННОГО ТЕЛЕВИДЕНИЯ

Ошаровская Е.В.

Одесская национальная академия связи им. А.С.Попова

В вещательных мультимедийных системах метаданные рассматриваются как описание дополнительных сервисов, так и управление самим контентом мультимедиа. Каждый медиа-объект имеет набор связанных с ним дескрипторов. Телевизионная сцена компонуется из нескольких объектов. Видео объекты целесообразно представлять триангуляционными сетками.

Модель что эквивалентно описанию метаданными, ориенти-рованными на контент, должна создавать меньший битовый поток, чем исходные данные самого сеточного объекта, по крайней мере, на порядок. Представление метаданных должно содержать достаточно информации для поддержки широкого класса запросов. Наконец, точность и скорость запросов должны быть в пределах допусков, требуемых пользователями. Метаданные органи-зованы в два уровня. Нижний уровень - это модель сетки при разном разрешении. Верхний уровень несет информацию о семантике сеточного объекта. Единица обмена представляет собой структурированный цифровой объект (цифровой элемент (DI)).

Метаданные в виде текстовой и визуальной информации составляют важнейшую часть цифровых информационных потоков между различными телевизионными системами. В этой статье предлагается использовать словарь веб-онтологии с терминами мультимедийного объектно-ориентированного вещания.

Evolution of Artificial Neural Networks

Alexey Averkin
Dorodnicyn Computing Centre
Federal Research Center "Computer Science and Control"
of Russian Academy of Sciences
Moscow, Russia
averkin2003@inbox.ru

Sergey Yarushev
Dubna State University
Dubna, Russia
Sergey.Yarushev@icloud.com

***Abstract*—The history of artificial neural networks originates from the distant 1943, when Warren McCulloch and Walter Pitts formalized the notion of a neural network in a fundamental article on the logical calculation of ideas and nervous activity. This was the starting point in the history of artificial neural networks. Later, a huge number of different architectures of neural networks, learning methods and optimization algorithms were proposed. There was a time when neural networks were left in oblivion, but after the 1980s they were talked about again when John Hopfield introduced his famous full-mesh network. Today, artificial neural networks are indispensable tools in a huge number of tasks, such as tasks of forecasting and analyzing time series, the task of recognizing images and even emotions, classification tasks, natural language recognition, neural networks are used in industry, defense and medicine; many enumerate areas of their application. All this tells us that artificial neural networks have become part of modern life.**

***Keywords*—artificial neural networks, forecasting, modular neural networks**

I. Introduction

What was once just a figment of the imagination of some our most famous science fiction writers, artificial intelligence (AI) is taking root in our everyday lives. We're still a few years away from having robots at our beck and call, but AI has already had a profound impact in more subtle ways. Weather forecasts, email spam filtering, Google's search predictions, and voice recognition, such Apple's Siri, are all examples. What these technologies have in common are machine-learning algorithms that enable them to react and respond in real time. There will be growing pains as AI technology evolves, but the positive effect it will have on society in terms of efficiency is immeasurable. In this paper we will talking about Artificial Neural Networks - great instrument of the Artificial Intelligence.

II. Early researches in Artificial Neural Networks

In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts wrote a paper on how neurons might work [8]. In order to describe how neurons in the brain might work, they modeled a simple neural network using electrical circuits. In 1949, Donald Hebb wrote "The Organization of Behavior", a work which pointed out the fact that neural pathways are strengthened each time they are used, a concept fundamentally essential to the ways in which humans learn. If two nerves fire at the same time, he argued, the connection between them is enhanced. As computers became more advanced in the 1950's, it was finally possible to simulate a hypothetical neural network. The first step towards this was made by Nathanial Rochester from the IBM research laboratories. Unfortunately for him, the first attempt to do so failed.

A. First Model of Artificial Neural Network

American neurophysiologist Frank Rosenblatt. He proposed a diagram of a device that simulates the process of human perception, and called it a "perceptron". The perceptron transmitted signals from photocells, which are a sensory field, to blocks of electromechanical memory cells. These cells were connected randomly in accordance with the principles of connectivity. In 1957, the Cornell Laboratory of Aeronautics successfully completed the modeling of perceptron work on the IBM 704 computer, and two years later, on June 23, 1960 at Cornell University, the first neurocomputer Mark-1 was demonstrated, which was able to recognize certain letters of the English alphabet [4].

Frank Rosenblatt with his creation - "Mark-1." To "teach" the perceptron to classify images, a special iterative method of learning trial and error was developed, reminiscent of the process of human training - the method of error correction [3]. In addition, when recognizing a particular letter, the perceptron could distinguish the characteristic features of letters that are statistically more frequent than minor differences in individual cases. Thus, the perceptron was able to generalize letters written in different ways (handwriting) into one generalized image. However, the Perceptron's possibilities were limited: the machine could not reliably recognize partially closed letters, as well as letters of a different size, located with a shift or rotation, rather than those used at the stage of its training. The report on the first results appeared in 1958 - then Rosenblatt published an article "Perceptron: A Probable Model of Storage and Organization of Information in the Brain" [11]. But he describes his theories and assumptions about perception processes and perceptrons in more detail in 196 in his book "Principles of Neurodynamics: Perceptrons and Theory of Brain Mechanisms" [12]. In the book, he examines not only ready-made perceptron models with one hidden layer, but also multilayer perceptrons with cross-sections (the third chapter) and inverse (fourth chapter) connections. The book also introduces a number of important ideas and theorems, for example, the theorem of convergence of a perceptron is proved.

B. Next Researches

In 1959, Bernard Widrow and Marcian Hoff of Stanford developed models called "ADALINE" and "MADALINE." In a typical display of Stanford's love for acronymns, the names come from their use of Multiple ADAptive LINear Elements. ADALINE was developed to recognize binary patterns so that if it was reading streaming bits from a phone line, it could predict the next bit. MADALINE was the first neural network applied to a real world problem, using an adaptive filter that eliminates echoes on phone lines. While the system is as ancient as air traffic control systems, like air traffic control systems, it is still in commercial use. In 1962, Widrow & Hoff developed a learning procedure that examines the value before the weight adjusts it (i.e. 0 or 1) according to the rule: Weight Change = (Pre-Weight line value) * (Error / (Number of Inputs)). It is based on the idea that while one active perceptron may have a big error, one can adjust the weight values to distribute it across the network, or at least to adjacent perceptrons. Applying this rule still results in an error if the line before the weight is 0, although this will eventually correct itself. If the error is conserved so that all of it is distributed to all of the weights than the error is eliminated. Despite the later success of the neural network, traditional von Neumann architecture took over the computing scene, and neural research was left behind. Ironically, John von Neumann himself suggested the imitation of neural functions by using telegraph relays or vacuum tubes. In the same time period, a paper was written that suggested there could not be an extension from the single layered neural network to a multiple layered neural network. In addition, many people in the field were using a learning function that was fundamentally flawed because it was not differentiable across the entire line. As a result, research and funding went drastically down. This was coupled with the fact that the early successes of some neural networks led to an exaggeration of the potential of neural networks, especially considering the practical technology at the time. Promises went unfulfilled, and at times greater philosophical questions led to fear. Writers pondered the effect that the so-called "thinking machines" would have on humans, ideas which are still around today. The idea of a computer which programs itself is very appealing. If Microsoft's Windows 2000 could reprogram itself, it might be able to repair the thousands of bugs that the programming staff made. Such ideas were appealing but very difficult to implement. In addition, von Neumann architecture was gaining in popularity. There were a few advances in the field, but for the most part research was few and far between. In 1972, Kohonen and Anderson developed a similar network independently of one another, which we will discuss more about later. They both used matrix mathematics to describe their ideas but did not realize that what they were doing was creating an array of analog ADALINE circuits. The neurons are supposed to activate a set of outputs instead of just one. The first multilayered network was developed in 1975, an unsupervised network.

III. From 1980 to present

In 1982, interest in the field was renewed. John Hopfield of Caltech presented a paper to the National Academy of Sciences. His approach was to create more useful machines by using bidirectional lines. Previously, the connections between neurons was only one way. That same year, Reilly and Cooper used a "Hybrid network" with multiple layers, each layer using a different problem-solving strategy. Also in 1982, there was a joint US-Japan conference on Cooperative/Competitive Neural Networks. Japan announced a new Fifth Generation effort on neural networks, and US papers generated worry that the US could be left behind in the field. (Fifth generation computing involves artificial intelligence. First generation used switches and wires, second generation used the transister, third state used solid-state technology like integrated circuits and higher level programming languages, and the fourth generation is code generators.) As a result, there was more funding and thus more research in the field. In 1986, with multiple layered neural networks in the news, the problem was how to extend the Widrow-Hoff rule to multiple layers. Three independent groups of researchers, one of which included David Rumelhart, a former member of Stanford's psychology department, came up with similar ideas which are now called back propagation networks because it distributes pattern recognition errors throughout the network. Hybrid networks used just two layers, these back-propagation networks use many. The result is that back-propagation networks are "slow learners," needing possibly thousands of iterations to learn. Now, neural networks are used in several applications, some of which we will describe later in our presentation. The fundamental idea behind the nature of neural networks is that if it works in nature, it must be able to work in computers. The future of neural networks, though, lies in the development of hardware. Much like the advanced chess-playing machines like Deep Blue, fast, efficient neural networks depend on hardware being specified for its eventual use. Research that concentrates on developing neural networks is relatively slow. Due to the limitations of processors, neural networks take weeks to learn. Some companies are trying to create what is called a "silicon compiler" to generate a specific type of integrated circuit that is optimized for the application of neural networks. Digital, analog, and optical chips are the different types of chips being developed. One might immediately discount analog signals as a thing of the past. However neurons in the brain actually work more like analog signals than digital signals. While digital signals have two distinct states (1 or 0, on or off), analog signals vary between minimum and maximum values. It may be awhile, though, before optical chips can be used in commercial applications.

IV. Examples of Nowdays Artificial Neural Networks

A. Modular Neural Networks

The core of the modular neural networks is based on the principle of decomposition of complex tasks into simpler ones. Separate modules make simple tasks. More simple subtasks are then carried through a series of special models. Each local model performs its own version of the problem according to its characteristics. The decision of the integrated object is achieved by combining the individual results of specialized local computer systems in a dependent task. The expansion of the overall problem into simpler subtasks can be either soft or hard unit subdivision. In the first case, two or more subtasks of local computer systems can simultaneously assigned while in the latter case, only one local computing model is responsible for each of the tasks crushed. Each modular system has a number of special modules that are working in small main tasks. Each module has the following characteristics:

1) The domain modules are specific and have specialized computational architectures to rec-ognize and respond to certain subsets of the overall task;
2) Each module is typically independent of other modules in its functioning and does not in-fluence or become influenced by other modules;
3) The modules generally have a simpler architecture as compared to the system as a whole. Thus, a module can respond to given input faster than a complex monolithic system;
4) The responses of the individual modules are simple and have to combine by some inte-grating mechanism in order to generate the complex overall system response.

The best example of modular system is human visual system. In this system, different mod-ules are responsible for special tasks, like a motion detection, color recognition and shape. The central nervous system, upon receiving responses of the individual modules, develops a complete realization of the object which was processed by the visual system.

B. ANN in Time Series Forecasting

The recent surge in research of artificial neural networks (ANN) showed that neural networks have a strong capability in predicting and classification problems. ANN successfully used for various tasks in many areas of business, industry and science [14]. Such high interest in neural networks caused by the rapid growth in the number of articles published in scientific journals in various disciplines. It suffices to consider several large databases to understand the huge number of articles published during the year on the theme of the study of neural networks, it is thousands of articles. A neural network is able to work parallel with input variables and consequently handle large sets of data quickly. The main advantage of neural networks is the ability to find patterns [2]. ANN is a promising alternative in the toolbox professionals involved in forecasting. In fact, the nonlinear structure of the neural networks is partially useful to identify complex relationships in most real world problems. Neural networks are perhaps the universal method of forecasting in connection with those that they cannot only find the non linear structures in problems, they can also simulate the processes of linear processes. For example, the possibility of neural networks in the modeling of linear time series line were studied and confirmed by a number of researchers [5]. One of the main applications of ANN is forecasting. In recent years, it was seen increasing interest in forecasting using neural networks. Forecasting has a long history, and its importance reflected in the application in a variety of disciplines from business to engineering. The ability to accurately predict the future is fundamental to many decision making processes in planning, developing strategies, building policy, as well as in the management of supply and stock prices. As such, forecasting is an area in which a lot of effort has invested in the past. In addition, it remains an important and active area of human activity in the present and will continue to evolve in the future. Review of the research needs in the prediction presented to the Armstrong [1]. For several decades in forecasting dominated by linear methods. Linear methods are simple in design and use, and they are easy to understand and interpret. However, linear models have significant limitations, owing to which they cannot discern any nonlinear relationships in data. Approximation of linear models to complex are not linear relationships do not always give a positive result. Earlier in 1980, there have been large scale competition for forecasting, in which most widely used linear methods were tested on more than 1,000 real time series [7]. Mixed results showed that none of the linear model did not show the best results worldwide, which can be interpreted as a failure of the linear models in the field of accounting with a certain degree of non linearity, which is common for the real world. Predicting financial markets is one of the most important trends in research due to their commercial appeal [6]. Unfortunately, the financial markets are dynamic, nonlinear, complex, nonparametric and chaotic by nature [13]. Time series of multistationary, noisy, casual, and have frequent structural breaks. In addition, the financial markets also affects a large number of macroeconomic factors [9], such as political developments, global economic developments, bank rating, the policy of large corporations, exchange rates, investment expectations, and events in other stock markets, and even psychological factors. Artificial neural networks are one of the technologies that have received significant progress in the study of stock markets. In general, the value of the shares is a random sequence with some noise, in turn, artificial neural networks are powerful parallel processors nonlinear systems depending on their internal relations. Development of techniques and methods that can approximate any nonlinear continuous function without a priori notions about the nature of the process itself seen in the work of P. Pino [10]. It is obvious that a number of factors demonstrate sufficient efficacy in the forecast prices, and most importantly a weak point in this is that they all contain a few limitations in forecasting stock prices and use linear methods, the relative of this fact, although previous investigation revealed the problem to some extent, none of them provides a comprehensive model for the valuation of shares. If we evaluate the cost and provide a model in order to remove the uncertainty, it is largely can help to increase the investment attractiveness of the stock exchanges. Conduct research to get the best method of forecasting financial time series is currently the most popular and promising task.

V. Conclusion

In this paper, we talked about evolution of artificial neural networks evolved from the first models of ANN, which presented Warren McCulloch and Walter Pitts. The history of ANN is not so long but evolution of architectures and methods is is impressive. Now we could use ANN for prediction and data analisys, for computer vision and emotions recognition. Neural networks are used in safety systems. We could see very wide fileds in our paper where ANN could be applied. In the development of new smartphones used artifitial intelligence, in automotive industry ANN is used for safety and assistant systems. All it demonstrate - Artificial Intelligence is a part of our nowdays life!

Acknowledgment

This work was supported by the Russian Foundation for Basic Research 17-07-01558 and 17-37-50013

References

[1] J Scott Armstrong. Research needs in forecasting. *International Journal of Forecasting*, 4(3):pp.449–465, 1988.

[2] Pei-Chann Chang, Yen-Wen Wang, and Chen-Hao Liu. The development of a weighted evolving fuzzy neural network for pcb sales forecasting. *Expert Systems with Applications*, 32(1):pp.86–96, 2007.

[3] Michele D Estebon. Perceptrons: An associative learning network. *Virginia Tech.—1997*, 1997.

[4] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.

[5] H Brian Hwarng. Insights into neural-network forecasting of time series corresponding to arma (p, q) structures. *Omega*, 29(3):pp.273–289, 2001.

[6] Ritanjah Majhi, Ganapati Panda, Gadadhar Sahoo, Pradipta K Dash, and Debi Prasad Das. Stock market prediction of s&p 500 and djia using bacterial foraging optimization technique. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pages pp.2569–2575. IEEE, 2007.

[7] Spyros Makridakis, A Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, 1(2):pp.111–153, 1982.

[8] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):pp.115–133, 1943.

[9] Kai Miao, Fang Chen, and ZG Zhao. Stock price forecast based on bacterial colony rbf neural network [j]. *Journal of Qingdao University (Natural Science Edition)*, 20(2):pp.50–54, 2007.

[10] Raúl Pino, José Parreno, Alberto Gomez, and Paolo Priore. Forecasting next-day price of electricity in the spanish energy market using artificial neural networks. *Engineering Applications of Artificial Intelligence*, 21(1):pp.53–62, 2008.

[11] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):p.386, 1958.

[12] Frank Rosenblatt. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. Technical report, CORNELL AERONAUTICAL LAB INC BUFFALO NY, 1961.

[13] Tuan Zea Tan, Chai Quek, and Geok See Ng. Brain-inspired genetic complementary learning for stock market prediction. In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, volume 3, pages pp.2653–2660. IEEE, 2005.

[14] Bernard Widrow, David E Rumelhart, and Michael A Lehr. Neural networks: applications in industry, business and science. *Communications of the ACM*, 37(3):pp.93–106, 1994.

ЭВОЛЮЦИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Аверкин А.Н.
Вычислительный центр им. А.А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук
г. Москва, Россия
Ярушев С.А.
Государственный Университет "Дубна"
г. Дубна, Россия

История искусственных нейронных сетей берет свое начало с далекого 1943 года, когда Уоррен Мак-Каллок и Уолтер Питтс ормализовали понятие нейронной сети в фундаментальной статье о логическом исчислении идей и нервной активности. Это и стало отсчетной точкой в истории искусственных нейронных сетей. В последствии было предложено огромное количество разнообразных архитектур нейронных сетей, методов обучения и алгоритмов оптимизации. Было время, когда нейронные сети подвергались забвению, но после 1980х годов о них заговорили снова, когда Джон Хопфилд представил свою знаменитую полносвязную сеть сеть. Сегодня искусственные нейронные сети являются незаменимым инструментов в огромном количестве задач, таких, как задачи прогнозирования и анализа временных рядов, задачи распознавания образов, изображений и даже эмоций, задачи классификации, распознавания естесственного языка, нейронные сети применяются в промышленности, обороне и медицине, можно очень много перечислять области их применения. Все это говорит нам о том, что искусственные нейронные сети стали частью современной жизни.

Neural Network Structures: Current and Future States

Maksim V. Davydov
and Anatoly N. Osipov
Belarusian State University
of Informatics and Radioelectronics
Minsk, Belarus
{davydov-mv, osipov}@bsuir.by

Sergei Y. Kilin
B.I.Stepanov Institute of Physics,
National Academy of Sciences of Belarus
Minsk, Belarus
sergei_kilin@yahoo.com

Vladimir A. Kulchitsky
Institute of Physiology,
National Academy of Sciences of Belarus,
Minsk, Belarus
vladi@fizio.bas-net.by

***Abstract*—Two prevailing points of view on the way of creating artificial intelligence have been presented: based on the consistent patterns the evolution of neurons of the natural nervous system went through and based on radically new patterns, different from the consistent patterns seen in the natural evolution, shall be used to develop neural network structures and programs. The main structural elements of the future artificial intelligence based on the laws of the development of human neural networks have been presented: the main types of neurotransmitters and their role in the modulation of informative signals; a principle of diffuse transmission of information through the intercellular matrix; a concept of improving efficiency of functioning neural network structures based on diffuse transmission of signals in the form of hormones, growth factors, cytokines. It was pointed out that the researchers who are aiming at the development of artificial intelligence must remember about potential negative consequences that should be taken into account and avoided at the stages of problem semantic comprehension.**

I. Introduction

Any discussions dedicated to artificial intelligence involve two prevailing points of view. As the case in point concerns the intelligence inherent in humans only, one of the points of view assumes that the approach for creating artificial intelligence shall be based on the consistent patterns the evolution of neurons of natural nervous system went through.

The other point of view is that some radically new patterns, different from the consistent patterns seen in the natural evolution, shall be used to develop neural network structures and programs. Since the organizers and the participants of the scientific conference Open Semantic Technologies for Intelligent Systems (OSTIS-2018) are basically the adherents of the latter point of view, this paper focuses on the former one.

One remark need to be made in the very beginning: the first approach assuming a certain intervention of humans in natural laws of the development of nervous system brings a potential danger. We mean here a phenomenon described in literary form by Percy Bysshe Shelly's wife Mary Shelly who wrote a novel about Frankenstein. Therefore, the researchers who are aiming at the development of artificial intelligence must remember about potential negative consequences that should be taken into account and avoided at the stages of problem semantic comprehension.

Now let us turn directly to the main structural elements of the future artificial intelligence based on the laws of the development of human neural networks.

In the neural networks of the human brain, the signals are predominantly transferred by means of electricity using the branching extensions of nerve cells. However, direct signal recognition is accomplished by means of chemical intermediaries, or neurotransmitters, allocated discretely in the form of quanta of matter, that ensure the modulation of informative signals (Figure 1). Considering a huge variety of intermediaries (neurotransmitters), ranging up to tens and hundreds of thousands, we can imagine the great variety of the modulating abilities of neural networks of the human brain. By the way, the emotional, cognitive diversity of human behavior illustrates the diversity of capabilities of the brain neural network structures demonstrated above [1-3].

The next important concept of the nervous system organization that should be considered when creating artificial intelligence, is the colossal ability of just a single neuron to process information, to store, sum and integrate various signals actually instantaneously (within fractions of milliseconds); this feature, according to a number of scientists, makes capabilities of just a single neuron (especially in the part of recognizing visual images) comparable to those of personal computer [4-6]. Human neuron network organization always includes functional diversity of individual populations of the neuron networks constituting entire neuron network structure of human brain. Specifically, neural network organization includes the groups of neurons that present acceptors of signals both from the periphery and from within the neural network. The second important neuron population mainly performs processing of incoming signals and makes decisions further transferred, in the form of electrical signals, to the neural population responsible for specific control functions in the organism. Norbert Wiener [7] and academician Pyotr Anokhin [8] deserve the credit for inventing an obligatory link in neural networks in the form of feedback, both at the level of individual populations of neural networks, and at the system level of neuron network – effector (executive subsystems) [9-10].

Functional activity of neurons is determined by efficiency

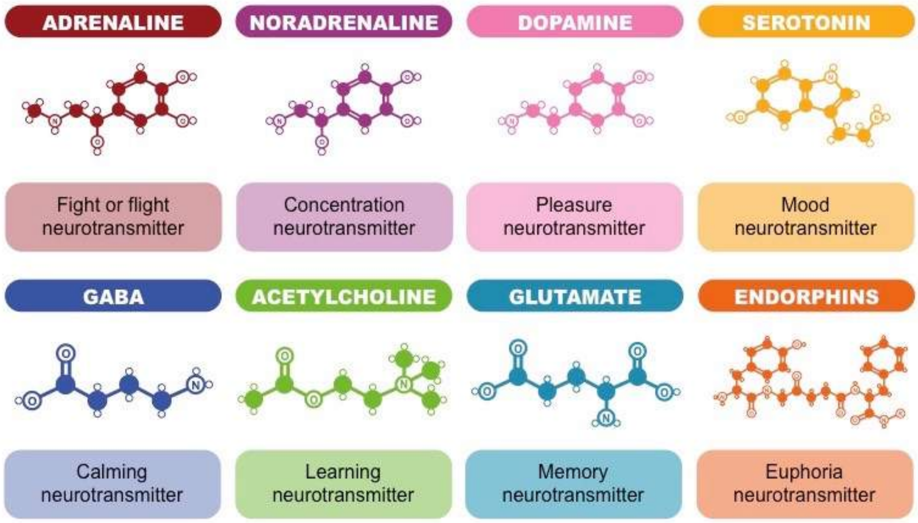


Figure 1. Chemical Structures of Neurotransmitters. Reproduced from http://www.compoundchem.com/2015/07/30/neurotransmitters/. Below is a list of the main neurotransmitters and their effect on the central nervous system of human: Adrenaline. Adrenaline is primarily a hormone released by the adrenal gland, but some neurons may secrete it as a neurotransmitter. Noradrenaline. In contrast to adrenaline, noradrenaline is predominantly a neurotransmitter that is occasionally released as a hormone. Dopamine. It is primarily responsible for feelings of pleasure, but is also involved in movement and motivation. Serotonin. Contributes to feelings of well-being and happiness. GABA. Inhibits neuron firing in the CNS – high levels improve focus whereas low levels cause anxiety. Acetylcholine. Involved in thought, learning and memory within the brain. Glutamate. Most common brain neurotransmitter. Endorphins. Release is associated with feelings of euphoria and a reduction in pain (body's natural 'pain killers').

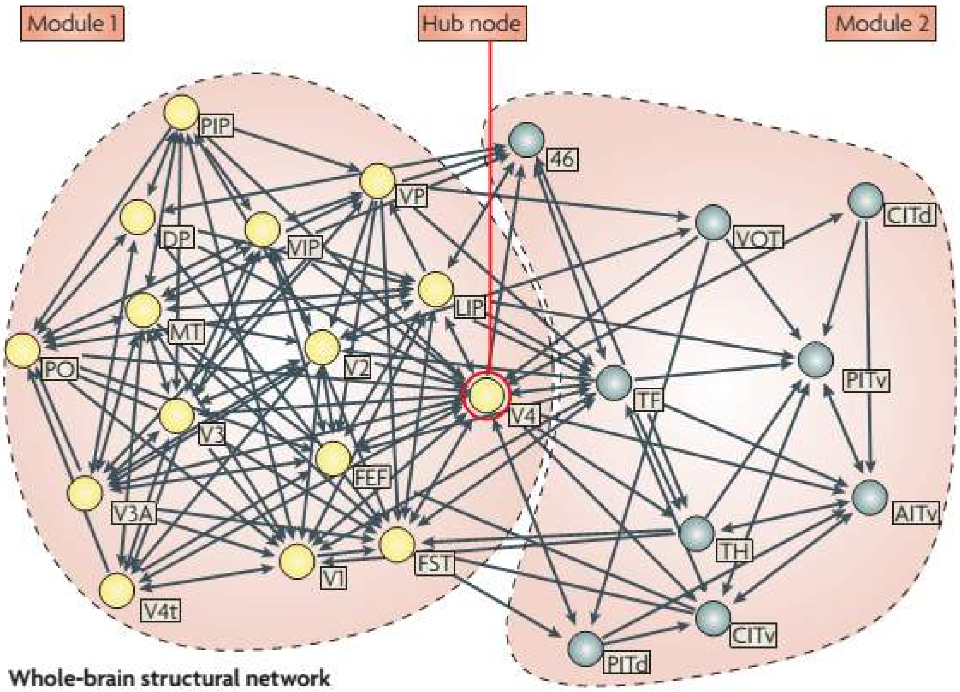


Figure 2. Whole-brain structural network constructed from histological data on the macaque cortex; each node corresponds to a brain area and the connections represent axonal projections between areas. Reproduced from [11].

of synaptic formations functioning. Plenty of signal molecules (from neurotransmitters till growth factors, amino acids and gas molecules) are involved into interneuron communications. By the way, every human has its own complex (pattern) of signal molecules which significantly determine intellectual features and personality. Unfortunately, namely the targets of signal molecules – neural networks and synapses – are the most sensitive to inflammatory factors, toxins, endogenous and synthetic neutrophilic substances exposure. Acute or chronic disruption of interneuron communications is always followed by failure in neural networks integrative work. Development of pathological processes in brain becomes the result of such failures. These processes may acquire functional character after initiation of plastic processes in nerve tissue including stem cells. On the other side, disturbance of brain plasticity leads to development of destructive transformations in neurons, synapses and glial cells. The process becomes irreversible, and patients usually visit doctor only at this stage. The standard treatment scheme is usually applied in clinical practice: medicines with neurotropic effect. Additional "intervention" in neural networks activity transforms their regular work. Pathological process becomes chronic. Then the doctor decides to change treatment scheme. But where to shift focuses in correction of neural networks functions? Special methods are applied in preclinical studies in order to assess the effects of new substances on functional state of neurons and interneuron communications. In particular, neurophysiological techniques allow quantitatively and dynamically establish the efficiency of signals transmission in nerve tissue and assess synaptic plasticity. Modern neurophysiological techniques of electrical events registration at the level of single synapse or part of neuron membrane (patch-clamp) increase opportunities in assessment of neural networks functions in health and disease. But there are difficulties in common analysis of revealed patterns at the levels of neuron membrane part and multicomponent neural network. Integrative analytic approach with computer modelling is needed. Otherwise the diversity of mechanisms of neurotropic substances action on central nervous system (from synapses and cells till integrative level of brain and spinal cord) cannot be ascertained.

Interneuron communications modelling at the level of integral brain is one of the technologies to answer the raised questions. Formation of such models became available after development of numerical methods of analysis of biophysical processes of data processing in biological neural networks. Data processing at this level requires special hardware and software design. For example, the project "Blue Brain" uses supercomputer Deep Blue (IBM) for detailed modelling of one part on cerebral cortex [12, 13]. One can imagine the required level of computer to model not the one part of cerebral cortex, but the functions of whole brain. That is why methods of computer modelling attract special attention of researchers and doctors.

In many cases models of neural networks with certain configuration are used in the analysis of experimental data. There is popular model [1] consisting of several hundred input neurons forming synaptic contacts with output neuron, which is used in numerous studies of synaptic plasticity investigation. This model considers spike-timing dependent plasticity, revealed in many brain regions [12]. Change of synaptic conductivity in the model is determined only by time parameters of pre- and postsynaptic activation [12-14]. Such relatively simple model provides insight into essential features of neural network. It is all about stability of output activity after change of different parameters and formation of competitive input signals during interaction of synapse groups. In this case, the increase in conductivity in one group of synapses is accompanied with the decrease in conductivity of other synapses.

Ideas of competitive learning appeared in theoretical studies dedicated to pattern recognition and self-management in neural networks [12, 13], as well as in modelling of the processes of topological mapping formation (correlation between receptor and the area of brain cortex) [12]. Hebb's theory of competitive learning is used in modelling of patterns formation, maps of brain cortex and selectivity columns and is characterized as "flexible, simple and useful" [12]. For example, only part of information presented at eye retina can be processed at any specific time (namely the information which is in the focus of attention). That is why populations of neurons responsible for mapping of certain patterns compete for priority of their information processing [12, 13]. The effect of medicinal substance in these models is considered as factor influencing on certain features of used model. Conditions of spike potentials generation and electrical processes at the level of pre- and postsynaptic membrane, characterizing synaptic transmission are considered the most important processes [15]. Influence of substance or other disturbing factors is manifests as change of threshold of spike potentials generation that can be easily detected experimentally and realized in model neuron. Main experimental studies in assessment of disturbing factors influence on the processes of synaptic plasticity are considered classic protocols of long-term potentiation induction based on the use of high frequency of train stimulation.

The rate of neural network learning is one of the parameters which characterize efficiency and productivity of neural network functioning. The rate of learning in models is determined by time needed for primary single-mode histogram of synaptic conductivities to become bimodal one. Such computer models adequately display real situations in living, but not computerized brain. High frequency of neural network activity and maximal conductivity of all synapses lead to "epileptiform" state, when the rate of learning increases under controlled conditions. The rate of learning decreases in the presence of pathological processes in brain or during uncontrolled use of neurotropic drugs; epileptiform activity accompanied with such motor effects as frank and absence seizures begins to appear in brain.

Therefore, the modelling of processes of biological neural networks functioning under conditions of neuromodulators factors action is promising technique for the development of new neuropharmacological substances action analysis. Above mentioned goes near development of adequate methods of correction of nerve system dysfunctions, because performing of unique calculations is one of brain features. Traditional conservative approaches cannot always afford tools suitable for analysis of biological neural networks functional activity. The time is ripe for integrative analytic approach with computer modelling use including topological approach to explain multisensory neurons functioning as nonlinear chaotic systems

[16, 17].

All neural network structures of the human brain mutually affect the functional state of each other (Figure 2). In addition, apart from the transmission of electrical signals by means of synapses (which ensures the targeting of information transfer) neuro network structures feature a principle of diffuse transmission of information through the intercellular matrix (glial cells and various substances filling the space between the neurons) (Figure 3). Association about the results of such diffuse transmission of information can be felt by anyone who experienced the emotions of unrestrained fear, danger, joy, grief, etc. [18-21].

One more concept of improving efficiency of functioning neural network structures is based on diffuse transmission of signals in the form of hormones, growth factors, cytokines that provide the so-called "emotional coloring" of the actions of neural network structures. By the way, emotional coloring of the process within the brain always features binary polarity as: yes/no, good/bad, safe /dangerous; this feature ensured the choice of appropriate conditions of existence in the process of evolution of living beings and, eventually, to minimize the risk of death. At the final stage of the discussions it makes sense to take a look at the problem of interaction of local neural networks with other neural network structures in the human brain and diversified elements of the surrounding world [23–40] (Figure 4). At this stage of the discussions the didactic warning made by Mary Shelly again rushed back upon us as we do not know what way artificially created neural networks structures (artificial intelligence) will interact with the surrounding human neural network structures. After all, if "yes/no, good/bad" concept has been implemented in the artificial intelligence we can hardly guess now what way the artificial intelligence will treat its creators.

II. Conclusion

For many readers this free style presentation of the ideas may seem rather romantic and narrative. However, the evolution of living nature evidences the survival of those creatures and beings that managed to adapt optimally to the surrounding conditions in the process of their development [8]. Therefore, when solving semantic, intellectual and informational problems at this stage of the development of artificial intelligence and complex semantic systems, it is necessary to look into the future, anticipating wherever possible the consequences of our bold decisions and innovations.

Acknowledgment

The authors gratefully acknowledge the assistance of Prof. V. Golenkov.

References

[1] Hughes P., Dragunow M. Induction of immediate-early genes and the control of neurotransmitter-regulated gene expression within the nervous system. *Pharmacological reviews*, 1995, vol. 47, No. 1, pp. 133-178.

[2] Steketee J. D. Neurotransmitter systems of the medial prefrontal cortex: potential role in sensitization to psychostimulants. *Brain Research Reviews*. 2003, vol. 41, No. 2, pp. 203-228.

[3] Myhrer T. Neurotransmitter systems involved in learning and memory in the rat: a meta-analysis based on studies of four behavioral tasks. *Brain Research Reviews*. 2003, vol. 41, No. 2, pp. 268-287.

[4] Bowers J. S. On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychological review*. 2009, vol. 116, No. 1, pp. 220.

[5] Foldiak P. Sparse coding in the primate cortex. The handbook of brain theory and neural networks. *MIT Press, Cambridge*, MA., 1995, pp. 777-780.

[6] Poggio T., Bizzi E. Generalization in vision and motor control. *Nature*. 2004, vol. 431, No. 7010, pp. 768-774.

[7] Wiener N. Cybernetics or Control and Communication in the Animal and the Machine. *MIT press*, 1961, vol. 25.

[8] Anokhin P.K., Shuleikina K.V. System organization of alimentary behavior in the newborn and the developing cat. *Dev. Psychobiol*, 1977, Vol. 10(5), pp. 385-419.

[9] Bullmore E., Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 2009, vol. 10, No. 3, pp. 186-198.

[10] Power J. D. et al. Functional network organization of the human brain. *Neuron*, 2011, vol. 72, No. 4, pp. 665-678.

[11] Sporns O., Honey PP. J., Kötter R. Identification and classification of hubs in brain networks. *PloS one*, 2007, vol. 2, No. 10, pp. e1049.

[12] Markram H. The blue brain project. *Nat. Rev, Neurosci*, 2006, Vol. 7, No. 2, P. 153–160.

[13] Hay E., Hill S., Schurmann F., Markram H., Segev I. Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Comput. Biol*, 2011, Vol. 7, No. 7, P. e1002107.

[14] Song S. Miller K.D., Abbott L.F. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci*, 2000, Vol. 3, No. 9, P. 919–926.

[15] Roberts P.D., Bell C.C. Spike timing dependent synaptic plasticity in biological systems. *Biological cybernetics*, 2002, Vol. 87, No. 5-6, P. 392–403.

[16] Peters J.F., Tozzi A. A topological approach explains multisensory neurons. *Technical Report*, 2015, DOI: 10.13140/RG.2.1.4491.7200.

[17] Peters J.F., Tozzi A. Topology plus logistic maps: nonlinear chaotic systems explained. *Technical Report*, 2015.

[18] Theocharis A. D. et al. Extracellular matrix structure. *Advanced drug delivery reviews*, 2016, vol. 97, pp. 4-27.

[19] Novak U., Kaye A. H. Extracellular matrix and the brain: components and function. *Journal of clinical neuroscience*, 2000, vol. 7, No. 4, pp. 280-290.

[20] Bellail A. PP. et al. Microregional extracellular matrix heterogeneity in brain modulates glioma cell invasion. T*he international journal of biochemistry and cell biology*, 2004, vol. 36, No. 6, pp. 1046-1069.

[21] Yamaguchi Y. Lecticans: organizers of the brain extracellular matrix. *Cellular and molecular life sciences*, 2000, vol. 57, No. 2, pp. 276-289.

[22] Vargová L., Syková E. Astrocytes and extracellular matrix in extrasynaptic volume transmission. *Phil. Trans. R. Sopp. B*, 2014, vol. 369, No. 1654, pp. 20130608.

[23] Wanner S. P. et al. Lipopolysaccharide-induced neuronal activation in the paraventricular and dorsomedial hypothalamus depends on ambient temperature. *PloS one*, 2013, vol. 8, No. 9, pp. e75733.

[24] Peters J. F. et al. The human brain from above: an increase in complexity from environmental stimuli to abstractions. *Cognitive Neurodynamics*, 2017, pp. 1-4.

[25] Déli E., Tozzi A., Peters J. F. Relationships between short and fast brain timescales. *Cognitive neurodynamics*, 2017, vol. 11, No. 6, pp. 539-552.

[26] Perin R., Telefont M., Markram H. Computing the size and number of neuronal clusters in local circuits. *Frontiers in Neuroanatomy*, 2013, vol. 7.

[27] Reimann M. W. et al. Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function. *Frontiers in Computational Neuroscience*, 2017, vol. 11, pp. 48.

[28] Clark A. Natural-born cyborgs?. Cognitive technology: Instruments of mind, 2001, pp. 17-24.

[29] Konar A. Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain. *CRC press*, 1999.

[30] Kennedy P. R. et al. Direct control of a computer from the human central nervous system. *IEEE Transactions on rehabilitation engineering*, 2000, vol. 8, No. 2, pp. 198-202.

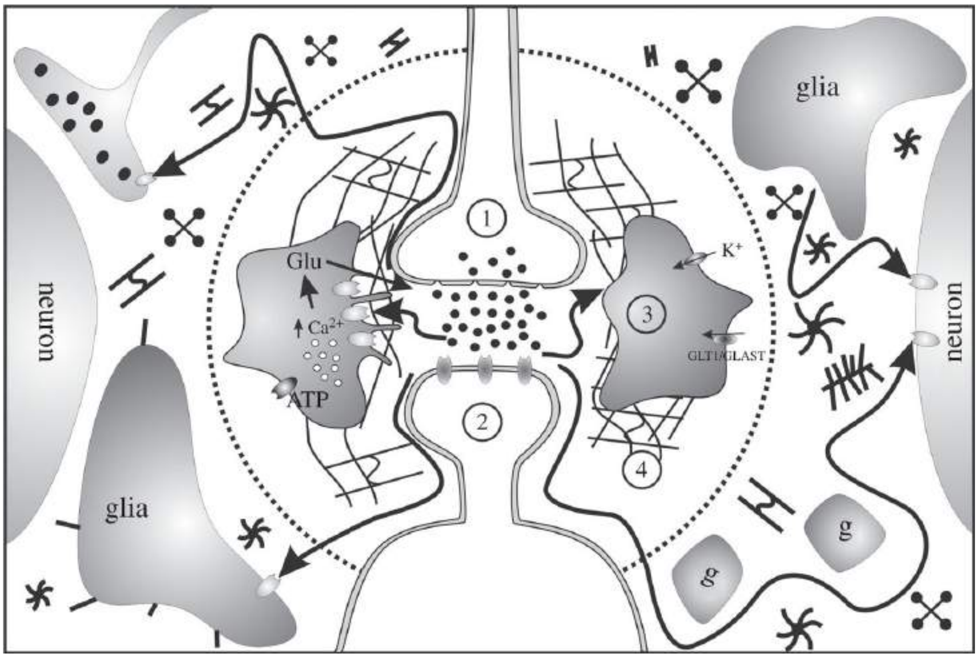


Figure 3. Schematic of intercellular communication via diffusion. Inside the dotted circle: short-distance diffusion of molecules mediates intercellular communication between presynaptic terminal (1), postsynaptic terminal (2) and glia and their processes ensheathing the synapse (3), and it is affected by the ECS properties and its content, particularly the ECM (4). Outside the dotted circle: neuroactive molecules, which have escaped from the synaptic cleft, reach high-affinity neuron or glia receptors at longer distance. This diffusion is restricted to the ECS volume and hindered by neuronal and astrocytic processes and ECM molecules. g, glial processes; Glu, glutamate; GLT1/GLAST, glutamate transporter. Reproduced from [22].

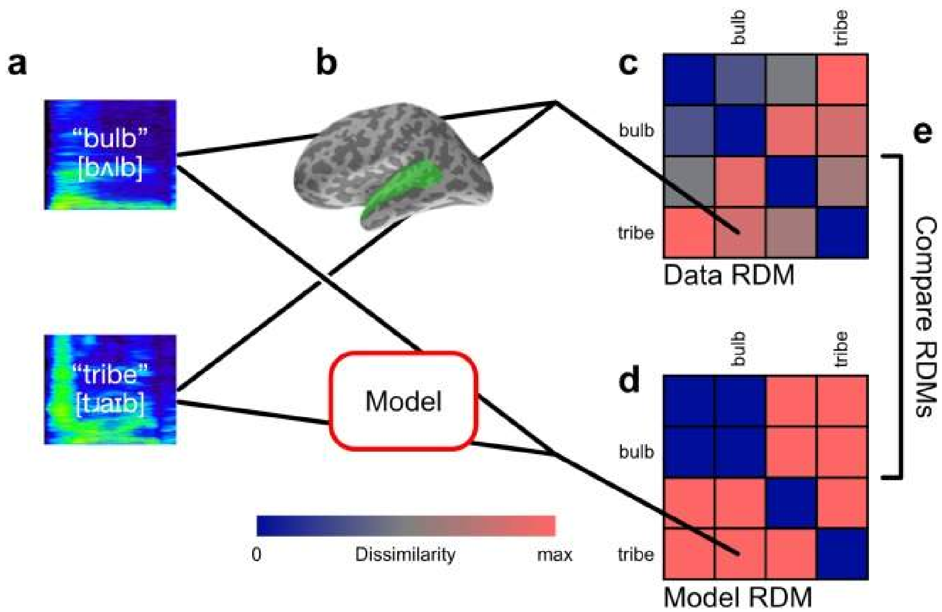


Figure 4. Representational similarity analysis (a) A set of experimental conditions or stimuli are presented to participants. In this example, recordings of English words are presented aurally. (b) For each experimental condition, EMEG data is collected from participants' regions of interest for a specified epoch. (c) Dissimilarities between each pair of responses are computed and stored in a representational dissimilarity matrix. Potential dissimilarity measures include Pearson's correlation distance or Euclidean distance between response vectors. Rows and columns of the matrix are indexed by the condition labels, making the matrix symmetric with diagonal entries all 0 by definition. In this example there are four conditions in total, and the responses to the condition pair (bulb, tribe) is compared, with the value stored in the indicated matrix entry, and its diagonally-symmetric counterpart. (d) A model of the experimental conditions or stimuli is used to compute a model RDM. The model RDM can be computed in several ways, e.g. by comparing representations of the stimuli under the model; or by modelling the dissimilarities directly. (e) Data and model RDMs are statistically compared, e.g. by computing Spearman's rank correlation of their upper-triangular vectors. Reproduced from [34].

[31] Hramov A. E. et al. Classifying the perceptual interpretations of a bistable image using EEG and artificial neural networks. *Frontiers in neuroscience*, 2017, vol. 11, pp. 674.

[32] Dewar A. D. M. et al. Neural coding in the visual system of Drosophila melanogaster: How do small neural populations support visually guided behaviours?. *PLoS Computational Biology*, 2017, vol. 13, No. 10, pp. e1005735.

[33] Nakajima K. et al. Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study. *European Journal of Nuclear Medicine and Molecular Imaging*, 2017, vol. 44, No. 13, pp. 2280-2289.

[34] Wingfield PP. et al. Relating dynamic brain states to dynamic machine states: Human and machine solutions to the speech recognition problem. *PLoS computational biology*, 2017, vol. 13, No. 9, pp. e1005617.

[35] Davydov M. V. et al. Anthropomorphic Control System for Robotic Complexes. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems],* 2017, pp. 321-324.

[36] Hassabis D. et al. Neuroscience-inspired artificial intelligence. *Neuron*, 2017, vol. 95, No. 2, pp. 245-258.

[37] Hajimani E., Ruano M. G., Ruano A. E. An Intelligent Support System for Automatic Detection of Cerebral Vascular Accidents from Brain CT Images. *Computer Methods and Programs in Biomedicine*, 2017.

[38] Tonutti M., Gras G., Yang G. Z. A machine learning approach for real-time modelling of tissue deformation in image-guided neurosurgery. *Artificial intelligence in medicine*, 2017, vol. 80, pp. 39-47.

[39] Zafar R., Dass S. PP., Malik A. S. Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion.*PloS one*, 2017, vol. 12, No. 5, pp. e0178410.

[40] Del Papa B., Priesemann V., Triesch J. Criticality meets learning: Criticality signatures in a self-organizing recurrent neural network. *PloS one*, 2017, vol. 12, No. 5, pp. e0178683.

НЕЙРОСЕТЕВЫЕ СТРУКТУРЫ: НАСТОЯЩЕЕ И БУДУЩЕЕ

М.В. Давыдов[1], А.Н. Осипов[1], С.Я. Килин[2], В.А. Кульчицкий[3]

[1]Белорусский государственный университет информатики и радиоэлектроники

[2]Институт физики имени Б.И.Степанова Национальной академии наук Беларуси

[3]Институт физиологии Национальной академии наук Беларуси

Представлены две основные точки зрения на пути создания искусственного интеллекта: разработка на основе тех закономерностей, которые прошла эволюция нейронов естественной нервной системы, и разработка на основе принципиально новых, отличающихся от путей природы закономерностей развития и конструирования нейросетевых структур и программ. Представлены основные структурные элементы будущего искусственного интеллекта, построенного на основе законов развития нейронных сетей человека: основные типы медиаторов и их роль в модуляции информативных сигналов; принцип диффузной передачи информации через межклеточный матрикс; принцип усиления эффективности работы нейросетевых структур за счет диффузной передачи сигналов в форме гормонов, ростовых факторов, цитокинов.

В нейронных сетях мозга человека доминирует принцип передачи сигналов электрическим путем по отросткам нервных клеток. Однако непосредственная акцепция сигнала осуществляется с помощью химических посредников – медиаторов, которые выделяются дискретно, в форме квантов вещества и обеспечивают модуляцию информационных сигналов. Учитывая огромное разнообразие посредников (медиаторов), число которых достигает десятков и сотен тысяч, можно представить огромное количество модулирующих способностей нейронных сетей мозга человека.

Организация нейронных сетей мозга человека всегда включает отдельные популяции нейронных сетей, из которых состоит вся нейросетевая структура мозга человека. Конкретно, в нейросетевой организации присутствуют группы нейронов, которые являются акцепторами сигналов как с периферии, так и внутренних сигналов нейросети. Вторая важная популяция нейронов преимущественно осуществляет обработку поступивших сигналов и принимает решения, которые в форме новых электрических сигналов поступают к популяции нейронов, ответственных за контроль конкретной функции в организме. По обратной связи от рецепторов на периферии, контролирующих эффективность выполнения функции, сигналы поступают в нейронные сети мозга, которые с учетом новой информации осуществляют коррекцию функции.

Все нейросетевые структуры мозга человека взаимно влияют на функциональное состояние друг друга. При этом помимо передачи электрических сигналов с помощью синапсов (что обеспечивает адресность передачи информации) в нейросетевых структурах присутствует принцип диффузной передачи информации через межклеточный матрикс (глиальные клетки и разнообразные субстанции, которые заполняют пространство между нейронами).

Есть еще один принцип усиления эффективности работы нейросетевых структур за счет диффузной передачи сигналов в форме гормонов, ростовых факторов, цитокинов которые обеспечивают так называемую «эмоциональную окраску» действий нейросетевых структур.

Указывается, что при создании искусственного интеллекта исследователи должны не забывать и о негативных последствиях, которые целесообразно предусмотреть и избежать уже на этапах семантического осмысления проблемы.

Patterns of Electrical Activity Generated by Biological Neural Network *in vitro*

Denisov A.A.,
Bulai P.M.,
Pitlik T.N.,
Molchanov P.G.,
Cherenkevich S.N.
Belarus State University
Minsk, Belarus
an.denisov@gmail.com

Pashkevich S.G.,
Dosina M.O.
Institute of Physiology
National Academy of Sciences
Minsk, Belarus
skypasht@mail.ru

***Abstract*—Cultured dissociated neurons forming network *in vitro* is a unique system representing living biological neural network developing in fully artificial conditions. This is a promising model for study of basic mechanisms of the brain functioning that requires special tools for interfacing and investigations. We have developed a set of devices and techniques for culturing of neural network on the surface of microelectrode sensor and registered specific patterns of electrical activity of the living neural network *in vitro*.**

***Keywords*—biological neural network, electrical activity patterns, electrophysiological data processing**

I. INTRODUCTION

Human brain can be considered as a highly sophisticated system for semantic processing [1], [2]. The basis of brain functioning is a complex neural network formed by biological cells. An ability to establish communications between neurons by means of electrical pulses and to process information in large neuronal ensembles are unique properties of neural tissue. There are numerous models of semantic processing based on neural networks [3], [4]. One particular class of these models is based on spiking neural networks [5] which closely resemble features of biological neural tissue. In this regard, studies of data processing mechanisms in brain have particular interest for development of intellectual systems design principles.

Mechanisms of brain functioning can be studied at different levels starting from molecular, cellular, network and up to behavioral and psychological. In the context of network level, cultured dissociated neurons forming connections *in vitro* is a unique system representing living biological neural network developing in fully artificial conditions. This is a promising model for study of basic mechanisms of the brain functioning that requires special tools for interfacing. Therefore, development of highly specialized investigation methods is required in order to obtain knowledge about deep mechanisms of neural ensembles functioning. We have developed a set of devices and techniques for culturing of neural network on the surface of microelectrode sensor and registered an electrical activity of the living neural network *in vitro*.

II. EXPERIMENTAL

We have developed the 64-channel microelectrode sensor of neuronal electrical activity suitable for multichannel interfacing with cultured neural networks. The sensor consists of planar glass base with transparent indium-tin-oxide conducting tracks serving as electrodes. The chamber for neuronal culture solution has been developed on the basis of 3D printing techniques as shown at the Fig. 1. For recording of electrical activity of neurons, a specialized 64-channel amplifier has been developed with built-in analog to digital converter and digital serial interface. Basic recording and visualization operations are controlled by the open-source software.

Dissociated neurons of the rat cortex were seeded on surface of a microelectrode array (64 electrodes) in recording chamber and placed in cell incubator. Cells formed dense network of interconnecting neurites after three weeks in culture as shown

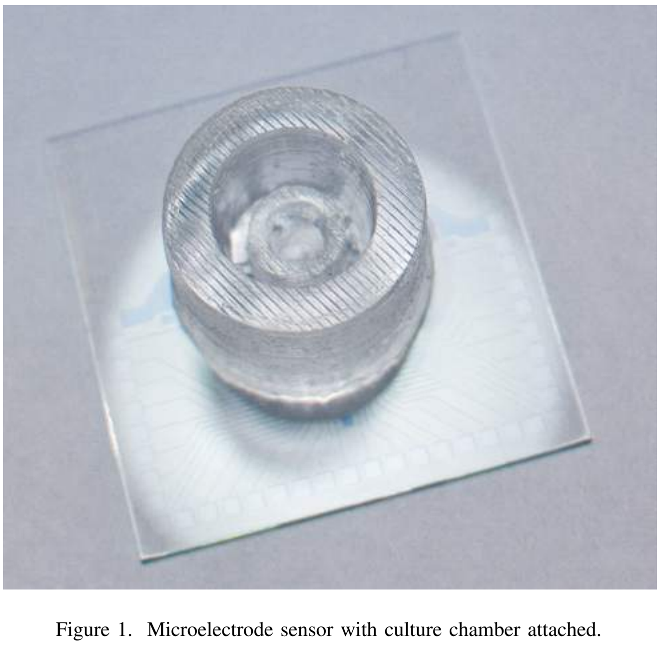


Figure 1. Microelectrode sensor with culture chamber attached.

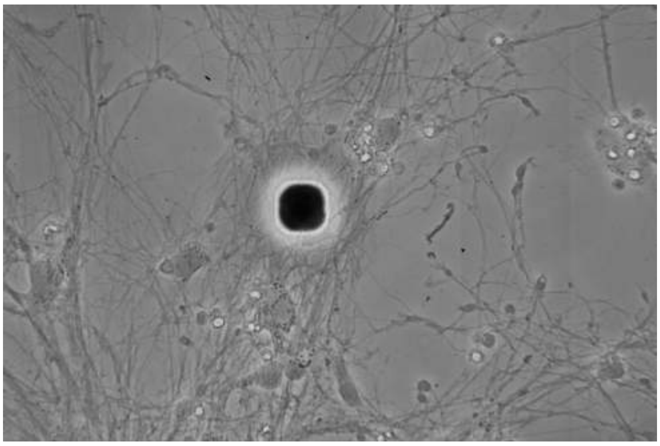


Figure 2. Neurons form dense network of interconnections around one of the working microelectrodes.

at the Fig. 2. Dark spot on the picture is electrode coated with opaque layer of electrodeposited conducting polymer poly(3,4-ethylenedioxythiophene) in order to reduce impedance and recording noise. Amplified extracellular activity of neurons was digitized and recorded to the hard disk of computer during experiments. The rate of data stream is about 10 gigabytes per hour in our case of 64 recording channels and 20 KHz sampling frequency so that specialized high-performance software tools for data processing are required. We used customized open-source module «Tridesclous» [6] for spike detection and classification. Spikes were distinguished from background noise on the basis of threshold crossing and then classified into clusters by principal component analysis.

III. Results

Typical examples of recorded multichannel neuronal electrical activity is shown at the Fig. 3. Each channel corresponds to different electrode of the array. Marks denote spikes divided into distinct clusters on the basis of amplitude and shape so that even different neurons near one electrode can be distinguished. Neuronal activity on the recordings is represented by background single sparse spikes and spike bursts. As shown at the Fig. 3a, a group of neurons at the channel 33 generates intense periodic bursts of spikes. A burst is composed by sum of repeating spikes of nearby neurons. Amplitudes of spikes are different due to different distances from electrode to neurons. This activity remains confined and spreads only partially and locally to channels 42 and 63.

Fig. 3b and Fig. 3c show another type of network behavior. Bursts at the channel 33 are followed by the bursts at the channel 63 and then less intense bursts and single spikes at the other channels. Therefore, activity generated by «pacemaker» group of neurons at the channel 33 propagates via neurites and synaptic connections over the network. Patterns of activation at Fig. 3b and Fig. 3c are similar but number of spikes in individual channels, exact timing of spikes and number of channels involved in network activation are different in sequential network bursts. This variability of the response is due to complex nature of biochemical processes running in living cells and thus constantly changing properties of individual neurons and synapses.

IV. Discussion

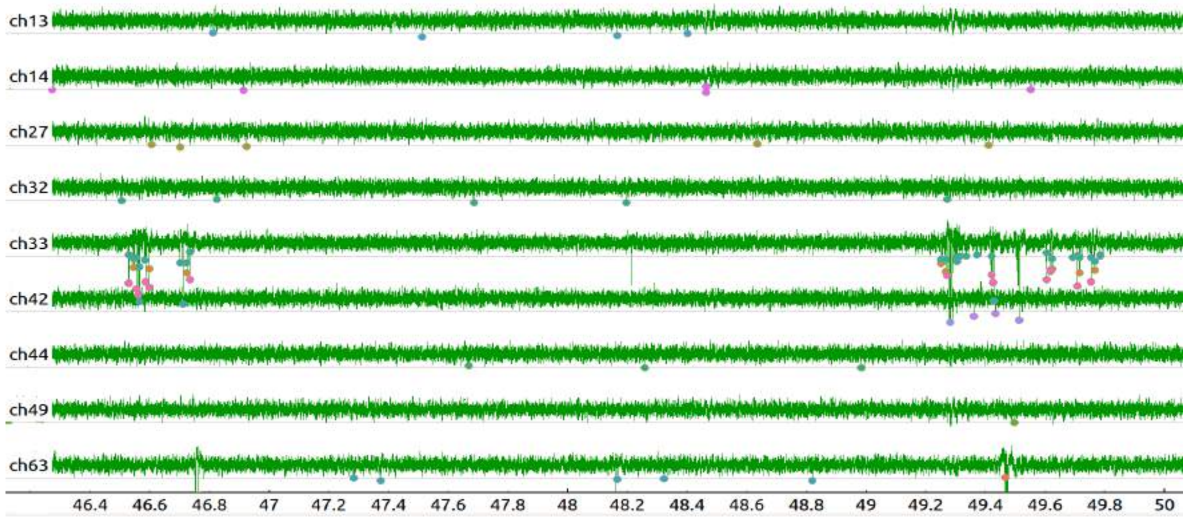
Analysis of electrical activity of biological neural networks is a complex problem including signal filtering, spike detection, classification and spike train analysis [7]. Many algorithms of spikes detection in background noise are based on threshold crossing. Choice of a threshold for spike detection is an important task because too high threshold will lead to losses of spikes, too low – to false positive detection of noise peaks as spikes. In our case, involvement of all channels in network events may indicate satisfactory functioning of the detection algorithm.

Spiking and bursting activation patterns observed in our experiments is a distinct feature of biological neural networks. Single sparse spikes between bursts may look like insignificant in network behavior, but investigations show that such single unit activity may play important role in shaping of bursting events [8]. Bursts of spikes increase reliability of neural information transmission and promote induction of synaptic efficacy changes [9]. Observed in our experiments two types of bursting (Fig. 3a and Fig. 3b, 3c) correspond to local and global network bursting indicating switching of the network behavior into different states.

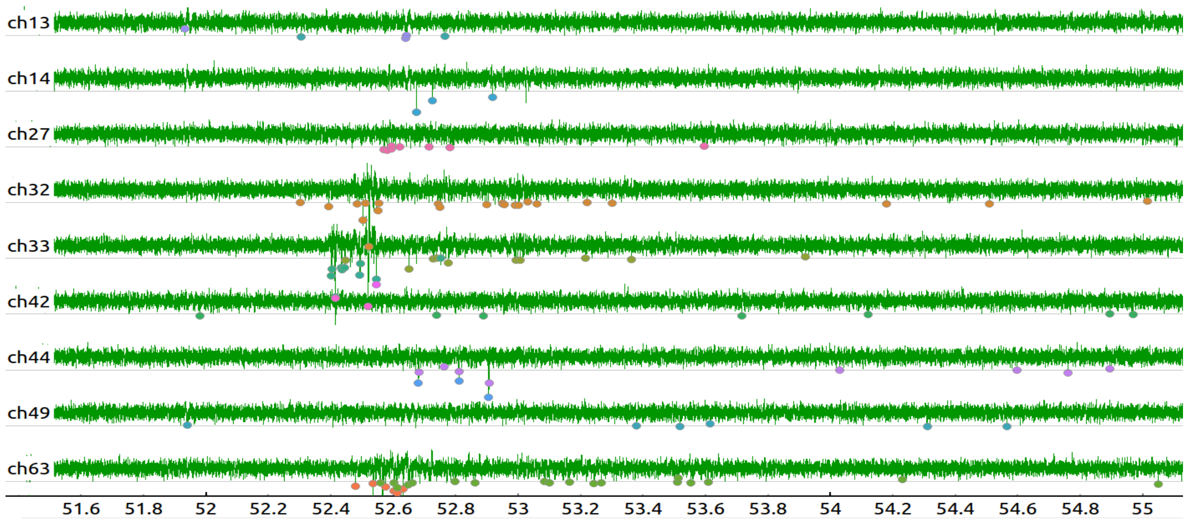
Using topological tools, it is possible to find subnetworks with repetitive and synchronous patterns of activity thus revealing connectivity features of network [10]. Data of [10] show that patterns of spiking propagation induced by external stimulation are topologically similar to spontaneous activation indicating that stimulating electrode can be treated as a biologically realistic input to the network.

A set of external stimulating electrodes can be treated as a vector input to the neural network and it is possible to attempt inducing learning processes [11] on the basis of biological learning rules [12]. In this regard, a question arises about detecting changes induced by learning protocols. Spontaneous activity can interfere with evoked one so several measures can be applied to reduce spontaneous activation such as pharmacological manipulation or low-frequency stimulation [13].

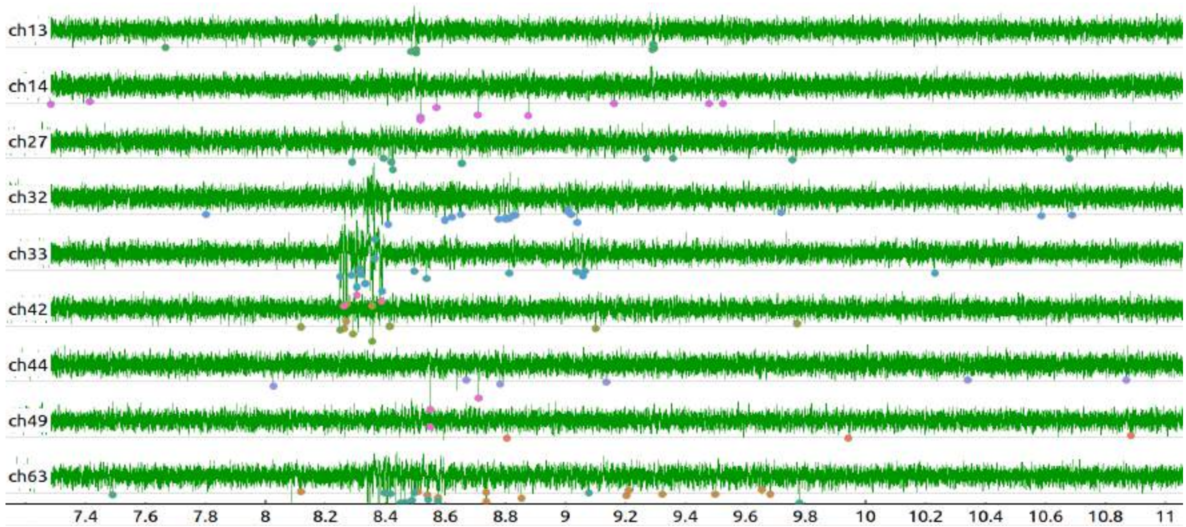
Despite confined simple structure, living network *in vitro* exhibit substantially complex behavior. One of the well-known and distinct features of developing *in vitro* neural networks is an extremely diverse set of electrical activity patterns [14]. Therefore, special techniques are required to distinguish between stochastic and determined network events. Different kinds of activity are distinguished such as spikes, avalanches and fluctuations [15] that can be treated as semantic representations of current network state. Activity patterns presented at Fig. 3b and Fig. 3c have similar overall features but different exact timing of individual spikes. Removal of stochastic part of the response and maximizing information from spike timing is important for analysis approach to completely describe network state.



(a)



(b)



(c)

Figure 3. Examples of electrical activity recorded at different time periods. (a) - Bursting activity at the channel 33. (b),(c) - Activity propagation through the network. Y axis - number of recording channel, X axis - time in seconds from recording frame start.

One of the approaches to reduce variability of electrical activity and to improve response predictability can be based on formation of ordered neural network [16].

The experimental model considered is two-dimensional network with because it is growing on a planar substrate. Modern cell culture techniques offer possibilities to engineer three-dimensional cell structures with topology resembling in vivo neural networks. Such network *in vitro* has more longer and complex response to external stimulation and spontaneous activity is characterized by spatial segregation of bursting with absence of global synchrony [17].

Increasing scale and complexity of modern electrophysiological experiments indicate need for new tools of 'big data' processing [18]. New approaches can be based on powerful tools inspired by intellectual systems designs such as deep learning [19], [20].

V. Conclusion

Diverse natural activity patterns of biological neural networks cultured on planar microelectrode arrays require utilization of sophisticated techniques in order to capture high-order features of the network state on the basis of recorded extracellular potentials. Several tasks are being solved for this purpose such as development of techniques for analysis of complex spatiotemporal patterns of network activation in order to detect response features induced by learning protocols.

Biological neural network *in vitro* is considered as a perspective experimental model for study of the mechanisms of learning. The ultimate goal of experiments with such networks is a reproduction of learning and memory processes specific to brain. However, in spite of extensive scientific work during past two decades this goal currently is not considered as clearly achieved [21]. Our current work is directed towards inducing learning processes by application of appropriate stimulating patterns to the patterned cultured neural network.

Acknowledgment

This work is supported by "Science around us" Foundation (London)

References

[1] A. Martin and L. L. Chao, “Semantic memory and the brain: structure and processes,” *Curr. Opin. Neurobiol.*, vol. 11, no. 2, pp. 194–201, apr 2001.

[2] J. R. Binder and R. H. Desai, “The neurobiology of semantic memory,” *Trends Cogn. Sci.*, vol. 15, no. 11, pp. 527–536, nov 2011.

[3] R. Collobert and J. Weston, “A unified architecture for natural language processing,” in *Proc. 25th Int. Conf. Mach. Learn. - ICML '08*. New York, New York, USA: ACM Press, 2008, pp. 160–167.

[4] G. E. Dahl, Dong Yu, Li Deng, and A. Acero, “Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition,” *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 20, no. 1, pp. 30–42, jan 2012.

[5] S. G. Wysoski, L. Benuskova, and N. Kasabov, “Evolving spiking neural networks for audiovisual information processing,” *Neural Networks*, vol. 23, no. 7, pp. 819–835, sep 2010.

[6] “Tridesclous.” [Online]. Available: https://github.com/tridesclous/tridesclous

[7] H. G. Rey, C. Pedreira, and R. Quian Quiroga, “Past, present and future of spike sorting techniques,” *Brain Res. Bull.*, vol. 119, pp. 106–117, oct 2015.

[8] M. Shein, V. Volman, N. Raichman, Y. Hanein, and E. Ben-Jacob, “Management of synchronized network activity by highly active neurons,” *Phys. Biol.*, vol. 5, no. 3, p. 036008, sep 2008.

[9] L. Chen, Y. Deng, W. Luo, Z. Wang, and S. Zeng, “Detection of bursts in neuronal spike trains by the mean inter-spike interval method,” *Prog. Nat. Sci.*, vol. 19, no. 2, pp. 229–235, feb 2009.

[10] V. Pirino, E. Riccomagno, S. Martinoia, and P. Massobrio, “A topological study of repetitive co-activation networks in in vitro cortical assemblies.” *Phys. Biol.*, vol. 12, no. 1, p. 016007, jan 2015.

[11] J. le Feber, J. Stegenga, and W. L. C. Rutten, “The Effect of Slow Electrical Stimuli to Achieve Learning in Cultured Networks of Rat Cortical Neurons,” *PLoS One*, vol. 5, no. 1, p. e8871, jan 2010.

[12] H. Markram, W. Gerstner, and P. J. Sjöström, “Spike-timing-dependent plasticity: a comprehensive overview.” *Front. Synaptic Neurosci.*, vol. 4, p. 2, 2012.

[13] C. M. Hales, J. D. Rolston, and S. M. Potter, “How to Culture, Record and Stimulate Neuronal Networks on Micro-electrode Arrays (MEAs),” *J. Vis. Exp.*, no. 39, pp. e2056–e2056, may 2010.

[14] D. A. Wagenaar, J. Pine, and S. M. Potter, “An extremely rich repertoire of bursting patterns during the development of cortical cultures.” *BMC Neurosci.*, vol. 7, no. 1, p. 11, feb 2006.

[15] G. Gigante, G. Deco, S. Marom, and P. Del Giudice, “Network Events on Multiple Space and Time Scales in Cultured Neural Networks and in a Stochastic Rate Model,” *PLOS Comput. Biol.*, vol. 11, no. 11, p. e1004547, nov 2015.

[16] B. C. Wheeler and G. J. Brewer, “Designing Neural Networks in Culture: Experiments are described for controlled growth, of nerve cells taken from rats, in predesigned geometrical patterns on laboratory culture dishes.” *Proc. IEEE. Inst. Electr. Electron. Eng.*, vol. 98, no. 3, pp. 398–406, mar 2010.

[17] M. Frega, M. Tedesco, P. Massobrio, M. Pesce, and S. Martinoia, “Network dynamics of 3D engineered neuronal cultures: a new experimental model for in-vitro electrophysiology.” *Sci. Rep.*, vol. 4, p. 5489, jun 2014.

[18] Y. Chen, Z.-y. Wang, G. Yuan, and L. Huang, “An overview of online based platforms for sharing and analyzing electrophysiology data from big data perspective,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 7, no. 4, p. e1206, jul 2017.

[19] M.-P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, and D. Pompili, “Cloud-based deep learning of big EEG data for epileptic seizure prediction,” in *2016 IEEE Glob. Conf. Signal Inf. Process.* IEEE, dec 2016, pp. 1151–1155.

[20] S. M. Plis, D. R. Hjelm, R. Salakhutdinov, E. A. Allen, H. J. Bockholt, J. D. Long, H. J. Johnson, J. S. Paulsen, J. A. Turner, and V. D. Calhoun, “Deep learning for neuroimaging: a validation study,” *Front. Neurosci.*, vol. 8, p. 229, aug 2014.

[21] Y.-T. Huang, Y.-L. Chang, C.-C. Chen, P.-Y. Lai, and C. K. Chan, “Positive feedback and synchronized bursts in neuronal cultures,” *PLoS One*, vol. 12, no. 11, p. e0187276, nov 2017.

ПАТТЕРНЫ ЭЛЕКТРИЧЕСКОЙ АКТИВНОСТИ БИОЛОГИЧЕСКОЙ НЕЙРОННОЙ СЕТИ in vitro

Денисов А.А., Булай П.М., Питлик Т.Н., Молчанов П.Г., Досина М.О., Пашкевич С.Г., Черенкевич С.Н.

Культивируемые диссоциированные нейроны, образующие синаптические связи in vitro, являются уникальной системой, представляющую собой живую биологическую нейронную сеть, развивающуюся в полностью искусственных условиях. Это многообещающая модель для изучения основных механизмов функционирования мозга, которая требует специальных инструментов для исследования и взаимодействия. Нами разработан комплекс устройств и методов для культивирования нейронной сети на поверхности микроэлектродного сенсора и получены специфические закономерности электрической активности живой нейронной сети in vitro.

Approach to determining the structural similarity of software projects

Gleb Guskov
Information system department
Ulyanovsk State Technical University
Ulyanovsk, Russia
guskovgleb@gmail.com

Alexey Namestnikov
Information system department
Ulyanovsk State Technical University
Ulyanovsk, Russia
am.namestnikov@gmail.com

***Abstract*—The paper proposes a method for comparing software projects on the basis of their structural similarity. An approach is proposed for obtaining an ontological representation of the project structure based on the source code. The paper considers several metrics for calculating the structural similarities applicable for different types of projects.**

***Keywords*—ontology, conceptual model, engineering design, hierahical clusstering**

I. Introduction

Human resources in modern software development are basic. Often the same tasks are solved several times, this leads to an ineffective waste of time. There are approaches for reusing source code at various stages of development. Basically, these approaches allow you to reuse only certain functions and classes. But such approaches do not allow finding the similarity of projects on the basis of their subject area.

Knowledge gained from already implemented projects in one subject area will allow borrowing and reusing much larger parts of projects and avoiding conceptually incorrect solutions. Quite often when updating the composition of developers, the implemented software solutions are forgotten and not used.

A tool capable of determining the similarities between projects can be very useful in software development. This tool can not be replaced by a version control system, because it only provides storage of all versions of the project with comments. Version control system provides comparison beetween file versions. Such an instrument will be able to work not only with projects from single organization, but also with projects from open repositories.

Search on open repositories is carried out on the basis of keywords. The results of such a search can be several thousand projects, which can not be handled by hand. The choice of projects based on their architectural solutions is a promising approach. To implement filtering on projects based on their architecture, you need to be able to analyze it. The architecture of the software project is built at the stage of its design, prior to the development. Elements of the UML language were developed to describe the architecture of the project with the required level of detail. Basing on the results of our previous research [1], we can conclude that the developers have different types of structural elements. Ontology as a knowledge storage system could well act as a reference for the project analysis tool. Attempts to integrate ontologies into software development were carried out at different levels: technical documents [2]–[7], maintenance and testing of the source code [8], the UML diagrams [9]–[13].

The minimal structural elements of UML, such as classes, interfaces, objects themselves, weakly convey the semantics and architectural solutions of the project. The combination of such elements is much better describes the architecture of the project. Stable combination of structural elements, known as design patterns. The design patterns appeared relatively long time in information technology and are still relevant. Design patterns are actively used by the developer community, thus representing a reliable benchmark in the analysis of the project. In addition, it makes sense to create local design patterns that solve this or that task in a given subject area. A template based on a specific subject area loses its main property - universality, but its semantic weight is a more important metric for solving the problem of constructing a tool for finding similarities between projects. There are many works devoted to the integration of software development with ontologies. There is a whole approach to development, based on a domain known as development based on the subject area [14]–[16].

II. Formulation of the problem

The results presented in this paper are a continuation of the work described in research[1]. The system described in research made it possible to extract information from conceptual models and save it as an ontology of a certain format. But the life cycle and manufacturing practices of IT companies show that conceptual models are created at best at each stage of the project, and in the worst once at the beginning of the project.

The state of the project is best described by the source code of the project. Developers try to maintain the source code in good condition, create documentation, provide comments and opportunely refactor. Another advantage of the source code is the widespread introduction of version control systems. Tracking all versions of software products allows effectively

The authors acknowledge that this paper was supported by the project no. 16-47-730742 and 16-47-732033 of the Russian Foundation for Basic Research.

manage the development of software and generates a large amount of information available for processing.

Information for comparison obtained from conceptual models of a new project at the design stage and information obtained from the source code of projects that have already been implemented will allow to determine the structural similarity of projects.

To get the projects structural similarity, it is necessary to translate information about projects from different sources to a single format. It is most convenient to present the extracted information in the form of ontology in the language of OWL. OWL ontology will allow preserving the semantics of complex architectural solutions, allowing to make changes to already existing data and to perform logical operations on statements.

The search for structural similarity of projects is part of the project comparison method. In addition to comparing the structure of projects, it is planned to compare still subject areas of projects. If a comparison is made between projects of the same enterprise in one subject area, then the comparison should be performed at the level of the processes and the components of the subject area. If the comparison of the project is carried out among the projects placed on the open repository, the structural similarity of the projects is more important metric then general subject area.

III. UML META-MODEL BASED ONTOLOGY

As a structure for storing UML class diagrams was chosen an OWL ontology, because this format is the most expressive for representation of knowledge from complex subject areas. The class diagram elements should be translated into ontology as concepts with considering to their semantics. Semantics of the whole diagram is being formed from the semantics of diagram elements and the semantics of their interaction. That is why the ontology was built on the basis of the UML meta-scheme, not a formal set of translated elements.

To solve the problem of intellectual analysis of project diagrams, included in the project documentation, it is necessary to have knowledge in the area of construction of formalized diagrams.

Ontology contains concepts that describe the most basic elements of the class diagram, but it can be expanded if necessary. When translating the meta-scheme of UML, the following notations were applied.

Formally, the ontology of project diagrams is represented as a set:

$$O^{prj} = \langle C^{prj}, R^{prj}, F^{prj} \rangle, \tag{1}$$

where : $C^{prj} = \{c_1^{prj}, ...c_i^{prj}\}$ – is a set of concepts that define main UML diagram elements such as : "Class", "Object", "Interface", "Relationship" and others;
R^{prj} – the set of connections between ontology concepts. These relationships allow us to correctly describe the rules of UML notation.
F^{prj} – is the set of interpretation functions defined on the relationships R^{prj}

IV. DESIGN PATTERNS AS STRUCTURAL PARTS OF SOFTWARE PROJECTS

Design patterns are insert into ontology as a set of individuals based only on the ontology classes described above. Semantic constraints and properties of design patterns are specified with by the ObjectProperties and DatatypeProperties of OWL ontology. Since many design patterns are stored in the ontology at the same time, it is necessary to enter rules of naming for their elements to avoid duplicate names. The name of the design pattern element begins with the design pattern name, and then if the element is the class, its name is written. If the element is a relationship, then the names of the elements that it connects are written through the underscore. One of the most commonly used design patterns is the Builder [17].

Builder is a creational pattern. The Builder pattern separates the algorithm for the step-by-step construction of a complex object from its external representation so that it is possible to obtain different representations of this object using the same algorithm.

In order to preserve this design pattern in the developed ontology, the following individuals were required.

- SimpleClass: Builder_Client, Builder_Director, Builder_ConcreteBuilder, Builder_Product.
- AbstractClass: Builder_AbstractBuilder.
- Association: Builder_Client_AbstractBuilder, Builder_Client_Director, Builder_Client_IProduct, Builder_ConcreteBuilder_Product.
- Generalization: Builder_ConcreteBuilder_AbstractBuilder.
- Realization: Builder_Product_IProduct.

Ontological representation of the design pattern:

$$O_{tmp_i}^{prj} = \{inst(C_1^{prj}), ...inst(r_1^{prj}), ..., r_{sameAs}\}, \tag{2}$$

In fact, the ontological representation of a single design pattern is a set of instances of concepts and relations from the ontology of project diagrams.

To calculate the structural similarity of projects based on ontology, the following expressions were proposed. The first metric gives priority to the maximum single expressed design pattern in both diagrams:

$$\mu_{dc_\gamma, dc_\delta} = \bigvee_{tmp \in (dc_\gamma \cap dc_\delta)} \mu_{dc_\gamma \cap dc_\delta}(tmp), \tag{3}$$

where dc_γ and dc_δ is project class diagrams presented as UML metamodel ontology Abox expressions,
$\mu_{dc_\gamma, dc_\delta (tmp)}$ - measure of similarity design pattern in project diagram.
The second metric considers the coincidence of all design patterns in equal proportions and does not considers design patterns with a measure of expression less than 0.3:

$$\mu_{dc_\gamma, dc_\delta} = (\sum_{tmp \in (dc_\gamma \cap dc_\delta) \geq 0.3} \mu_{dc_\gamma \cap dc_\delta})/N, \tag{4}$$

where N - count of design patterns with a measure of expression greater than 0.3 for each of both projects.

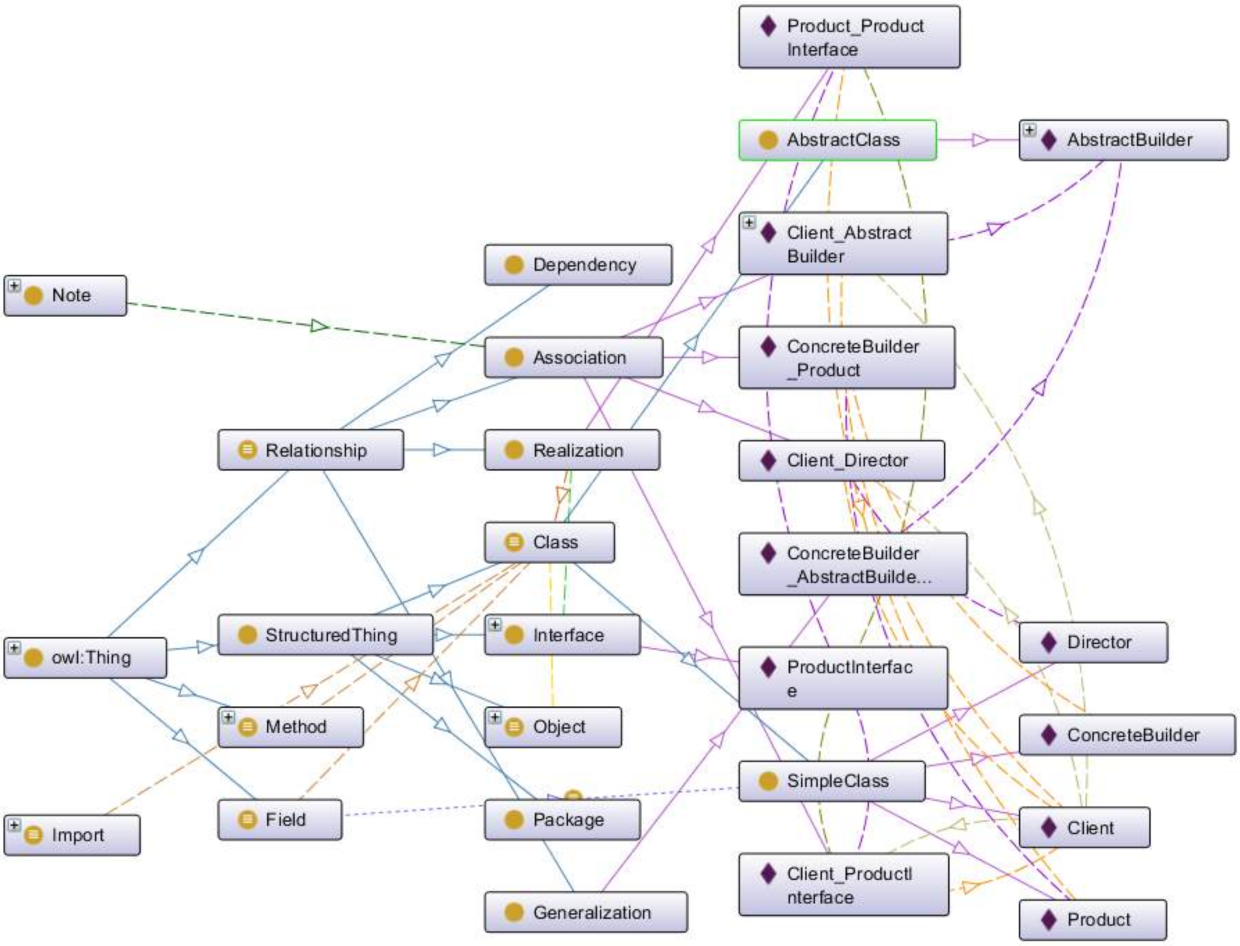


Figure 1. "Builder" design pattern ontology presentation in Protege editor.

The third metric works the same as the second one, but the contribution to the evaluation by design patterns depends on the number of elements in the design pattern (the design pattern with 20 elements means more than a design pattern with 5 elements):

$$\mu_{dc_\gamma, dc_\delta} = \Big(\sum_{tmp \in (dc_\gamma \cap dc_\delta) \geq 0.3} \widetilde{\mu}_{dc_\gamma \cap dc_\delta} \Big) / N, \qquad (5)$$

where $\widetilde{\mu}_{dc_\gamma \cap dc_\delta}$ - weighted measure of expression.

V. The results of searching structurally similar software projects

A. *Searching design patterns in projects*

To determine the measure of similarity between the two projects, it is necessary to calculate the degree of expression of the each design pattern in each of the project. The measure of the expression of the design pattern in the project can be calculated by mapping a project ontology Abox on a design pattern ontology Abox. The Table I contains degree of expression of the each design pattern in each of the project.

B. *Results of searching structurally similar software projects by different metrics*

The results of calculating the similarity between projects by three metrics are presented in the Table II.

Estimates based on the results of comparison projects are quite high. The estimates are normalized from 0 to 1. The estimates for the first metric are always equal to 1. This is easy to explain, because it chooses the most expressed design pattern in both projects. Among the design patterns participating in the test there are templates with a small number of elements, for example: Absatrast superclass, interface and delegator. The results for the second and third metrics are also quite high. Design patterns with a degree of expression less than 0.3 are excluded from consideration. All projects that participated in the comparison are downloaded from the open repository Github and somehow interact with the public API of the well-known social network vkontate.

Table I
EXPRESSION OF DESIGN PATTERNS IN PROJECTS

Project name / Design pattern name	Delegator (3)	Adapter (8)	Builder (12)	Abstract superclass (3)	Interface (5)
Android-MVP	1.0	0.875	0.83	1.0	1.0
cordova-social-vk	1.0	0.875	0.83	1.0	0.8
cvk	1.0	0.875	0.83	1.0	0.8
DroidFM	1.0	0.875	0.92	1.0	1.0
VK-Small-API	1	0.625	0.42	0.33	0.6
VKontakteAPI	1.0	0.875	0.83	1.0	0.8
VK_TEST	1	0.75	0.58	0.66	0.6

Table II
SIMILARITY BETWEEN PROJECTS

<table>
<tr><th>First project / Second project</th><th>Android-MVP</th><th>cordova-social-vk</th><th>cvk</th><th>DroidFM</th><th>VK-Small-API</th><th>VKontakteAPI</th><th>VK_TEST</th></tr>
<tr><td>Android-MVP</td><td>–</td><td>1 | 0.96 | 0.96</td><td>1 | 0.96 | 0.96</td><td>1 | 0.98 | 0.96</td><td>1 | 0.78 | 0.64</td><td>1 | 0.96 | 0.96</td><td>1 | 0.78 | 0.77</td></tr>
<tr><td>cordova-social-vk</td><td>1 | 0.96 | 0.96</td><td>–</td><td>1 | 1 | 0.99</td><td>1 | 0.94 | 0.93</td><td>1 | 0.85 | 0.67</td><td>1 | 1 | 0.99</td><td>1 | 0.83 | 0.80</td></tr>
<tr><td>cvk</td><td>1 | 0.96 | 0.97</td><td>1 | 1 | 0.99</td><td>–</td><td>1 | 0.94 | 0.93</td><td>1 | 0.85 | 0.67</td><td>1 | 1 | 0.99</td><td>1 | 0.83 | 0.80</td></tr>
<tr><td>DroidFM</td><td>1 | 0.98 | 0.97</td><td>1 | 0.94 | 0.93</td><td>1 | 0.94 | 0.93</td><td>–</td><td>1 | 0.78 | 0.61</td><td>1 | 0.94 | 0.93</td><td>1 | 0.78 | 0.74</td></tr>
<tr><td>VK-Small-API</td><td>1 | 0.78 | 0.64</td><td>1 | 0.85 | 0.67</td><td>1 | 0.85 | 0.68</td><td>1 | 0.78 | 0.61</td><td>–</td><td>1 | 0.85 | 0.67</td><td>1 | 0.96 | 0.87</td></tr>
<tr><td>VKontakteAPI</td><td>1 | 0.96 | 0.97</td><td>1 | 1 | 0.99</td><td>1 | 1 | 0.99</td><td>1 | 0.94 | 0.93</td><td>1 | 0.85 | 0.67</td><td>–</td><td>1 | 0.83 | 0.80</td></tr>
<tr><td>VK_TEST</td><td>1 | 0.79 | 0.77</td><td>1 | 0.83 | 0.80</td><td>1 | 0.83 | 0.80</td><td>1 | 0.78 | 0.74</td><td>1 | 0.95 | 0.87</td><td>1 | 0.83 | 0.80</td><td>–</td></tr>
</table>

CONCLUSION

The work presented in this paper have great potential for further research. Number of projects could be expanded. It is possible to include new design patterns in consideration. Ontologies obtained in the intermediate stages could be used separately in Protege editor. Expanding the system by using ontologies of subject areas can significantly increase the relevance of the similar projects selection.

REFERENCES

[1] G.Guskov A. Namestnikov, *Ontological mapping for conceptual models of software system*, OSTIS-2017 proceedings, 2017.

[2] A.Namestnikov , A. Filippov, V. Avvakumova, *An ontology based model of technical documentation fuzzy structuring.*, CEUR Workshop Proceedings. SCAKD 2016. Moscow. Russian Federation. Volume 1687, pp. 63-74 (2016).

[3] A.Goy, D. Magro, *Towards an ontology-based software documentation management - a case study.*, KMIS (Kecheng Liu and Joaquim Filipe, eds.), SciTePress, pp. 125–131 (2012).

[4] A. Koukias, D.Kiritsis, *Rule-based mechanism to optimize asset management using a technical documentation ontology*, IFAC-PapersOnLine 48, no. 3, 1001 – 1006 (2015).

[5] U. Zagorulko, I. Ahmadeeva, A. Serii, V. Shestakov, *Postroenie tematicheskih intellektualnih nauchnih internet-resursov sedstvami semantic web*, Trudi 15 nacionalnoi konferencii po iscusstvennomu intelektu KII-2016, Smolensk, vol 2, 2016, pp. 47-55.

[6] A. Namestnikov, G. Guskov, *Programnaya sistema preobrazovaniya UML-diagram v ontologii na yazike OWL* , Trudi 15 nacionalnoi konferencii po iscusstvennomu intelektu KII-2016, Smolensk, vol 3, 2016, pp. 270-278.

[7] Filippov A.A., Moshkin V.S., Shalaev D. O., Yarushkina N.G. *Edinaya ontologicheskaya platforma intellectualnogo analiza dannih*// Materiali VI mejgunarodnoi nauchno-tehnicheskoikonferencii OSTIS-2016, Minsk, Respublica Belarus, 2016.

[8] Hossein S., Sartipi K., *Dynamic analysis of software systems using execution pattern mining.*, ICPC, IEEE Computer Society, pp. 84–88 (2006).

[9] Ma, Z., Zhang, F., Yan, L., Cheng, J. *Representing and reasoning on fuzzy UML models: A description logic approach.* Expert Systems with Applications,38(3), 2536–2549 (2011).

[10] J. Zedlitz, J. Jorke,N. Luttenberger, *From UML to OWL 2*, In: Proceedings of Knowledge Technology Week 2011. Springer (2012).

[11] Almeida Ferreira D., Silva A., *UML to OWL Mapping Overview An analysis of the translation process and supporting tools.* Conference: 7th Conference of Portuguese Association of Information Systems(2013).

[12] Bobillo F., Straccia U., *Representing Fuzzy Ontologies in OWL 2*, WCCI IEEE World Congress on Computational Intellegence, Barcelona, Spain, July 18-23, pp.2696- 2700. (2009)

[13] N.Dorodnich i A. Urin, *Ispolzovanie diagram klassov UML dlya formirovaniya produccionnih baz znanii*, Programnaya ingeneriya vol. 4, 2015.

[14] Wongthongtham P., Pakdeetrakulwong U.,Marzooq S., *Ontology annotation for software engineering project management in multisite distributed software development environments*, pp. 315–343, Springer International Publishing, Cham, 2017.

[15] Emdad A., *Use of ontologies in software engineering.*, SEDE (Hisham Al-Mubaid and Rym Zalila-Wenkstern, eds.), ISCA, pp. 145–150 (2008).

[16] Dillon T., Chang E., Wongthongtham P., *Ontology-based software engineering- software engineering 2.0.*, Australian Software Engineering Conference, IEEE Computer Society, pp. 13–23 (2008).

[17] Mark Grand, *Java enterprise design patterns: Patterns in java (patterns in java)*, John Wiley and Sons, 2002

ПОДХОД К ПОИСКУ ПРОГРАММНЫХ ПРОДУКТОВ СО СХОЖЕЙ СТРУКТУРОЙ

Глеб Гуськов, Алексей Наместников
Кафедра «Информационные системы»
Ульяновский государственный технический университет, Россия, Ульяновск

В статье описан метод сравнения программных продуктов на основе их структурного сходства. Подход основан на использовании онтологических представлений структуры программных продуктов. Структура программных продуктов извлекается из исходного кода или концептуальных моделей. В статье рассматривается несколько различных метрик для расчёта структурного сходства, которые могут быть использованы для проектов различных типов.

Conceptual Indexing of Project Diagrams in an Electronic Archive

Alexey Namestnikov
Ulyanovsk State
Technical University
Ulyanovsk, Russian Federation
am.namestnikov@gmail.com

Valeria Avvakumova
FRPC JSC «RPA «Mars»
Ulyanovsk, Russian Federation
valeria.avvakumova73@gmail.com

***Abstract*—The objective of the research, results of which are present in the article, is the development of a method for conceptual indexing of semistructured project diagrams. The task arouses interest in development of intelligent repositories of design organizations and allows grouping the realized projects. The proposed method of conceptual indexing is based on the use of the special type of ontologies – the project ones. In essence, an ontology plays a role of a compatible model for knowledge representation in a project organization. The ontology includes semantic description of project diagram notations and design pattern used in design activity. Conceptual indexing of project diagrams are carried with the following steps: defining the project context, defining the subset of the electronic archive diagrams corresponding to the project context, defining a degree of conformity of design patterns from an ontology to the project diagram concerning each element of the obtained subset of the electronic archive diagrams. In case of conceptual indexing of project diagrams representing composite elements of information resources of the electronic archive, both a text component (comments in program code, different instructions, etc.) and elements of semistructured notations of project diagrams are taken into account. Technically, the result of conceptual indexing is represented in the form of a fuzzy hypergraph defined on sets of terms of a domain ontology and terms corresponding to design patterns. The ontology of project diagrams describing semantics of UML class diagrams and including «delegation» as a design pattern in Java.**

***Keywords*—ontology, conceptual indexing, UML, design pattern, project diagram**

I. INTRODUCTION

In recent years, researchers in the field of program engineering have been interested in intelligent systems that have their origins in ontological principle of special knowledge representation [1], [2], [3]. The ontological representation of artifacts of software development (models, source codes of program modules) allows to carry out automated analysis of program systems and software intensive systems. Realization of the ontological approach assumes constructing the ontology model taking into account aspects of semistructured modelling notations of program systems and features of architectural solutions or design patterns. The domain context can be extracted from comments texts relating to the source code. The factor of fundamental incompleteness of project diagrams and comments in the source code on the natural language implies the necessity of using corresponding mathematical models in the purpose of formalization. The article proposes to use formalism of fuzzy hypergraphs for such cases.

II. THE FORMAL MODEL OF AN ONTOLOGY OF PROJECT DIAGRAMS

In order to solve the problem of intelligent analysis of project diagrams included in project documentation, it's necessary to have knowledge in the field of constructing the formalized diagrams (using notations). Fig. 1 shows the structure of the fragment of the ontology of project diagrams, particularly, class diagrams in UML (Unified Modelling Language).

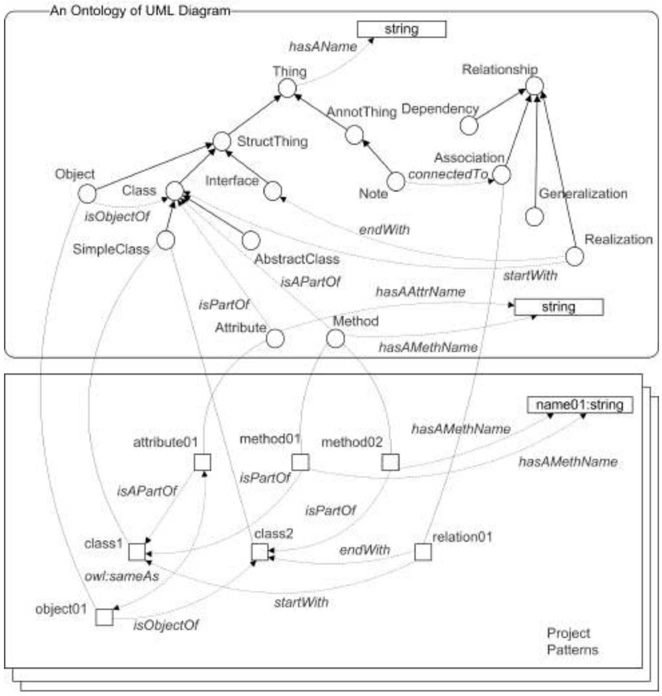


Figure 1. The structure of an ontology of project diagrams (including an example of a design pattern)

Such knowledge allows to identify design patterns used in different projects and, consequently, find projects with similar architectural solutions and approaches to realization of program subsystems of automated systems (AS). Fig. 1

cites a design pattern called «Delegation» as an example [4]. Formally, the ontology of project diagrams can be written as follows:

$$O^{prj} = \{O^{prj}_{dc}, O^{prj}_{tmp_1}, O^{prj}_{tmp_2}, \ldots, R^{prj}, F^{prj}\}, \quad (1)$$

where O^{prj}_{dc} – is an ontology of the UML diagram (a class diagram is used in the working process), $O^{prj}_{tmp_i}$ – is an ontological representation of the i-th design pattern of program systems, R^{prj} – is a relation associating a concept from O^{prj}_{dc} and an instance appertaining to $O^{prj}_{tmp_i}$, F^{prj} – is an interpretation function setting up a correspondance between instances from $O^{prj}_{tmp_i}$ and O^{prj}_{dc} classes.

Let us consider the main components of the UML diagram ontology:

$$O^{prj}_{dc} = \langle C^{prj}, R^{prj}, F^{prj} \rangle,$$

where $C^{prj} = \{c^{prj}_1, \ldots, c^{prj}_l\}$ – is a set of terms defining the main elements of UML class diagram (e.g. «Thing», «Class», «Object», «Interface», «Relationship», etc.); R^{prj} – is a set of relations allowing to construct ontological representations of project diagrams in compliance with the appropriate notations (e.g. relations: defining the name of the diagram element of string type «hasAName», inheritance relation «isA», relation connecting a class with an object of this class «isObjectOf», etc.); F^{prj} – is a set of interpretation functions defining on relations R^{prj}.

The ontological representation of the i-th design pattern of program systems is defined as follows:

$$O^{prj}_{tmp_i} = \{instance(c^{prj}_1), \ldots, instance(r^{prj}_1), \ldots, r_{sameAs}\}.$$

Actually, the ontological representation of an individual design pattern represents a set of terms and relations from the ontology of project diagrams with addition of relation r_{sameAs}, representing the «owl:sameAs» relation embedded in OWL. If the specified relation associates two individuals of an ontological representation of a design pattern, individuals are considered as the same object.

Fig. 1 shows attribute01 and object01 associated by the «owl:sameAs» relation. This means that the specified object of a class (class2) is an attribute of another class (class1).

III. The Method of Conceptual Indexing of Project Diagrams

The objective of the conceptual indexing of project diagrams is calculating an index of electronic archive. In order to solve the task, it's necessary to define instances of ontology classes of project diagrams and calculate the degree of conformity of design patterns to a project diagram for the indexable diagram.

Let us consider the following data as source ones:

- $\{\langle cs_1, dc_1\rangle, \langle cs_2, dc_2\rangle, \ldots, \langle cs_n, dc_n\rangle\}$ – is a set of analyzed projects of an electronic archive, each of them includes the source code cs_i and a class diagram dc_i, i – is the project number;
- the ontology $O^{prj} = \{\langle C^{prj}, R^{prj}\rangle, \{tmp_1, tmp_2, \ldots, tmp_m\}\}$ including a set of concepts of a notation of project diagrams C^{prj} (UML class diagrams as elements), a set of relations between classes R^{prj} and a set of design patterns $\{tmp_1, tmp_2, \ldots, tmp_m\}$;
- $Tz^p = \{\langle t^p_1, f^p_1\rangle, \langle t^p_2, f^p_2\rangle, \ldots, \langle t^p_l, f^p_l\rangle\}$ – is a technical assignment on a new project of an automated system p preprocessed and defined as a number of terms $t^p_1, \ldots, t^p_l$ with corresponding frequencies $f^p_1, \ldots, f^p_l$;
- dc^p – is a project diagram as a part of a new project of the AS (the project diagram corresponds the technical assignment) Tz^p.

According to the scheme of calculating the conceptual index of a design organization, the conceptual indexing of project diagrams in carried as follows:

1) Defining the project context.
2) Defining a subset of the electronic archive diagrams corresponding to the project context.
3) Defining the degree of conformity of design patterns from the ontology O^{prj} to the project diagram dc^p with relation to each element of the obtained subset of diagrams of an electronic archive.

Defining the context of the project is carried out on the basis of the method of conceptual indexing of textual information resources [5], [6]:

$$oV_{tz} = F_{oV}(Tz, O^{dom}, O^{tz}).$$

The input for the function of the conceptual indexing of the textual information F_{oV} is the technical assignment Tz preprocessed, the domain ontology O^{dom} and the thesaurus O^{tz}.

The following set is the result of conceptual indexing

$$oV_{tz} = \{\mu(c^{tz}_1)/c^{tz}_1, \mu(c^{tz}_2)/c^{tz}_2, \ldots, \mu(c^{tz}_k)/c^{tz}_k\} = \mu_{oV}(c^{tz}).$$

The set includes the concept $c^{tz}_i \in C^{dom}$ with the corresponding value of the function of membership of the i-th concept $\mu_i(c^{tz}_i)$ to the technical assignment Tz (the degree of manifestation of the concept in the text of the technical assignment). Let us consider the obtained set oV_{tz} as the realized project concept.

The conceptual indexing is carried analogically for the set of projects of the electronic archive $\{\langle Sc, Dc\rangle\}$. The comment text is extracted for each text of the source code $sc_i \in Sc$:

$$\forall i : tc_i = F_{extcomm}(sc_i),$$

where tc_i is the preprocessed textual representation of the program module sc_i:

$$tc_i = \{\langle t^{sc_i}_1, f^{sc_i}_1\rangle, \langle t^{sc_i}_2, f^{sc_i}_2\rangle, \ldots, \langle t^{sc_i}_s, f^{sc_i}_s\rangle\}.$$

The result of performing a function of the conceptual indexing is the ontological representation of comments of the source code for each program module:

$$oV_{sc_i} = F_{oV}(tc_i, C^{dom}, C^{tz}),$$
$$oV_{sc_i} = \{\mu(c^{sc_i}_1)/c^{sc_i}_1, \mu(c^{sc_i}_2)/c^{sc_i}_2, \ldots, \mu(c^{sc_i}_l)/c^{sc_i}_l\}.$$

Let us define a set of project diagrams of an electronic archive of a project organization compliant with the context oV_{tz} as follows:

$$Dc|_{oV_{tz}} = \{dc_i : \&(\mu_{oV}(c^{sc_i}) \leftrightarrow \mu_{oV}(c^{tz})) \geq 0.5\},$$

where «$\leftrightarrow$» is the equality operator and «&» is the conjunction operator for all $c^{sc_i}, c^{tz} \in C^{dom}$.

In other words, the specified set includes project diagrams for which the condition of fuzzy equality of ontological representations on source codes of programs and a technical assignment is fulfilled.

Let us consider the process of defining the degree of manifestation of patterns of an ontology of project diagrams. That allows to calculate the conceptual index of a project organization according to UML project diagrams presented in the electronic archive.

Let us denote the degree of membership $\mu_{tmp_j}(dc_i)$ of a project diagram dc_i to a pattern tmp_j. $\mu_{tmp_j}(dc_i)$ is defined analytically as follows:

$$\mu_{tmp_j}(dc_i) = \frac{N(ABox^{prj}_{dc_i})}{N(ABox^{prj}_{tmp_j})},$$

where $N(ABox^{prj}_{dc_i})$ is a number of facts that are true if terminology $TBox^{prj}$ is true, and corresponds to the base of facts $ABox^{prj}_{tmp_j}$; $N(ABox^{prj}_{tmp_j})$ is the number of facts of the tmp_j.

The number of facts $N(ABox^{prj}_{tmp_j})$ of a pattern tmp_j is quite simple to define (by summation of the number of facts of the j-th design pattern), but in order to define $N(ABox^{prj}_{dc_i})$, the following algorithm should be used.

Step 1. Transformation of a project diagram of an electronic archive dc_i to a number of facts $ABox^{prj}_{dc_i}$ of the type shown below:

$$elem^{dc_i}_k : Concept$$
$$\langle elem^{dc_i}_k, elem^{dc_i}_s \rangle : Role,$$

where $Concept$ is the concept defined in $TBox^{prj}$ and $Role$ is the role defined in $TBox^{prj}$; $elem^{dc_i}_k, elem^{dc_i}_s$ are terms instances extracted from a project diagram dc_i.

Step 2. Defining the set of basic classes from $ABox^{prj}_{dc_i}$ regarding to a pattern tmp_j.

The basic class is such an instance $elem^{dc_i}_k$ of a term «Class» (or its subsidiary term «Subclass» from $ABox^{prj}_{dc_i}$, that corresponds to an instance $cls^{tmp_j}_l \in Class$ from $ABox^{prj}_{tmp_j}$ and in case of which a pattern tmp_j includes the maximum number of fact of the following type:

$$elem^{dc_i}_k : Concept$$
$$\langle elem^{dc_i}_k, * \rangle : Role, \quad \langle *, elem^{dc_i}_k \rangle : Role.$$

The obtained set of basic classes of a project diagram dc_i regarding to a pattern tmp_j can be denoted as follows:

$$\{\langle elem^{dc_i}_1, cls^{tmp_j}_1 \rangle, \langle elem^{dc_i}_2, cls^{tmp_j}_2 \rangle, \dots\}$$

where a sequence $\langle elem^{dc_i}_k, cls^{tmp_j}_k \rangle$ means that a term instance of a project diagram $elem^{dc_i}_k$ is equal to an instance of a class of a project pattern $cls^{tmp_j}_k$.

Step 3. Calculating the number of true facts by pairwise replacement of instances of classes of the j-th pattern tmp_j and the i-th project diagram dc_i:

$$\forall k : cls^{tmp_j}_k \leftrightarrow elem^{dc_i}_k. \quad (2)$$

The fact of a design pattern is true regarding to a project diagram if the relation to it can be found in the set of facts of a project diagram taking into account the fact that different term names of instances denotes different individuals. Different names of terms belong to the same ontology instance only when the instances are associated through the relation «owl:sameAs».

The stated steps of the algorithm of conceptual indexing of a project diagram dc_i are carried out for each design pattern of an ontology of project diagrams. As a result, ontological representation of a project diagram of an electronic archive is defined as follows:

$$oV_{dc_i} = \{\mu_{tmp_1}(dc_i)/tmp_1, \mu_{tmp_2}(dc_i)/tmp_2, \dots, \mu_{tmp_s}(dc_i)/tmp_s\}. \quad (3)$$

Practically, the expression (3) represents a fuzzy set based on a set of design patterns of an ontology of project diagrams, where $\mu_{tmp_j}(dc_i)$ is a degree of membership of a project diagram dc_i to a design pattern tmp_j.

IV. The model of a conceptual index of project diagrams

The formal structure of a conceptual index of project diagrams is more difficult that the one of textual technical documents. The reason lies in the fact that, in conceptual indexing of project diagrams which are the composite elements of information resources of an electronic archive, both textual component (comments in the source code, different instructions, etc.) and elements of semiformalized notations of a representation language for project diagrams are taken into account.

As in case of document resources, $C = \{c_i\}$, $i \in I = \{1, 2, 3, ..., n\}$ is a finite set of the domain terms fixed in the ontology. A set of patterns of project diagrams in the ontology is denoted as $T = \{tmp_k\}$, $k \in K = \{1, 2, 3, ..., l\}$. A set of project diagrams is denoted as $Dc = \{\tilde{dc}_j\}$, $j \in J = \{1, 2, 3, ..., m\}$ is a family of fuzzy subsets in $C \cup T$. $\widetilde{CI}_{prj} = (C, T, Dc)$ is called a fuzzy undirected hypergraph if $\tilde{dc}_j \neq \varnothing, j \in J$ and $\bigcup_{j \in J} \tilde{dc}_j = C \cup T$; herewith $c_1, c_2, ..., c_n \in C$ and $tmp_1, tmp_2, ..., tmp_l \in T$ are vertices of a hypergraph, a set Dc consisting of $\tilde{dc}_1, \tilde{dc}_2, ..., \tilde{dc}_m$ is a set of fuzzy edges of a hypergraph.

A project diagram has an ontological representation as a result of conceptual indexing. Hence, let us denote a set $\tilde{Dc} = \{\tilde{dc}_j\}$ as a set of project diagrams in the conceptual index, $\tilde{dc}_j$ is an individual ontological representation of the

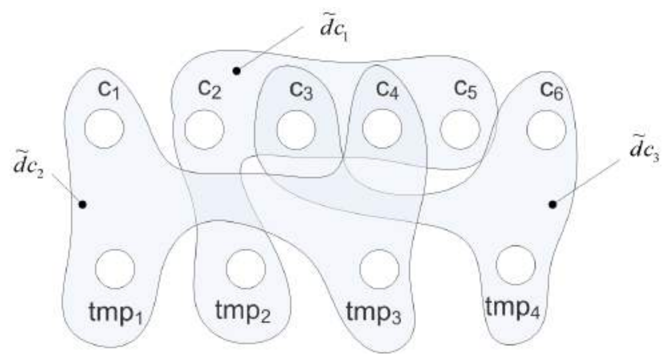


Figure 2. An example of a part of conceptual index

j-th diagram of an electronic archive. The following fuzzy undirected hypergraph

$$\widetilde{CI}_{prj} = (C, T, Dc) \tag{4}$$

defines the conceptual index of a base of projects (Fig. 2) practically.

Two project diagrams $\tilde{dc_\gamma}$ and $\tilde{dc_\delta}$ are called unfuzzy adjacent if $\tilde{dc_\gamma} \cap \tilde{dc_\delta} \neq \varnothing$. The following value

$$\mu(\tilde{dc_\gamma}, \tilde{dc_\delta}) = \bigvee_{c \in (dc_\gamma \cap dc_\delta)} \mu_{dc_\gamma \cap dc_\delta}(c) \tag{5}$$
$$\& \bigvee_{tmp \in (dc_\gamma \cap dc_\delta)} \mu_{dc_\gamma \cap dc_\delta}(tmp)$$

is called a degree of adjacency of project diagrams $\tilde{dc_\gamma}$ and $\tilde{dc_\delta}$. The value $1 - \mu(\tilde{dc_\gamma}, \tilde{dc_\delta})$ describes the distance between the project diagrams in an information base on the basis of contexts of projects and the degree of membership of a project diagram to design patterns from an ontology.

V. Conclusion

The paper presents the formal model of a project diagram ontology allowing to describe diagram notations at the semantic level and design patterns used in the organization. The developed method of conceptual indexing of semiformalized UML project diagrams allows to reduce the task of information resource analysis to operations with hypergraphs. The fuzzy metrics of distance between ontological representations of project diagrams is used of program subsystems for structuring the content of the electronic archive of FRPC JSC «RPA «Mars» (Ulyanovsk, Russia).

Acknowledgment

The authors acknowledge that this paper was supported by the project no. 16-47-730742 and 16-47-732033 of the Russian Foundation for Basic Research.

References

[1] Wongthongtham P., Pakdeetrakulwong U.,Marzooq S. Ontology annotation for software engineering project management in multisite distributed software development environments. Springer International Publishing, Cham, 2017, pp. 315–343.

[2] Emdad A. Use of ontologies in software engineering SEDE (Hisham Al-Mubaid and Rym Zalila-Wenkstern, eds.), ISCA, pp. 145–150 (2008).

[3] Dillon T., Chang E., Wongthongtham P. Ontology-based software engineering- software engineering 2.0. Australian Software Engineering Conference, IEEE Computer Society, pp. 13–23 (2008).

[4] Mark Grand. Java enterprise design patterns: Patterns in java, John Wiley & Sons, 2002.

[5] Guskov G., Namestnikov A. Ontological mapping for conceptual models of software system: Seventh Conference, OSTIS 2017, Minsk, Belarus, February 18-20, 2017, Proceedings, pp. 111-117 (2017).

[6] Namestnikov A., Filippov A., Avvakumova V. An ontology based model of technical documentation fuzzy structuring. CEUR Workshop Proceedings. SCAKD 2016. Moscow. Russian Federation. Volume 1687, pp. 63-74 (2016).

КОНЦЕПТУАЛЬНОЕ ИНДЕКСИРОВАНИЕ ПРОЕКТНЫХ ДИАГРАММ В ЭЛЕКТРОННЫХ АРХИВАХ

Наместников А.М., УлГТУ
Аввакумова В.С., ФНПЦ АО «НПО «Марс»

Целью исследования, результаты которого представлены в данной работе, является разработка метода концептуального индексирования слабоструктурированных проектных диаграмм. Данная задача представляет интерес при разработке интеллектуальных репозиториев проектных организаций и позволяет на семантическом уровне производить группировку реализованных проектов. Предложенный метод концептуального индексирования основан на использовании специального вида онтологий – онтологии проектных диаграмм. Фактически, онтология выполняет роль согласованной модели представления знаний в проектной организации. Данная онтология включает в себя семантическое описание нотаций проектных диаграмм и применяемых в проектной деятельности шаблонов проектирования.

Концептуальное индексирование проектных диаграмм выполняется посредством следующих шагов: определение контекста проекта, определение подмножества диаграмм электронного архива, соответствующих контексту проекта, определение степени соответствия шаблонов проектирования из онтологии проектной диаграмме относительно каждого элемента найденного подмножества диаграмм электронного архива. При концептуальном индексировании проектных диаграмм, являющихся составными элементами информационных ресурсов электронного архива, учитывается как текстовая составляющая (комментарии в программном коде, различные инструкции и т.д.), так и элементы слабоформализованных нотаций представления проектных диаграмм. Формально результат концептуального индексирования представляется в виде нечеткого гиперграфа, определенного на множествах понятий онтологии предметной области и понятий, соответствующих шаблонам проектирования.

Разработана онтология проектных диаграмм, описывающей семантику диаграмм классов языка UML и включающая представление шаблона проектирования на языке Java «делегирование».

Using ontologies for data processing scenarios in aircraft designing

Javier Flores
Department of Aircraft Design
Samara University
Samara, Russia
flores.javier@yandex.ru

Kamil Osseni
Department of Aircraft Design
Samara University
Samara, Russia
ossenikamil@gmail.com

Nikolai Borgest
Department of Aircraft Design
Samara University
Samara, Russia
borgest@yandex.ru

***Abstract*—The use of Ontologies in the present is becoming very common and widely used in different education areas, some areas of the engineering such as aerospace, can improve with the application of this informatics tools. Classification of aircraft as an object and all its components has become an important research area for the statistics analysis previous of designing a new concept, in this document we are going to explain, how Fluent Editor can help the designer to improve the classification and information analysis used to choose the statistical characteristics of the new concept. On this paper databases are created for storing, processing, performing calculations, sorting, sampling and presentation of data arrays according to various criteria. Then the created array of data on aircraft wing airfoil can be further used as a basis for selecting the airfoil according to the technical task.**

***Keywords*—design, ontology, education, aircraft, airfoil**

I. Introduction

It is known that the key element of the process of creating an airplane is its project, that is, the development of the project both in manual mode and in automated mode. To develop a modern airplane project means to develop a complete set of design and technological documentation both for paper technology and for machine carriers, which makes it possible to create an aircraft in metal and to operate it.

During the design of a new aircraft, designers and analysts create and refine several aircraft models using different software tools. Each model covers parts of the whole aircraft and usually focuses on one aspect of the aircraft. Especially during conceptual aircraft design the degree of diversity and content overlap are high compared to later design phases.

It is clear that modern design of aircraft and other complex engineering techniques cannot be created without the use of automation systems, so the degree of automation of design processes is largely determined not so much by the capabilities of modern CAD, as by the ability to formalize a particular design task, that is, the designer's ability to give a fairly strict formulation design tasks and a clear completed algorithm for solving it, using the maximum information about a typical design task, then is the use of information about the projected product (airplane).

Classically, the technical documentation is only the end result of a complex and long process of design work by the creators of the aircraft, aimed at developing a project of a previously non-existent object (aircraft), system and process.

Classification of digitized documents nowadays gains a higher significance due to the rapid growth of digital content. With respect to the growth, organizing them is a big challenge for efficient retrieval information. Therefore, finding and improving solutions for text classification has considerable importance [1].

II. About airfoil design concepts in general

An airfoil is a surface designed to obtain lift from the air through which it moves. Thus, it can be stated that any part of the aircraft that converts air resistance into lift is an airfoil. The profile of a conventional wing is an excellent example of an airfoil. (Fig. 1) Notice that the top surface of the wing profile has greater curvature than the lower surface.

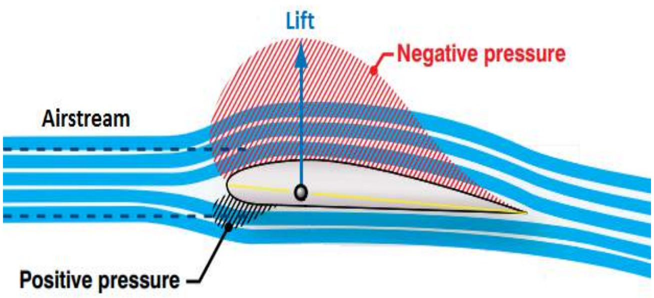


Figure 1. Airfoil flow.

The difference in curvature of the upper and lower surfaces of the wing builds up the lift force. Air flowing over the top surface of the wing must reach the trailing edge of the wing in the same amount of time as the air flowing under the wing. To do this, the air passing over the top surface moves at a greater velocity than the air passing below the wing because of the greater distance it must travel along the top surface.

This increased velocity, according to Bernoulli's Principle, means a corresponding decrease in pressure on the surface. Thus, a pressure differential is created between the upper and lower surfaces of the wing, forcing the wing upward in the direction of the lower pressure.

An aircraft in flight is the center of a continuous battle of forces. Actually, this conflict is not as violent as it sounds, but it is the key to all maneuvers performed in the air. There is nothing mysterious about these forces; they are definite and known. The directions in which they act can be calculated, and the aircraft itself is designed to take advantage of each of them. In all types of flying, flight calculations are based on the magnitude and direction of four forces: weight, lift, drag, and thrust (Fig. 2).

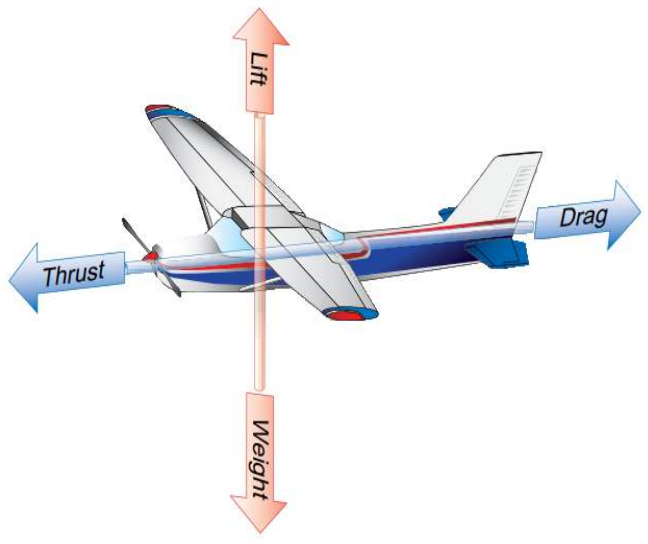


Figure 2. Magnitude and direction forces.

An aircraft in flight is acted upon by four forces:

1) Gravity or weight—the force that pulls the aircraft toward the earth. Weight is the force of gravity acting downward upon everything that goes into the aircraft, such as the aircraft itself, crew, fuel, and cargo.
2) Lift—the force that pushes the aircraft upward. Lift acts vertically and counteracts the effects of weight.
3) Thrust—the force that moves the aircraft forward. Thrust is the forward force produced by the powerplant that overcomes the force of drag.
4) Drag—the force that exerts a braking action to hold the aircraft back. Drag is a backward deterrent force and is caused by the disruption of the airflow by the wings, fuselage, and protruding objects.

III. Ontology basics

Ontology originates from Greek philosophy, namely the study of being and existence, dealing with the questions what kinds of things exist and how they relate to one another. This concept has been adapted for use in computer science. Studer [3] define an ontology as "a formal explicit specification of a shared conceptualization of a domain of interest", emphasizing formality which is needed for automated processing, a consensus about the contents, and the focus on a specific domain whereas the view on that domain is influenced by a certain interest for the ontology in mind.

A. *Types of Ontologies*

Depending on their purpose, ontologies can be categorized into the following types [4]:

- Top-level ontologies cover general and abstract concepts, e.g notions of time and space that can be reused and refined in other ontologies.
- Domain or task ontologies cover knowledge about a specific domain (e.g. aircraft) or a specific task (e.g. cooking); since this distinction is somewhat imprecise, both are normally referred to as domain ontology.
- Application ontologies are typically developed in complement to an application and with certain usage scenarios in mind. They cover and refine specific aspects of domain ontologies for use in that specific application.
- The ontology developed in the context of this paper can be categorized as domain ontology.

B. *The Ontology Language OWL*

As a language for describing ontologies, the World Wide Web Consortium W3C1 recommends the Web Ontology Language (OWL) [5], respectively its revision OWL 2 [6]. The foundations for defining semantics between concepts in OWL are logical declarations which can be evaluated by reasoners. Reasoners are programs which provide services such as checking the consistency of logical declarations in an ontology and inference of new knowledge from explicitly declared knowledge. In general, ontologies consist of concepts and roles. The concepts are organized in a hierarchical structure formed by is-a relations between these concepts. In OWL, these concepts are called classes, e.g. an Airbus A340-500 is a specialized sub-class of its superclass Airplane.

With the use of roles, more context can be added to classes in form of semantic relations. OWL expresses roles by properties which represent relations between two concepts. Possible sources and targets of these relations can be defined by the specification of appropriate domains and ranges of properties. OWL stipulates two kinds of properties: object properties relate two classes, whereas data type properties relate classes to data types.

C. *The Open World Assumption*

A peculiarity of ontologies is the open world assumption. Essentially, it states that all knowledge, that is not explicitly or implicitly stated in an ontology, has to be regarded as unknown. In contrast, according to the closed world assumption, which is normally used with traditional data models such as database systems, missing knowledge would be regarded as non-existent.

Take for example a knowledge base that consists of the single statement "Airbus is an aircraft manufacturer", and the question "Is Boeing an aircraft manufacturer?". According to the closed world assumption, the answer to the question would be ”No”, whereas the open world assumption would result in ”Unknown”. According to the open world assumption, no conclusions are made until more knowledge is available which might result in an unambiguous answer. Practically, it means that in principle it always possible to add new logically consistent knowledge to an ontology without invalidating its conceptualization or content, whereas with traditional data models in a closed world this might possibly require a complete overhaul of the model or its content.

IV. Ontology development

In the following, we provide an introduction into the ”Solution of design tasks with the help of ontological systems” methodology, show its application in our project, and conclude with the evaluation of the developed ontology.

Divide the ontology into 3 parts:

- Database of airplane (Table 1, describe the five aircraft and assign them the values of Vc)
- Database of the wing airfoil (Table 2, describe the 6 wing profiles and assign them the profile thickness values)
- The selection of the thickness and the wing airfoil (Table 3, set the conditions for selecting the wing profile for the 5 aircraft).

For symmetric profiles are selected empennage, the thickness of which is determined by the Mach number for the cruising flight regime. The most important factor influencing the choice of the wing profile is also the Mach number. Through the condition given in Table 3, the relative thickness of the profile is chosen. Based on the relative thickness of the profile, a set of existing wing profiles for an aircraft with a given Mach number is selected [6].

Table I
Five selected aircraft

Aircraft	Vc, km/h
A 340-500	890
Boeing 717	810
Yak-42	700
Bombardier Jet 100	786
An-24	460

Table II
Six selected airfoils

Airfoil	c,%
A-15	15
B-14	14
TsAGI 6-13	13
A-12	12
Clark-YH-11	11
N-10	10

A. Creating parts of the ontology

Hierarchical levels are related to each other, this characterizes the structure of the system and regulates the composition of its elements, blocks, aggregates and the relationship of the constituent structures to each other. At the same time, any structure is created to perform certain functions (useful actions, states or properties). For example, the wing performs the functions of ”lift the airplane”, ”maneuver the airplane”, etc.

Table III
Recommended values thicknesses bearing surfaces

Range of numbers M	c-relative thickness, %		
	wing	*horizontal stabilizer*	*Vertical stabilizer*
M>0,7	15 - 13	12 - 6	12 - 6
M<0,7	12 - 10	8 - 6	8 - 6

Thus, all the elements of the subsystems and their individual elements over the hierarchy levels are related to each other by functional relationships.

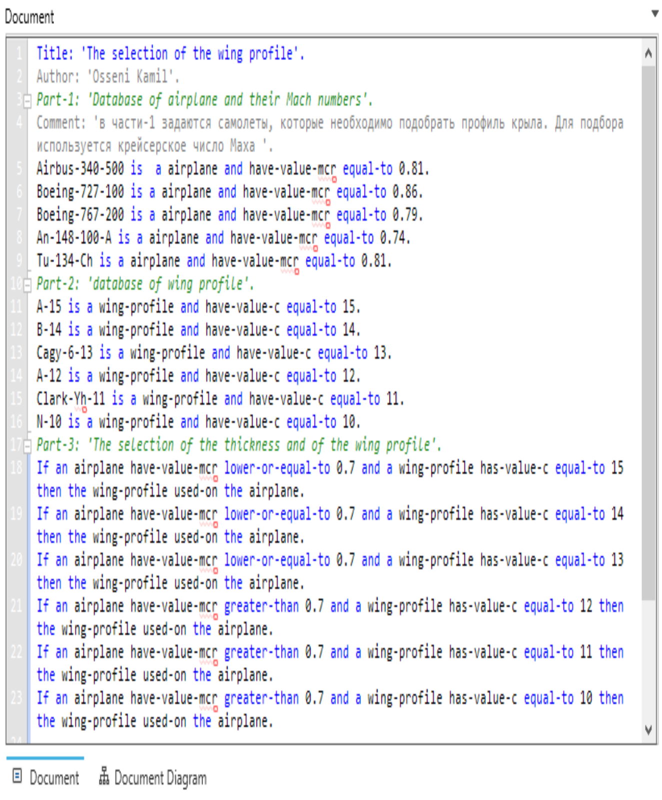


Figure 3. Fragment of the document.

Each hierarchical level corresponds to its list of tasks, the solution of which is necessary for making design decisions corresponding to this level, and even more so in the conditions of the CAD functioning when creating the aircraft design. Therefore, in the automated design of aircraft design, an important aspect from the point of view of formalization is its hierarchical structure and the multi-stage design that follows from it [7].

Figure 3 shows CNL Editor window is the main part of Fluent Editor where you can actually view and edit ontology files. It shows all the CNL phrases, both from the edited file and from the reference ontologies that correspond to every OWL statement. You can click on any phrase from the edited OWL file to modify it, or can also add new phrases [7].

B. Taxonomy Tree

Fluent Editor shows in a way of Taxonomy Tree in for of hierarchical levels and how each of the elements are related to each other, this structure facilitates the visualization of the Database. Taxonomy tree is displayed for each OWL file being edited and is built upon data from this OWL file and all included ontologies. Selecting element on the Taxonomy tree will filter all expressions in CNL Editor window to those, which are related explicitly to the selected phrase [7].

The figure 4 shows the Taxonomy tree that helps to navigate and to search systems or subsystems in the document.

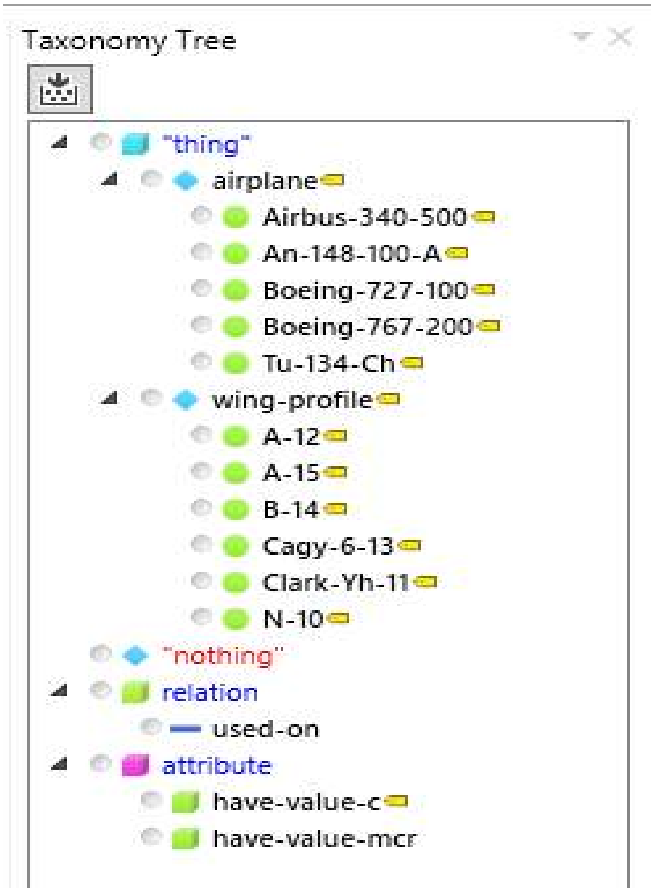


Figure 4. Fragment of the Taxonomy Tree.

C. Presentation of ontology as a diagram

Checking the completeness and correctness of the established Database is carried out by visualization of ontology, namely, to create relationships and instances represented graphically by CNL-diagram (Figure 5). One skilled in the subject area will be easier to assess the correctness of the prepared guidance by imaging in CNL-diagram terms and relations between them [6].

The CNL-diagram help us to visualize in a more accessible way all the information, so in this way is more simple to understand the relation of every object, for example, (Figure 5) this diagram shows the 5 different airplanes,and also the 6 airfoils, and how they can relate as a thing.

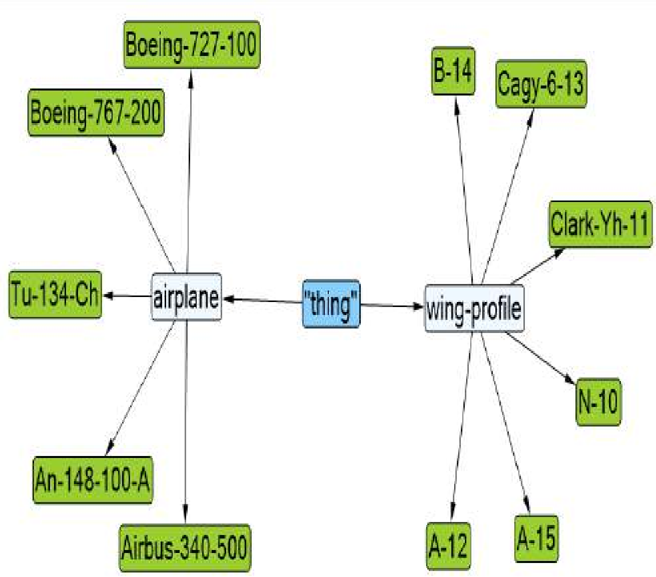


Figure 5. CNL-diagram.

D. Reasoning

Regarding reasoning, the two most relevant benefits are consistency checks and automated classification. By using the restrictive capabilities provided by OWL, we were able to enforce rather rigid consistency constraints.

Trying to assign a value-c to a wing-profile, the subsumption service of reasoners can classify an profile that is used-on the Airplane. we use the logical selection of the thickness and of the wing profile.

- if an airplane have value-mcr lower or equal to n and a wing-profile has-value-c equal-to n then the wing profile used-on the airplane
- if an airplane have value-mcr greater-than n and a wing-profile has-value-c equal-to n then the wing profile used-on the airplane

Given an individual of an profile that was accidentally created as a direct member of the most general class Thing. When this individual is used-on the Airplane individual, the reasoner can classify it as profile.

SWRL is a powerful mechanism to build actual ontologies and express complex conditional relations for instances. Yet it is quite intractable to trace, especially for larger ontologies. In a brief it allows to inspect which rules were executed during materialization and what entities was substituted for their head clauses (and used in the body)

'Active Rules' In a brief it allows to deploy within the knowledge-base imperative code that will be executed when a certain conditions are met, expressed with the SWRL, the body of the Active Rule may perform any arbitrary action or update the knowledge base (Fig. 6).

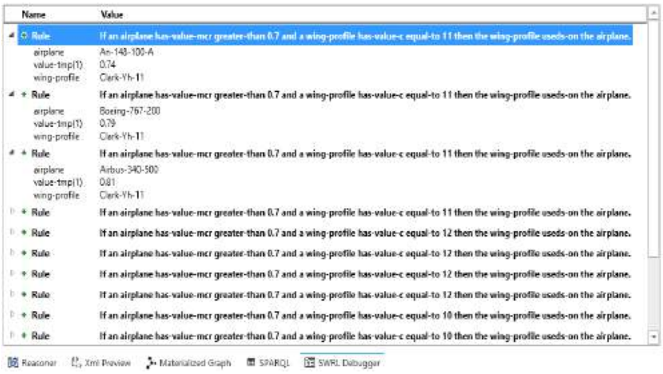


Figure 6. SWRL Debugger classification

After the running of the SWLR debugger the new CNL diagram (Fig.7) shows the relations of the wing profile and the airplane, and what type of profile is the best to use in each aircraft, acording to the relations of c and mcr values of all the items in the ontology, so we can see that there are some wing profiles that dont fit with an specific airplane and other fit with two or more, an then in an easyest visualy form this will help to chose the best of the wing profiles for the airplane.

V. Related work

There has been extensive work on the benefits of applying semantic technologies for the efficiency of model driven systems engineering, which has been the motivational background for this ontology development project.

For example, in the development of the aircraft ontology following the NEON process model. In particular, is described the experiences from applying the NEON methodology and the resulting aircraft ontology. The aircraft ontology is an OWL ontology that covers system decomposition and component parameters of a single aisle civil transport aircraft. It can be used as a common semantic reference during model comparison and transformation [10].

Representing design alternatives as configurations of port-based objects is useful at the conceptual design stage when the geometry and spatial layout is still ill-defined. During the design process, as the designer makes additional decisions about the components and their interactions, these initial placeholders will be gradually transformed into specific port definitions. In terms of the port ontology, the incremental decisions of the designer will result in the addition of attributes to the port definitions, the sub-classing or refinement of attributes, or the addition of constraints on the attribute values [11].

Another application that would work well together with this type of ontology is the Robot Designer actually in development in Samara University, this investigated the possibility of applying methods and techniques of artificial intelligence for accelerated training of student for preliminary design of the aircraft. It is aimed to help in the process of preliminary design of an aircraft [12].

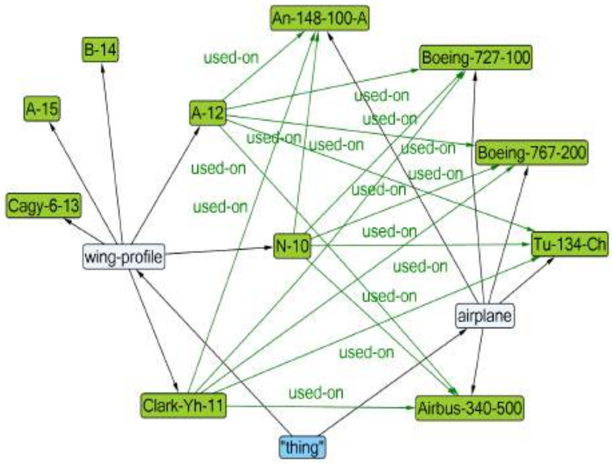


Figure 7. new CNL diagram

VI. Conclusion

Classification of information on an easy way, effectively and fast, then working with it, will help in the future to cut times of design and research, so this will have a positive effect on the cost of development of new technology.

"Fluent Editor" has the necessary mechanisms for the implementation, conceptualization and formalization of ontology. This program facilitates the use of ontologies for classification of objects and to visualize them so the student or researcher, could do an easiest and faster conclusions on themes related to the design area.

We also intend to exploit the aircraft ontology in existing aircraft design tools beyond model consistency checking and integration. Our long term goal is to contribute to the integration of semantic technologies into system design tools and to the establishment of knowledge engineering as a natural part of systems design.

Future research is supposed to explore a more theoretical way of design concept definition than the presented one. It is empirical in that it followed a speculative way to find out the content and contextual dependencies of a design ontology and design concepts. It is also for future research how to make concept generation unique and unambiguous.

It is an absolutely open issue is how to create a population of materializable design concepts quasi automatically, if possible at all. As well, it needs further research how to further articulate possible interdependencies and conceivable interactions of design concepts, especially in unknown circumstances (situations).

References

[1] Ch.M. Wijewickrema. Impact of an Ontology for Automatic Text Classifications, Annals of Library and Information Studies. Vol. 61, December 2014, pp. 263-272. - http://nopr.niscair.res.in/handle/123456789/30334.

[2] FAA Aerodynamics, Aircraft Assembly, and Rigging

[3] Studer, R., Benjamins, V., and Fensel, D. (1998). "Knowledge engineering: principles and methods". In: IEEE Trans. Knowl. Data Eng. 25 (1-2), pp. 161–197

[4] Studer, R., Grimm, S., and Abecker, A. (2007). Semantic Web Services: Concepts, Technologies, and Applications. Springer-Verlag New York, Inc..

[5] Zhitomirsky G.I. The design of aircraft: A textbook for students of aviation specialties of universities. / 3rd ed., And additional. / M .: Mechanical Engineering, 2005. 406 p.: Ill.

[6] Great Soviet Encyclopedia: [in 30 t.] / Ch. Ed. A. M. Prokhorov. / 3rd ed. / M.: Soviet Encyclopedia, 1969-1978.

[7] A.A. Orlova, N.M. Borgest. Text verification instruction manual using the ontologies / Proceedings of the 2016 Conference on Information Technologies in Science, Management, Social Sphere and Medicine (ITSMSSM 2016). ACSR - Advances in Comptuer Science Research. Vol. 51. 2016, pp.238-241.

[8] N.M. Borgest, M.D. Korovin. Formalization of design procedures in the engineer's educational strategy / Proceedings of the 2016 Conference on Information Technologies in Science, Management, Social Sphere and Medicine (ITSMSSM 2016). ACSR - Advances in Comptuer Science Research. Vol. 51. 2016, pp.524-527.

[9] OWL - https://www.w3.org/TR/owl-semantics/

[10] M. Ast, Martin Glas, Tobias Roehm, Creating an Ontology for Aircraft Design, Deutscher Luft- und Raumfahrtkongress 2013. DocumentID: 301356. – 11 p.

[11] Vei-Chung Liang, Christiaan J.J. Paredis. A port ontology for conceptual design of systems. Journal of Computing and Information science in engineering. Vol.4, September, 2004, pp.206-217. - DOI: 10.1115/1.17781

[12] N.M. Borgest, S.A. Vlasov, Al.A. Gromov, An.A. Gromov, M.D. Korovin, D.V. Shustova, Robot-designer: on the road to reality. Ontology of Designing, 5(4): 429-449, 2015. DOI: 10.18287/2223-9537-2015-5-4-429-449.

ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ ДЛЯ СЦЕНАРИЕВ ОБРАБОТКИ ДАННЫХ при ПРОЕКТИРОВАНИИ САМОЛЕТОВ

Хавиер Флорес, Оссени Камиль, Николай Боргест

Самарский Университет

В настоящее время использование онтологий становится очень распространенным. Они широко используются в различных областях образования, некоторых инжиниринговых областях, таких как аэрокосмическая промышленность, и могут улучшаться с применением этих инструментов информатики. Классификация самолетов как объекта и его компонентов стала важной областью исследований для анализа статистики, предшествующей разработке новой концепции. В этой статье делается попытка объяснить, как Fluent Editor может помочь дизайнеру улучшить классификацию и анализ информации, используемые для выбора характеристик новой концепции самолета. В работе были созданы базы данных для хранения, обработки, выполнения вычислений, сортировки, выборки и представления массивов данных в соответствии с различными критериями. Созданный массив данных по аэродинамическим профилям летательного аппарата может быть далее использован в качестве основы для выбора аэродинамического профиля в соответствии с требованиями технического задания.

Informational support of the process of mastering the ontology design tool

Anastasiya Malochkina
Aircraft design department
Samara University
Samara, Russia
malochkina.anastasia@gmail.com

Nikolai Borgest
Aircraft design department
Samara University,
ICCS RAS
Samara, Russia
borgest@yandex.ru

***Abstract*—This paper gives information about implementation of information support both for students and teachers of the discipline "Ontology of production sphere". The goal is to introduce methods of ontology creation into the process of airplane design to teach student find best solutions of design tasks. Vital part of information support is an effective communication between teacher and student. Usage of information support helps to reduce time for performing routine tasks. Much attention is given to content of this course. That provides a maximum level of adoption in electronic view.**

Methods and tools of ontology creating are described, recommendations for information support content are given. Developed data sourse is proposed to use as information support in Samara University.

***Keywords*—ontology, information support, education, aircraft**

I. Introduction

At the first stage of the aircraft design, an analysis is made of the already existing parameters and characteristics of the aircraft. Analysis of the experience of predecessors helps determine the characteristics of the new aircraft. This reduces the time and money spent on design and production.

An ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an ontology is a systematic account of Existence. For knowledge-based systems, what “exists” is exactly that which can be represented [1].

As the toolkit is developed to create ontologies, students can create their own ontologies for solving complex problems. This is made possible by the fact that the discipline "Ontology of the production sphere" provides key knowledge in the field of artificial intelligence, data representation models and the opportunities provided by modern ontology editors. Thus, information support is used as a monitoring tool for the teacher and a source of information for the student.

II. Description of the subject area

A. Course Content

Ontologies can be used in any branch of science. They can be used to describe the subject area. Ontologies can also contain a thesaurus, necessary for studying the conceptual apparatus of the domain. The purpose of creating a thesaurus is to create a semantically integrated basis for creating an ontology. Its implementation allows, at least reduce, and sometimes avoid problems associated with the ambiguity of the interpretation of one term in different spheres [2]. Reasons for creating ontologies:

- For shared usage by people and software agents each;
- Improving the perception of information through its formalization;
- Identification of the assumptions in the domain;
- Review of accumulated knowledge.

B. Review of the article and review of other ontological editors

Students in the course are given a list of philosophical works, from which they must choose one article and to work with her. The work consists in:

- Forming a student’s personal view of the article;
- Summing up and conclusions about the read;
- An explanation of how this article relates to the understanding of ontology by a student.

A review of other ontology creation products gives learner a more complete understanding of the possibilities for construction of ontologies.

C. Practice

Practice in this course is carried out with the help of the editor of ontologies "Magenta" [3]. Objectives of practice:

- Studying fundamental concepts that show how multi-agent systems help in designing;
- Studying the principles of building multi-agent systems for representing the functioning of a complex system;
- Mastering ontology tools;
- Assimilate approaches to design;
- Developing technologies to solve more complex problems.

Figure 1 shows the results of one-way matching. The ontology in this case contains the main characteristics of the prototypes and the requirements for the new aircraft. Figure 1 shows how they will relate to each other.

The possibilities of two-way matching on the example of solving the logistics problem are presented in Figure 2. The aircraft have a certain capacity, departure time and departure point. Each cargo has its mass and its delivery time. According to these data, cargo is distributed to aircrafts.

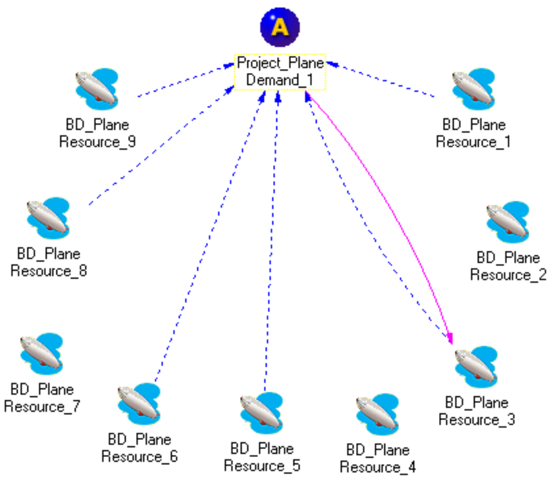


Figure 1. One-way matching.

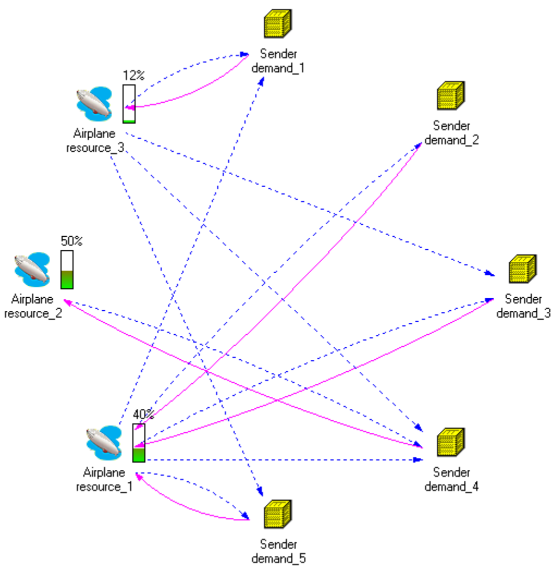


Figure 2. Two-way matching.

A piece of the practice is performed in the SmartSolutions software. It allows you to describe knowledge about the subject area, build conceptual models of enterprise activity, and also create situations models (scenes) used for situational management in intelligent decision support systems of a new class based on multi-agent technologies based on the principles of self-organization and evolution.

III. Formalization of the subject area

A. The formalization model

The aim of the training is to assimilate the knowledge accumulated in the studied subject area and the skills necessary for solving problems in this area. Information support is a "database" in which accumulated data is kept. To represent them in a form understandable to students, the scheme shown in Figure 3 is used.

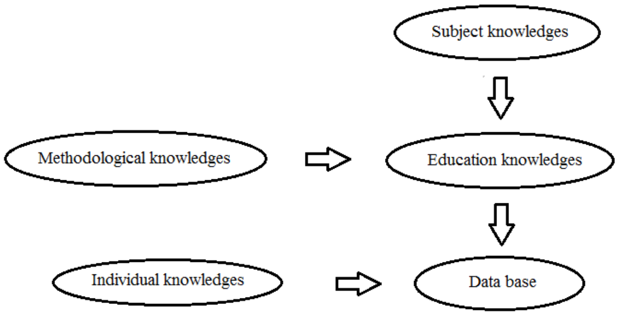


Figure 3. Composition data base scheme of information system.

Subject knowledge is an expert's knowledge about content and structure of teaching material. It is developed by teacher in accordance with state standard and his or her practical knowledge. Education knowledge reflects regularities of a certain subject educational process and content subject and methodological knowledge. Methodological knowledge is knowledge about adopting of subject knowledge. Individual knowledge belong to student, this term used to name experts' representation about students [4].

B. Subject area formalization

To visualize the content of the course "ontology of the production sphere" with the help of the ontology editor Fluent Editor [5],which uses "controlled English" to create ontologies, the course content is formalized. Visual representation helped to divide information into topics and create course scenario. Diagram created in Fluent Editor for the discipline "Ontology of production sphere" is presented in Figure 4.

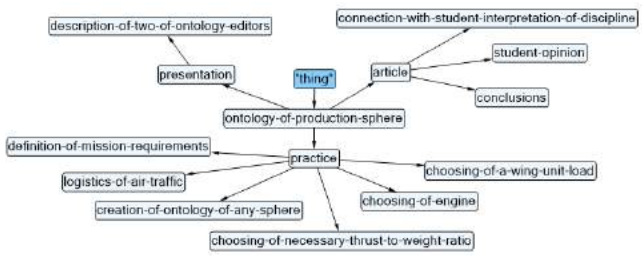


Figure 4. Diagram for the discipline "Ontology of production sphere".

IV. Approbation

A. Description of the tool

To design information support, a distance learning system Moodle was chosen. Its advantage is a free distributing and an open system code, which gives great opportunities for creating a course [6]. The training course consists of tests, databases, forums with frequently asked questions and glossaries with the possibility of adding new terms. Grades are set in semi-automatic mode (automatically based on the test results and manually when checking written assignments). A separate rating scale is defined for each kind of tasks. When creating the system, according to Figure 3, various types of knowledge were introduced into it. Let us consider in more detail:

- Subject knowledge - tasks descriptions and comments for each section.
- Educational knowledge - the experience of teaching the course is reflected in dividing the course into sections located in a certain order.
- Methodical knowledge - the first lecture in simple language and examples from life gives complex definitions and forms an approximate idea of the subject area.
- Individual knowledge - the answers to questions in the test have a variation on the correctness, which can be confusing if the understanding of the material is insufficient.

In addition, introductory testing has been introduced to help identify gaps in previous training, which are impeding successful completion of the course.

Statistics of student errors can be used to further optimize the course. Such information support stimulates students to self-preparation. The information is divided into parts in a certain way, each of them has its own time frame to stimulate preparing of everything on time. The end dates for studying the segments are indicated in the "calendar" block. Notifications about the end of the course are included by default. Figure 5 shows the division of the course into sections.

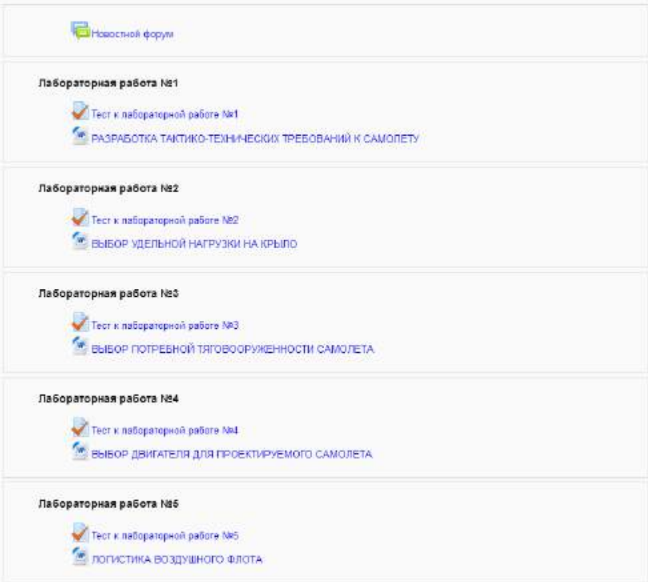


Figure 5. Course interface.

B. Newness

The use of information technology defines a new approach to the learning process. For the design of this course, ontology creation technologies were used [7]. This helped to highlight the main characteristics of this discipline. Students also have access to ontology. The capabilities of the Fluent Editor not only helped visualize the data, they allow you to ask ontology questions about its content. For this, you need to know "controlled English".

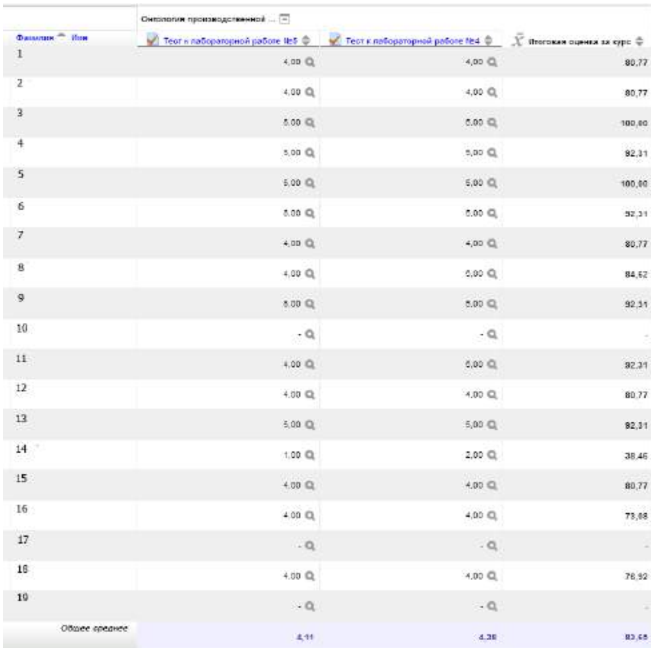
Ontology has become a tool for solving some problems that arose at the design stage of information support of the educational process:

- Lack of an acceptable classification of knowledge of the subject area and ambiguity of terms;
- The invariance of the presentation models and the review of the expert's ambiguous knowledge of the learning process.

C. Results

At first stage elements of control of knowledge were implemented. Within the course each element is defined with its own scale of assessment, but the success in mastering the course is estimated on a 100-point scale. Also, it is possible to see success level of every test in percent form. Visual results are given in figure 6. Moodle gives statistic data such a number of attempts, percent of right answers, marks median, etc.

Implementation of information support in an education process gives opportunity to analyze statistical data to reveal poor questions and replace it [8]. All the statistic can be downloaded as an Excel file. Next advantage is a control of an individual knowledge. System presents person's right and wrong answers for certain test. It put into practice individual training. In addition, teacher is able to see how much time student have spent on test fulfillment. Cheating become obvious. Review of students' right and wrong answers is presented in figure 7.



	Онтологии производственной ...		
Фамилия / Имя	Тест к лабораторной работе №5	Тест к лабораторной работе №4	Итоговая оценка за курс
1	4,00	4,00	80,77
2	4,00	4,00	80,77
3	5,00	5,00	100,00
4	5,00	5,00	92,31
5	5,00	5,00	100,00
6	5,00	5,00	92,31
7	4,00	4,00	80,77
8	4,00	5,00	84,62
9	5,00	5,00	92,31
10	-	-	-
11	4,00	5,00	92,31
12	4,00	4,00	80,77
13	5,00	5,00	92,31
14	1,00	2,00	38,46
15	4,00	4,00	80,77
16	4,00	4,00	73,08
17	-	-	-
18	4,00	4,00	76,92
19	-	-	-
Общее среднее	4,11	4,28	83,65

Figure 6. Representation of marks.

Адрес электронной почты	Состояние	Тест начат	Завершено	Затраченное время	Оценка/5,00	В. 1 /1,00	В. 2 /1,00	В. 3 /1,00	В. 4 /1,00	В. 5 /1,00
z	Завершенные	9 Ноябрь 2017 12:37	9 Ноябрь 2017 12:37	18 сек.	1,00	✗ 0,00	✗ 0,00	✓ 1,00	✗ 0,00	✗ 0,00
a	Завершенные	21 Декабрь 2017 08:56	21 Декабрь 2017 08:57	1 мин. 55 сек.	3,00	✓ 1,00	✓ 1,00	✓ 1,00	✗ 0,00	✗ 0,00
s	Завершенные	21 Декабрь 2017 09:08	21 Декабрь 2017 09:08	24 сек.	4,00	✓ 1,00	✓ 1,00	✓ 1,00	✗ 0,00	✓ 1,00
p	Завершенные	25 Декабрь 2017 08:38	25 Декабрь 2017 08:39	53 сек.	5,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00
c	Завершенные	25 Декабрь 2017 08:39	25 Декабрь 2017 08:39	51 сек.	5,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00
n	Завершенные	25 Декабрь 2017 08:48	25 Декабрь 2017 08:50	1 мин. 53 сек.	4,00	✓ 1,00	✓ 1,00	✓ 1,00	✗ 0,00	✓ 1,00
e	Завершенные	25 Декабрь 2017 08:49	25 Декабрь 2017 08:51	1 мин. 32 сек.	5,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00
q	Завершенные	25 Декабрь 2017 08:54	25 Декабрь 2017 08:55	1 мин. 25 сек.	4,00	✓ 1,00	✓ 1,00	✓ 1,00	✗ 0,00	✓ 1,00
r	Завершенные	25 Декабрь 2017 08:56	25 Декабрь 2017 08:56	1 мин. 3 сек.	1,00	✗ 0,00	✓ 1,00	✗ 0,00	✗ 0,00	✗ 0,00
r	Завершенные	25 Декабрь 2017 08:56	25 Декабрь 2017 08:57	42 сек.	4,00	✓ 1,00	✓ 1,00	✓ 1,00	✗ 0,00	✓ 1,00
b	Завершенные	25 Декабрь 2017	25 Декабрь 2017 09:01	1 мин. 56 сек.	5,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00	✓ 1,00

Figure 7. Representation of marks.

V. Conclusion

At present, artificial intelligence is widely available, capable of learning and adapting in computer systems [10]. The software for e-learning helps to find new solutions to improve the learning process. E-learning is continuously developing, which allows you to get knowledge at any convenient time of day and in any convenient place. Adaptation of the studied material in this work is done using ontology technology. They provide a lot of opportunities in any branch of science. The information support created for the course "ontology of the production sphere" in Samara University is just one example of their application. Successful mastering of this discipline helps in the study of other disciplines. Upon completion of the course, students can apply the acquired knowledge in practice: formalize the design (development) process, think through descriptions of ontologies and create their own ontologies.

References

[1] Thomas R. Gruber, A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, 1993, vol. 5 (2), pp. 199-220.

[2] N.M. Borgest and M.D. Korovin, Ontologii: sovremennoe sostoyanie, kratkii obzor [ONTOLOGIES: CURRENT STATE, SHORT REVIEW]. *Ontologiya proektirovaniya [Ontology of Designing]*,2013, vol. 2 (8), pp. 49–55.

[3] MagentA. Multi-Agent Applications. Version 2/0 build 2.23. 2002. Available at: http://www.magenta-corp.com/ (accessed 2017, Dec)

[4] I.Yu. Denisova and P.P. Makarychev, Matematicheskie modeli ontologii bazy znanii informatsionnoi obuchayushchei sistem [MATHEMATICAL MODELS OF ONTOLOGY KNOWLEDGE BASE OF INFORMATION TRAINING SYSTEM]. *Ontologiya proektirovaniya [Ontology of Designing]*, 2012, vol. 3 (5), pp. 62-78.

[5] FluentEditor 2014. Available at: http://www.cognitum.eu/semantics /Fluent Editor/ (accessed 2017, Dec)

[6] Moodle. Available at: https://moodle.org/ (accessed 2017, Dec)

[7] I.Yu. Denisova and P.P. Makarychev, Ontologicheskoe issledovanie protsessa elektronnogo obucheniya i proektirovanie sredstv podderzhki [ONTOLOGICAL RESEARCH OF E-LEARNING PROCESS AND DESIGN SUPPORTS]. *Ontologiya proektirovaniya [Ontology of Designing]*, 2012, vol. 4 (6), pp. 61-72.

[8] S.A. Nesterov and M.V. Smetanina, Otsenka kachestva testovykh zadanii sredstvami sredy distantsionnogo obucheniya MOODLE [QUALITY MEASUREMENT OF THE TEST TASKS IN THE LEARNING MANAGEMENT SYSTEM MOODLE]. *Nauchno-tekhnicheskie vedomosti SPbGPU* , 2013, vol. 5 (181), pp. 87-92.

[9] K. Baclawski, M. Bennett, G. Berg-Cross, D. Fritzsche, T. Schneider, R. Sharma, Ram D. Sriram and A. Westerninen, ONTOLOGY SUMMIT 2017 COMMUNIQU'E - AI, LEARNING, REASONING AND ONTOLOGIES. Available at: https://s3.amazonaws.com/ontologforum/OntologySummit2017/ Communique/OntologySummit2017Communique_v13.pdf (accessed 2017, Dec)

[10] A. Malochkina, N. Borgest, Information support in education process of an ontology design, Proceedings of the IV International research conference "Information technologies in Science, Management, Social sphere and Medicine" (ITSMSSM 2017), vol. 72, 2017, pp. 404-407.

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА ПРОЦЕССА ОСВОЕНИЯ ИНСТРУМЕНТА ПРОЕКТИРОВАНИЯ ОНТОЛОГИЙ

Малочкина А.В., Самарский университет

Боргест Н.М., Самарский университет, ИПУСС РАН

В статье описаны пути реализации информационной поддержки для дисциплины «Онтология производственной сферы», предназначенной как для студентов, так и для преподавателей. Цель данной работы внедрить инструменты для создания онтологий в процесс проектирования самолета с целью научить студентов находить наилучшие решения конструкторских задач. Информационная поддержка должна обеспечить эффективное взаимодействие студентов и преподавателя. Использование информационной поддержки снижает временные затраты на рутинные задачи. Большое внимание уделяется содержанию данного курса с целью максимально адаптировать его для представления в электронном виде. Здесь описаны методы и средства создания онтологий, выдвинуты рекомендации к содержимому информационной поддержки. Разработанная поддержка будет использоваться как информационный ресурс для вышеупомянутой дисциплины в Самарском университете.

Fuzzy logic usage for the data processing in the Internet of Things networks

Larysa Globa
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine
lgloba@hotmail.com

Vasyl Kurdecha
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine
kvv.vasyl@gmail.com

Ivan Ishchenko
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine
ivanishchenkoo@gmail.com

Andrii Zakharchuk
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine
96andriy@ukr.net

Nataliia Kunieva
Igor Sikorsky Kyiv Polytechnic Institute
Kyiv, Ukraine
kuneva.nataly@gmail.com

***Abstract*—An overloaded Internet of Things (IoT) network needs data processing mechanisms to analyze the state of the system from the input parameters. There is not enough simple and statistical data analysis methods when operating with a large volume of data. The paper deals with the approach to big data analyses that use of fuzzy logic and fuzzy knowledge base for big data processing. The mathematical methods that allow to obtain efficiently the required solution in conditions of high performance requirements are considered. The proposed approach is quite complex on computational costs in training but it allows further to use not complex algorithms for big data processing. This is very important for Internet of Things systems.**

***Keywords*—Internet of Things, fuzzy logic, fuzzy knowledge base, data processing, peer-to-peer connection.**

I. Introduction

The Internet of Things (IoT) paradigm enables interconnected among devices-anytime, anywhere on the planet-providing the Internets advantages in all aspects of daily life. Analysts predict that the IoT will comprise up to 26 billion interconnected devices by 2020 [1]. Modern computing and distributed information systems (for example, cloud computing systems, networks function on the dynamic architectures of VANET and MANET, smart-type networks, intelligent controllers and sensors networks, etc.) are complex telecommunication networks that include many different types of network devices, which are integrated into the information and computing block, operate with high network and computational load. The conventional Internet has proved valuable in almost all trends by giving people the ability to interact with global information and services.

Extensive data has led to an explosive growth in the popularity of a modern data processing methods and the effective utilization of the bandwidth of the telecommunication network, because the amount of information has become much larger and it is by its nature and content becoming more diverse and large.

The feature of these systems is that they should control of the devices functioning or make decisions about the behavior of the system in different situations. Developing such kind of the systems requires powerful computers usage. The feature of IoT systems is the impossibility of using complex calculations, since the client device is quite often a microcontroller that can not perform complex operations. Therefore, the paper deals with the approaches to processing a large number of data generated by the IoT system, and discuss possible approaches for managing IoT systems that can directly functioning on the client devices.

A. *Maintaining the Integrity of the Specifications*

Currently, when all mobile and most consumer digital devices have network interfaces that are used to exchange user data, branching and multicomponent of such networks, heterogeneity and virtually unlimited volume of nodes can determine the new phenomenon of network technologies - Internet of Things.

According to the Cisco IBSG Consulting Group the number of devices which are connected to the Internet by 2020 will increase to 50 billion [2]. For the implementation of the Internet of Things systems, we need to further develop specialized methods put forward for such systems. Goals that can be used for verification, and transmitted in accordance with the rules included in the system. The usual mathematical methods are too complicated for IoT systems, so there is a need for simpler algorithms that can provide the quality of the received technical solutions. In addition, it is important that during the adoption of IoT systems, it is often not possible to use Boolean logic, because sometimes it is not possible to describe the system parameter with two possible states.

The mathematical apparatus of fuzzy logic is used when choosing the best moment to increase or decrease the temperature of the environment, taking into account the external conditions. In addition, fuzzy logic can help reduce the energy consumption of urban buildings to achieve zero energy potential [3]. The basic idea of fuzzy logic is that the intellectual way of reasoning based on the natural language of human communication can't be described by traditional mathematical formulas. A formal approach is characterized by a strict uniqueness of the interpretation, and all that is associated with the use of natural language, has a significant interpretation. Fuzzy logic is intended to formalize human abilities to inaccurate or approximate considerations that allow more adequately describe situations with uncertainty [4].

L. Zade introduces the concept of fuzzy set and, along with this, he also proposes to generalize classical logic with an infinite number of truth values. In the theory of fuzzy logic, the true meanings of statements can take any truth values from the actual interval numbers [0; 1]. This provision allows us to construct a logical system in which judgments can be made with uncertainty and to assess the degree of truth of statements. One of the concepts of fuzzy logic is the concept of elementary fuzzy statement. An elementary fuzzy statement is a narrative sentence, which expresses a complete opinion about which one can judge whether it is true or false with some definite degree of confidence.

II. Basic concepts of the fuzzy sets theory

A. *Structure identification of the fuzzy system*

Identification of fuzzy systems is based on observational data, but it is still not possible to exclude from this process the participation of an expert who solves the following tasks: determining the type of fuzzy system (Singleton, Mamdani or Takagi-Sugeno), choice of t-normal functions for fuzzy logic operations, choice of methods for fuzzy output (for a system like Mamdani) [5].

The choice of the system type is determined by the task that needs to be solved. If the problem of interpolation or approximation is solved, and accuracy is a determining factor, then the choice should be made in favor of the fuzzy system of the Takagi-Sugeno type. If the task of obtaining knowledge of data (in the form of linguistic rules) or the search for associative connections on a plurality of data is solved then for these purposes it is necessary to use a fuzzy system of the type Mamdani. The undoubted advantage of such models is their comprehensiveness and interpretation. The disadvantage is that in high computational costs. The type of the Singleton system can be used both for solving problems of approximation and obtaining knowledge [6].

Fuzzy model is defined as a system with n input variables defined on the input field of attributes, and one output variable Y defined on the original domain of DY. The exact value that the input variable X_i, accepts is denoted as x_i , and y is used for the output variable Y.

The fuzzy area of definition of the *i-th* input variable X_i is denoted by, where p_i – is the number of linguistic terms on which the input variable is defined, $LX_{i,k}$ defines the membership function and the name of the *k-th* linguistic term. Similarly – fuzzy field of definition of the output variable, q – number of fuzzy values, LY_i – membership function and the name of the original linguistic term.

The rule base in a fuzzy system like Mamdani is the set of fuzzy rules of the form:

$$R_j : LX_{1,j_1} \text{ AND } \ldots \text{ AND } LX_{n,j_n} \rightarrow LY_j. \quad (1)$$

The fuzzy $j-th$ rule in the Singleton system has the form:

$$R_j : LX_{1,j_1} \text{ AND } \ldots \text{ AND } LX_{n,j_n} \rightarrow \ r_j, \quad (2)$$

where r_j is the number by which y is evaluated. Fuzzy $j-th$ rule in the Takagi model – Sugeno has the form:

$$R_j : LX_1, j_1 \ldots \ LX_n, j_n \rightarrow r_0 j + r_{1j} x_1 + \ldots + r_{nj} x_n, \quad (3)$$

where the output y is evaluated by a linear function.

An important role is played by the membership function $\mu_{LXij}(x_i)$, which indicate the degree of belonging to a clear variable x_i fuzzy concept LX_{ij}. The training of the system is based on observation tables or test functions $f(x)$. In general, the construction of a fuzzy system consists of the following main stages:

- Expert evaluation – the type of fuzzy system and its associated parameters.
- Structure identification – the choice of variables (X, DX, Y, DY, pi, q) and fuzzy rules R_j
- Parameters identification – the search for optimal values of all parameters involved in the fuzzy system: searching of the values of the consequent (THAN-parts of the rule) and the parameters of membership functions in the antecedent (IF-part) of each rule (FX_i, FY) based on the given criteria of quality and optimization method selected criteria.
- Verifying the correctness of the model.

To make decisions in a fuzzy system, it is proposed to use the process of identifying a structure - the definition of the structural characteristics of a fuzzy system, such as the number of fuzzy rules, the number of linguistic terms, to which the incoming variables are divided. This identification is performed with the help of a fuzzy cluster analysis or a method of selection [7].

The usage of cluster analysis algorithms is aimed at dividing the set of data into clusters so that each of them has the closest objects. Fuzzy clustering is one of the most interesting methods for identifying possible groups and testing hypotheses about the data structure. Methods of fuzzy clustering allow one and the same object to belong simultaneously to several clusters, but with varying degrees of affiliation. Usually, each cluster is characterized by some prefix, which is described by the cluster center and some additional information, such as the size and form of the cluster. There are a number of clustering techniques. The main methods of fuzzy clustering: fuzzy decision trees, fuzzy Petri nets, fuzzy self-organizing maps [7].

Fuzzy clustering in many situations is more "natural" than clear, for example, for objects located on the border of clusters. The most common: the algorithm of fuzzy self-organization of c-means and its generalization in the form of the Gustafson-Kessel algorithm.

B. Algorithm of fuzzy self-organization c-means

Purpose: clustering large sets of numeric data.

Advantages: fuzzy when assigning objects to the cluster, allows you to identify objects that are on the border of the cluster.

Disadvantages: computational complexity, specifying the number of clusters at the input of the algorithm, there is uncertainty with objects that are significantly removed from the centers of all clusters.

In the classical algorithm of k-means (c-means), elements are chosen using the ordinary Euclidean distance between the vector x and the center of the cluster c. With this assignment, the distance between two vectors of a set of points equidistant from the center takes the form of a sphere with the same scale along all the axes. But if the data creates groups whose form differs from the spherical one or if the scales of the individual vector coordinates are very different, in this case the metric becomes inadequate. In this case, the quality of clustering can be significantly improved by using an improved version of the self-organization algorithm, which is called the Gustavsson-Kessel algorithm [8].

C. Fuzzy decision trees

Fuzzy decision trees are used in Data Mining to solve classification problems and to solve the regression problem when it is necessary to know the degree of belonging to a particular outcome. They can be used in various fields: in banking for solving the problem of scoring, in medicine for diagnosing various diseases, in the industry for quality control of products and so on.

Unconditional advantage of this approach is the high accuracy of classification achieved by combining the advantages of fuzzy logic and decision trees. The learning process takes place quickly, and the result is simple for interpretation. Since the algorithm is capable of issuing for the new object not only the class, but also the degree of belonging to it, it allows to control the classification threshold [9].

However, the complexity of applying the algorithm is the necessity of having a representative set of learning examples, otherwise the decision tree generated by the algorithm will slightly reflect the reality and, as a result, produce wrong results.

D. Fuzzy Petri nets

In the Petri time networks, conditions are represented by a set of positions, and their execution is represented by marking the corresponding position. Placing a certain number of labels in the given position after a specified time. In the basis of the study of the listed properties lies the feasibility analysis. Methods for analyzing the properties of Petri nets are based on using graphs of achievable (covering) markings, solving the equation of network states and calculating linear invariants of positions and transitions. Also, auxiliary reduction methods are used, which make it possible to reduce the size of the Petri net while preserving its properties, and the decomposition that divide the source network into subnets [10]. A feature of Petri nets is the possibility of presenting fuzzy processes and the dynamics of their interaction. The disadvantage is that many parameters, indicators and characteristics are not taken into account, without which it is difficult to imagine the real processes of practical implementations. In addition, the possibility of specifying the indistinctness of the marking and the components of the incidence function is limited, which significantly limits the possibilities of the researcher.

E. The result of the fuzzy algorithms

All of the considered algorithms have a common disadvantage of the algorithm complexity and this is very important as the IoT network has a rigid frame of time delays in processing information, obtaining inference, etc. Because of this problem, an approach is taken to form the basis of fuzzy knowledge base which allows using previously obtained knowledge about the behaviour of the system or process, which allows us to do the analysis of data based on certain rules [11].

With this approach, the algorithm for processing information for large data is as follows:

- The behavior of the system is determined based on the input set of data obtained in the form of rules or relationships that can be represented in the form of fuzzy knowledge base.
- According to the obtained model of the behavior of the system of fuzzy logic output conducts its state or development trends. This approach is most applicable for the IoT system because it allows you to reduce the number of computational operations for the system.

The formed fuzzy knowledge base, which displays the dependencies of the system with known input parameters, can be used to determine its system states. This is possible due to fuzzy logic output, which in turn can be implemented in different ways.

A fuzzy logical inference is the approximation of the dependence $y = f(x_1, x_2 ..., x_n)$ with the help of a fuzzy knowledge base and operations on fuzzy sets. In order to fulfill a fuzzy logical inference, the following conditions are required [12]:

- There should be at least one rule for each term of the output variable.
- For any term of an input variable there should be at least one rule in which this term is used as a prerequisite.
- There should be no contradictions and correlations between the rules

In "Fig. 1", shows the sequence of actions using the process of fuzzy logical inference

The process of fuzzy logical inference is a procedure or algorithm for obtaining fuzzy inference based on fuzzy conditions or prerequisites.

Systems of fuzzy logical inference can be considered as a separate case of so-called production fuzzy systems. In such systems, the conditions and logical inference of different rules are formed in the form of fuzzy statements made regarding the values of some linguistic variables. The development and

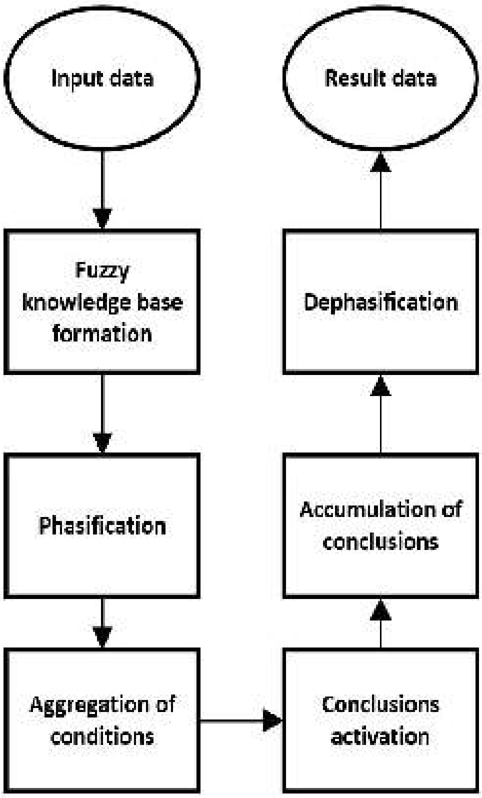


Figure 1. The sequence of actions using the process of fuzzy logical inference.

application of fuzzy logical inference systems consists of several stages, implementation of which is carried out with the help of the basic provisions of the theory of fuzzy sets.

Input of the fuzzy logic system are variables that carry information, which are obtained in some way, for example, by measuring some physical parameter of the system. These parameters are considered as real variables in the management processes [13]. The output of the system is a formed control variable of the fuzzy logic inference. Thus, fuzzy logic systems change the value of the input variables of the control process into output variables based on certain fuzzy rules. For IoT systems, the most effective is the application of fuzzy logic inference in the form of a set of rules such as fuzzy products, which are written in the form:

$$Rule < \# >: IF\ "\beta 1\ is\ \alpha 1", THAN\ "\beta 2\ is\ \alpha 2" \quad (4)$$

In the equation (4), the fuzzy statement $"\beta 1\ is\ \alpha 1"$ is a condition of this rule. The false statement $"\beta 2\ is\ \alpha 2"$ is the fuzzy inference of this rule. They are formulated in terms of fuzzy linguistic statements. It is assumed that $"\beta 1\ \neq \beta 2"$. The main stages of obtaining a fuzzy logical inference of the Mamdani system and the details of each stage are discussed in more detail below:

- Fuzzy logic base formation. The rules base of fuzzy logical systems is intended for a formal description of empirical knowledge or experts knowledge in a problem area and is a set of fuzzy rules of the form:

$$Rule_1 : IF\ "Condition_1", THAN\ "Conclusion_1"$$
$$Rule_n : IF\ "Condition_n", THAN\ "Conclusion_n" \quad (5)$$

where F_i (i belongs to 1, 2, ..., n) are the coefficients of certainty or the weighting coefficients of the corresponding rules. They can acquire values from the interval $[0; 1]$. Unless otherwise specified, $F_i = 1$. The rule base is given if it has set many rules for fuzzy products, as well as many input linguistic variables and the set of output linguistic variables.
- Phasification is a process or procedure for obtaining the values of the functions of fuzzy sets (terms) belonging to the basis of given input data. As a result of this phase completion, for all input variables, the specific values of the membership functions for each of the linguistic terms used in the set of conditions of the rules base of the fuzzy logical inference should be determined.
- Aggregation of conditions - is the procedure that determines the conditions truth degree for each of the rules of the fuzzy logical inference system. When the rule condition has a simple form, its truth is equal to the corresponding value of the function of the input of the input variable to the term used in this condition.
- Activation is the process of finding the degree of truth of each of the underlying conditions of fuzzy rules. Before the start of this stage it is assumed that the known degree of truth and weight coefficient F_i for each rule. Next, each of the outputs of the rules of the fuzzy logic inference system is considered. When the inference of the rule is one vague statement, the degree of its truth is equivalent to the algebraic product of the corresponding degree of truth of the condition of the weight coefficient.
- Accumulation is the process of finding an accessory function for each of the output linguistic variables. The purpose of the accumulation phase is to combine all the degrees of truth of the outputs to obtain the function of belonging to each of the output variables. The conditions that relate to the same source linguistic variable belong to different rules of the system of fuzzy logical inference.
- Dephasification (clarification) is a procedure for determining the usual (strict) value for each of the output linguistic variables. The purpose of this phase is to obtain, using the results of the accumulation of all output linguistic variables, the usual quantitative value of each of the output variables. This value can be used by special devices that do not belong to the fuzzy logic system.

After analyzing the main stages of the fuzzy logical inference , we can conclude that the most important stage of the logical output is the phase of defazification, since it is the stage in which the numerical value of the resulting variable is determined. But there are several ways to carry out this phase. The use of one algorithm for various systems is at least not efficient, as it is unacceptable for some systems to use heavy mathematical operations, and the requirements for the accuracy of calculations are insignificant. This in turn allows you to use more simple algorithms. In particular, this is especially important if you implement a pre-processing of information on the client part of the Internet of Things system. Therefore, it is simply necessary to somehow choose the algorithm of data processing [13].

Developments in this direction have been under way for a long time, and already a number of systems and regulators that are capable of self-learning are built and the choice of algorithms for processing in such systems goes back to the background, as during the training of the system the parameters of its accessories functions change to achieve the maximum possible accuracy. Such a process is quite long, it is most efficient to run on servers, and if there are several fuzzy knowledge bases in the system, and for each one it is necessary to undergo training, even for a powerful computer, this will be a difficult task. And as we are talking about IoT systems, to such databases, besides, quite a lot will be completely different. In addition, such an approach is not entirely appropriate for its use in systems with pre-processing information on client devices of the IoT system.

Each of the algorithms can be used in the real system, the main difference between these algorithms is the accuracy of the calculation and the operating time, they are different in each algorithm. But this is not the only criterion for analysis, a more detailed analysis can be seen in the table:

Table I
CRITERIONS OF CHOOSING DEPHASIFICATION ALGORITHMS

Method	Complexity	Sensitivity	Precision
Medium the maximum	low	absent	middle
First of the maximum	middle	absent	middle
The last maximum	low	middle	below average
Centre of Gravity	high	significant	high
Centre of Area	very high	significant	high
Height middle	low	high	above average

An overloaded Internet of Things (IoT) network needs data processing to analyze the state of the system from the input parameters [14]. When working with large volumes of data, there is not enough simple and statistic data analysis. Needed a more sophisticated data analysis mechanism [15].

III. CONCLUSIONS

Thus, the analysis of the mathematical methods of data analysis that can be carried out in IoT systems allows us to draw the following conclusions:

- Different data analysis methods allow you to obtain different data, which in turn are close to real ones, which is evidence that all these algorithms can be used in different systems. For real systems it is necessary to carry out the analysis of algorithms for the formation of fuzzy knowledge base and application of fuzzy logical inference, as well as to conduct the training of the system during the knowledge base formation and during the selection of algorithms for fuzzy logical inference separately on real data sets).
- The task of choosing the correct fuzzy logical inference algorithm remains relevant.
- A possible direction for further development of this topic may be the search for selection criteria, and the construction of a system that, with the help of the developed criterion, will be able to choose the best configuration of the system of fuzzy databases.

REFERENCES

[1] R. Lea, M. Blackstock, Smart cities: An iot-centric approach, in: Proceedings of the 2014 International Workshop on Web Intelligence and Smart Sensing, IWWISS'14, ACM,NewYork, NY, USA, 2014, pp. 12–15.

[2] M. Nikravesh, Evolution of fuzzy logic: From intelligent systems and computation to human mind, Stud. Fuzziness Soft Comput. 2007, pp. 37–53.

[3] G. Cueva-Fernandez, J. Pascual Espada, V. García-Díaz, R. Gonzalez-Crespo, Fuzzy decision method to improve the information exchange in a vehicle sensor tracking system, Appl. Soft Comput. 2015, pp. 708–716.

[4] Sugeno M., Yasukawa T. A fuzzy-logic-based approach to qualitative modeling // IEEE Trans. on Fuzzy Systems, 1993, vol. 1, pp. 7–31.

[5] Emami M.R., Turksen I.B., Goldenberg A.A. A unified parameterized formulation of reasoning in fuzzy modeling and control // Fuzzy Sets and Systems, 1999, vol. 108, pp. 59–81.

[6] Espinosa J., Vandewalle J., Wertz V. Fuzzy logic, identification and predictive control. — London: Springer-Verlag, 2005, 263 p.

[7] Kurdecha V.V., Ishchenko I.O., Zakharchuk A.G. An approach to the Internet of Things system architecture // CADSM'2017 14th International conference the experience of design and application of CAD systems in microelectronics, 2017, pp. 1–3.

[8] D.F. Larios, J. Barbancho, F.J. Molina, C. León, Lis: Localization based on an intelligent distributed fuzzy system applied to a wsn, Ad Hoc Networks, 2012, pp. 604–622.

[9] Globa L.S., Ishchenko I.O., Zakharchuk A.G. Intelligent support system for e-Health // 20th International Multi-Conference on Advanced Computer Systems Miedzyzdroje, Poland, pp. 1–4, October 2016.

[10] Fuzzy knowledge bases integration based on ontology Ternovoy, M., Shtogrina, O. 2012 Modern Problems of Radio Engineering, Telecommunications and Computer Science - Proceedings of the 11th International Conference, TCSET'2012, pp. 1–8.

[11] Method of heterogeneous information resources structuring and systematizing for Internet portals development Globa, L., Novogrudska, R., Oriekhov, O. 2013 IEEE EuroCon, pp. 10–20.

[12] Menard M. Fuzzy clustering and switching regression models using ambiguity and distance rejects / M. Menard // Fuzzy Sets and Systems, 2001, pp. 363–399.

[13] Chuang C.C. TSK fuzzy modeling for function approximation with outliers / C. C. Chuang, S. F. Su, S. S. Chen // IEEE Trans. on Fuzzy Systems, 2001, pp. 810–821.

[14] Andrzej Piegat, Fuzzy Modeling and Control //BINOMIAL. Knowledge lab , 2013, pp. 179–400.

[15] Koucheryavy A. The QoS Estimation for Physiological Monitoring Service in the M2M network / A. Koucheryavy, Y. Al-Naggar // Internet of Things and its Enablers (INTHITEN): Proceedings. Conference, State University of Telecommunication (SUT), St. Petersburg, Russia, 2013, pp. 133-139.

ИСПОЛЬЗОВАНИЕ НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ОБРАБОТКИ ДАННЫХ В IoT СЕТИ

Глоба Л.С., Курдеча В.В., Ищенко И.А., Захарчук А.Г., Кунева Н.В.
Национальный технический университет Украины "Киевский политехнический институт имени Игоря Сикорского"

Современные IoT (Интернет вещей) системы продуцируют значительные объемы данных, называемые «большими» данными. Объемы «больших» данных возрастают экспоненциально, поэтому для IoT систем нужны эффективные механизмы обработки и анализа данных, быстрого определения состояния системы по входным параметрам.

Особенностью IoT систем является то, что они должны контролировать работу устройств или принимать решения о поведении системы в разных ситуациях. При создании программного обеспечения для IoT систем следует учитывать невозможность использования сложных вычислений, поскольку клиентское устройство довольно часто является микроконтроллером, который не может выполнять сложные операции. В связи с этим в статье рассматриваются подходы к обработке большого количества данных, генерируемых IoT системой, обсуждаются возможные подходы к созданию алгоритмов управления ее состоянием, которые могут непосредственно функционировать на клиентских устройствах или узлах с малой вычислительной производительностью.

Анализ математических методов обработки и анализа данных, которые может использовать IoT система, позволяет сделать такие выводы:

- различные методы анализа данных позволяют получать выводы, которые, в свою очередь, близки к реальным, что свидетельствует о том, что все множество алгоритмов можно использоваться в разных системах, однако временные и вычислительные затраты при этом будут значительными;
- не все алгоритмы и методы учитывают возможную нечеткость границ изменения состояния системы и могут формировать правила перехода между этими состояниями;
- анализ методов обработки «больших» данных показывает, что недостаточно простых и статистических методов анализа данных при работе с большим объемом данных. В статье предлагается подход к анализу «больших» данных, использующий методы нечеткой логики и нечеткую базу знаний для их последующей обработки. Рассмотрены математические методы, позволяющие эффективно получить требуемое решение в условиях высоких требований к производительности.

Однако, в каждом конкретном случае для создания реальных систем необходим анализ наиболее эффективных алгоритмов формирования нечеткой базы знаний и проведения нечеткого логического вывода, а также предварительное обучение системы во время формирования базы нечетких знаний и при выборе алгоритмов для проведения нечеткого логического вывода. Проведение обучения всей системы на реальных наборах данных возможно выполнить в автономном режиме. Таким образом, предлагаемый подход не смотря на то, что является довольно сложным по затратам на обучение на значительных реальных наборах данных, все же позволяет использовать не сложные алгоритмы для обработки «больших» данных, что очень важно для IoT систем. Возможным направлением дальнейшего развития этих исследований является поиск критериев выбора наилучшей конфигурации системы нечетких баз данных при ее формировании.

Patent Landscapes & New Technology Trends in IoT: Extracting and Visualizing Data Patterns

Nikolay Stulov
Department of Control and Applied Mathematics
Moscow Institute of Physics and technology
Moscow, Russia
nikolay.stulov@phystech.edu

***Abstract*—Extracting complex relations in unstructured data is a challenging and promising task in any field, especially fast-growing like Internet of Things (IoT). In this work we research different methods to extract and represent these relations. As a result, we present a set of text mining and patent mining tools and an approach to further building knowledge-based decision support system.**

***Keywords*—patent mining, text mining, ontology, decision support, patent landscape, information extraction, ontology driven information extraction, Internet of Things**

I. Introduction

A. Objective and relevance

In this work we research the mechanism of patent analysis and propose an intelligent system for this task. Formally, the problems are stated as follows:

- Extract and store knowledge of a specific field from a set of partially structured documents.
- Analyze and compare tools for visual representation of a subset of this knowledge from model.
- Propose a function to calculate probabilities of existence or appearance of model-unknown connections.

B. Existing Approaches

Currently a huge part of patent analysis in industry is done manually with the help of experts. Some approaches are known to automate this task using co-citation analysis [1], [2], [3]. Other approaches include ontology driven analysis as in [4], [5]. This work continues and improves discussed approach by taking implicit linguistic data from abstracts, claims and full texts into consideration.

Visualization of big bibliographic networks is usually done with the help of Visualization of Similarities (VOS) [6] algorithm, which is basically specifically weighted Multidimensional Scaling (MDS) [7]. This work proposes a number of different ways to set weights and compares them to VOS.

The rest of the paper is organized as follows. The next section describes steps of proposed approach in detail. Also, some examples and visualizations are given. After that there is evaluation section where analysis results are discussed. We conclude with further research review and comparison with above stated analogies.

II. Our approach detailed

A. Semantic-aware knowledge extraction

Since the main feature of our approach is implicit linguistic information retrieval, this step requires usage of semantic technologies. In this case we use Part of Speech (POS) tagging and semantic features of patent genre to make use of different entities that appear in the texts. Extraction was done as follows. Pre-structured data is extracted from XML as-is, but abstract and full text are treated individually. With POS tagging we filter out only noun groups (NG) of two words (bigrams). After that the sentences are encoded by the number of occurrences of each NG, a matrix representation X of text is built. Then we use outer product on X to build first order collocation matrix T, which is then scaled with Term Frequency — Inverse Document Frequency (TF-IDF) scale. Finally we run PageRank on T and choose $n = 100$ highest ranked noun groups as output (keywords).

B. Ontology

To represent knowledge we designed an ontology of our field, which includes entities "Patent", "Author", "Assignee", "Region", "Class" and "Keyword", trivial accessory relations, a citing relation and analogy relation between two patents. For implementation purposes graph database was used. The graph is very complicated due to the number of entities, as expected. For example, the citation network is depicted in Fig. 1.

Some interesting analysis can be applied at this point already. Sorting out authors who have more than one patent (active authors) leads to obvious clustering (see Fig. 2), which appears to be regional first (see Fig. 3).

Some predictions can also be made. As Fig. 3 shows, patents with unavailable regional information can be assigned to a certain region.

C. Building landscapes

Images above only depict subsets of ontology graph and posses no information about likeness between entities. To build landscapes means to project multidimensional data into two-dimensional. There are many methods to tackle this problem.

The MDS approach optimizes loss function:

$$\mathcal{L}(\mathbf{x}_1, ..., \mathbf{x}_n) = \frac{\sum\limits_{i<j} w_{ij} \left(f(p_{ij}) - ||\mathbf{x}_i - \mathbf{x}_j||_2 \right)}{\sum\limits_{i<j} w_{ij} f(p_{ij})^2} \quad (1)$$

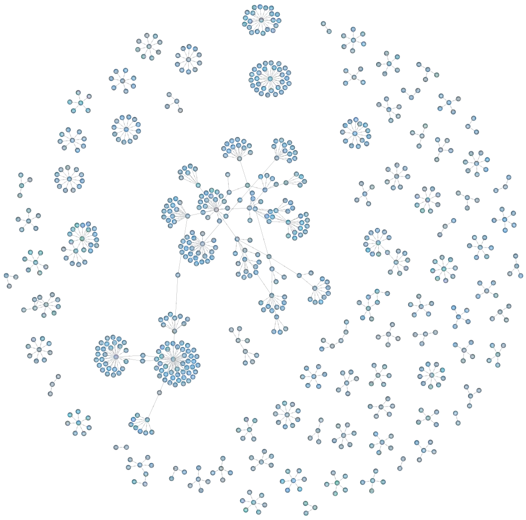


Figure 1. Patents (blue) co-citation network.

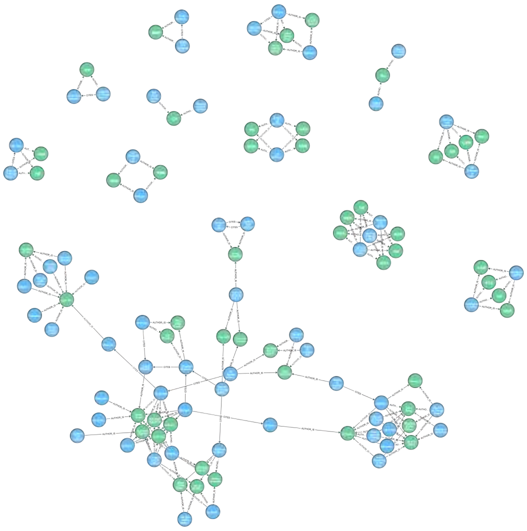


Figure 2. Patents (blue) and active authors (green) network.

where w_{ij} are weights, f denotes transformation function of proximity values p_{ij}. Usually the weights w_{ij} are set to 1.

In [7] they show that VOS solution is equivalent to MDS solution with $p_{ij} = \frac{1}{s_{ij}}$ and $w_{ij} = s_{ij}$, where

$$s_{ij} = \frac{2mc_{ij}}{c_i c_j} \tag{2}$$

where c_i denotes the total number of links of node i and m denotes the total number of links in the network.

This approach proved good at building bibliographic maps. But as soon as network contains information of different types from different sources, it becomes possible to use other metrics more efficiently. To do this we need to vectorize objects in some way. We propose using euclidean distance (ED), cosine similarity and traingle's area similarity - sector's area similarity (TS-SS) [8] on concatenation of attribute vectors.

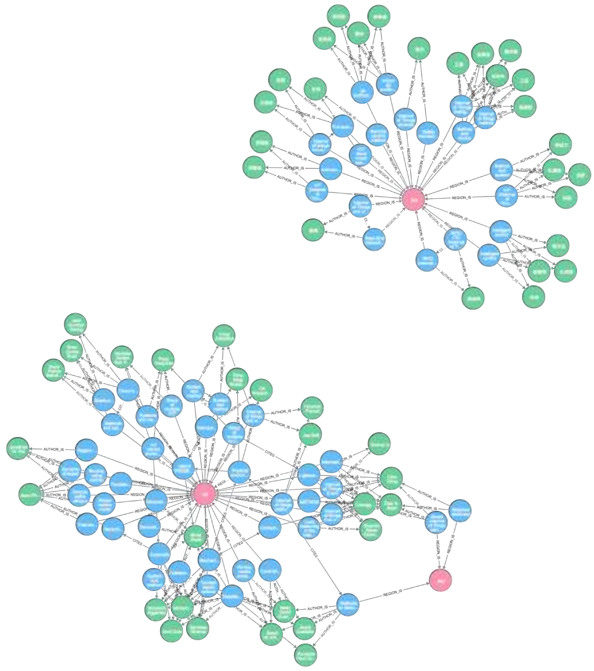


Figure 3. Regional subnetworks.

Cosine similarity is given by:

$$V = \text{cosine}(\mathbf{a}, \mathbf{b}) = \frac{(\mathbf{a} \cdot \mathbf{b})}{||\mathbf{a}||_2 \cdot ||\mathbf{b}||_2} \tag{3}$$

Euclidean distance (ED) is given by:

$$\text{ED}(\mathbf{a}, \mathbf{b}) = ||\mathbf{a} - \mathbf{b}||_2 = \sqrt{\sum_{i=0}^{n} (a_i - b_i)^2} \tag{4}$$

TS-SS is given by:

$$\begin{aligned} \text{TS-SS}(\mathbf{a}, \mathbf{b}) = \frac{\pi \sin(\theta')\theta'}{720} ||\mathbf{a}||_2 ||\mathbf{b}||_2 \cdot \\ \cdot (\text{ED}(\mathbf{a}, \mathbf{b}) + \text{MD}(\mathbf{a}, \mathbf{b}))^2 \end{aligned} \tag{5}$$

where Magnitude distance (MD) is:

$$\text{MD}(\mathbf{a}, \mathbf{b}) = \left| \sqrt{\sum_{i=0}^{n} a_i^2} - \sqrt{\sum_{i=0}^{n} b_i^2} \right| \tag{6}$$

In most trivial cases attribute vectors are one-hot encoded attributes. In case of textual attributes collocation matrix row is used. In case of citation attributes co-citation matrix row is used.

For clustering VOS solves the same task [9], meaning same drawbacks. We propose to run metric clustering algorithms on resulting vectors.

Finally, to estimate the probability of co-authorship we propose to use Bayes rule. Let A be event that two authors have am article and B the event of occurrence of their meta-data together. We estimate prior distribution of A with:

$$\mathbb{P}\{A\} = \frac{\#\{\text{patents of this author}\}}{\#\{\text{patents total}\}} \tag{7}$$

Then, according to Bayes rule:

$$\mathbb{P}\{A \,|\, B\} = \frac{\mathbb{P}\{B \,|\, A\}\mathbb{P}\{A\}}{\sum\limits_{C} \mathbb{P}\{B \,|\, C\}\mathbb{P}\{C\}} \tag{8}$$

III. Evaluation

For demonstration purposes Internet of Things field was chosen. Patent data for research was acquired through European Patent Office (EPO) API. This tool allows to get data such as inventor names, assignee, different dates, classification, citations both ways, abstracts, claims and full texts. A corpus of 150 documents is used.

After vectorizing patent data in the way discussed earlier, we apply MDS using three proposed dissimilarity metrics and VOS-original association strength for comparison. All the images share the legend: blue dots account for US and WO region, red dots for CN region and green dots for KR region. The edges correspond to large number of common keywords (at least 0.1%).

Applying euclidean, cosine and TS-SS distance leads to graphs shown in Fig. 4, Fig. 5, Fig 6 respectively. All the graphs feature visible separation of US and CN regions.

In detail, euclidean distance unsurprisingly draws attention to extracted keywords as this subvector is the most dense. Therefore patents group when they share similar semantic profile, leading to topic distinction.

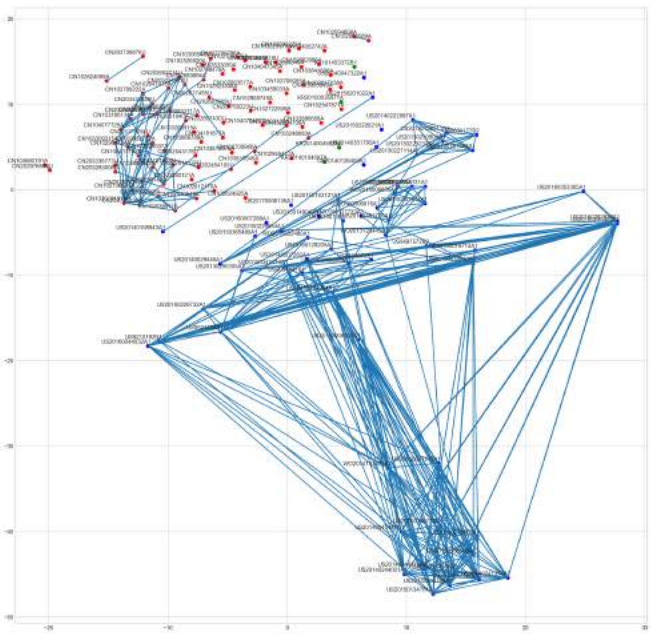


Figure 4. Euclidean distance.

Cosine distance, on the contrary, draws attention to classes, assignees and regions as they are represented as one-hot subvectors and affect angle rather than magnitude.

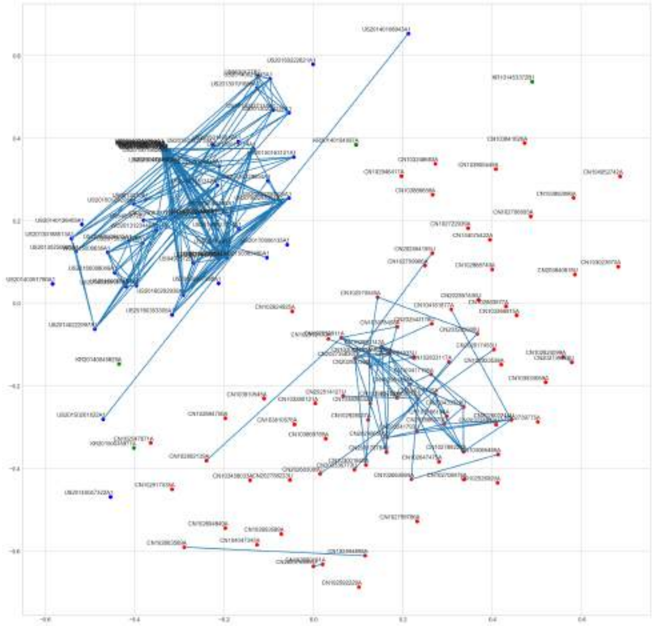


Figure 5. Cosine distance.

The TS-SS distance combines both features and is generally harder to interpret.

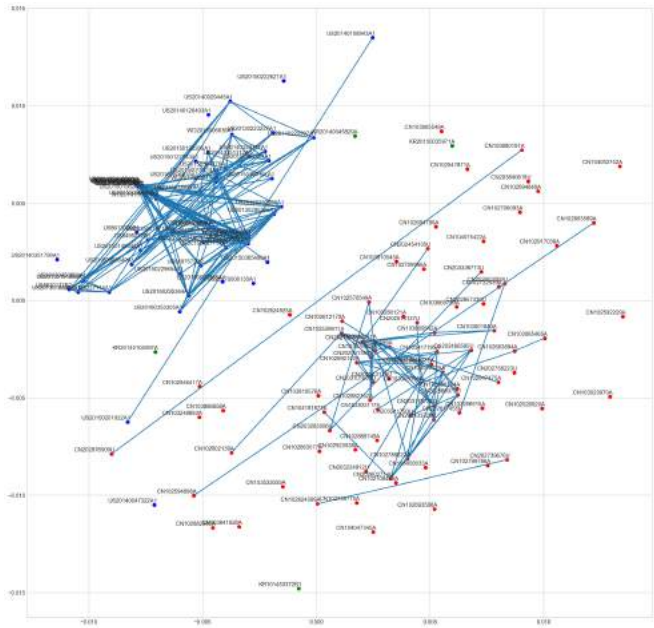


Figure 6. TS-SS distance.

IV. Further Research

As an application of proposed model we see decision support systems in patent analysis. Russian GOST for patent analysis states the tasks, that should be included in this research, including:

- Research of a technical level of objects of economic activity, revealing of tendencies, a substantiation of the forecast of their development
- A study of the state of the markets for these products, the prevailing patent situation, the nature of national production in the countries of study
- A study of directions of research and production activity of organizations and firms that operate or can operate on the market of products under study
- Justification of proposals on the feasibility of developing new industrial property for use in facilities that ensure the achievement of technical indicators foreseen in the technical task (tactical and technical task)

Most of them are directly linked to analysis and forecasting of patent landscapes, which is successfully achieved with the help of proposed system. To even improve the system and minimize expert involvement, Deep Learning can be used to mine relations as in [10]. Some improvements may also be achieved with the help of latent semantics and topic modelling, as existing patent classification was only briefly introduced to proposed model.

V. Conclusion

In comparison to current research, the proposed method includes more complex and detailed intellectual analysis of patents, including implicit linguistic factors.

Acknowledgment

Author would like to thank Vladimir Khoroshevsky for problem statement and useful discussion.

References

[1] D. Bonino, A. Ciaramella, and F. Corno, "Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics," *World Patent Information*, vol. 32, no. 1, pp. 30 – 38, 2010.

[2] I. V. Efimenko, V. F. Khoroshevsky, and E. C. M. Noyons, *Anticipating Future Pathways of Science, Technologies, and Innovations: (Map of Science)2 Approach*. Cham: Springer International Publishing, 2016, pp. 71–96.

[3] R. Winiarczyk, P. Gawron, J. Miszczak, L. Pawela, and Z. Puchala, "Analysis of patent activity in the field of quantum information processing," vol. 11, 12 2012.

[4] I. Efimenko and V. Khoroshevsky, "Peaks, slopes, canyons and plateaus: Identifying technology trends throughout the life cycle," vol. 14, p. 1740012, 11 2016.

[5] V. F. Khoroshevsky and I. V. Efimenko, "Russia among the world centers of excellence," in *Open Semantic Technologies for Intelligent Systems (OSTIS-2016)*. BSUIR, February 2016, pp. 223 – 232.

[6] N. J. van Eck and L. Waltman, *VOS: A New Method for Visualizing Similarities Between Objects*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 299–306.

[7] N. J. van Eck, L. Waltman, R. Dekker, and J. van den Berg, "A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS," *CoRR*, vol. abs/1003.2551, 2010.

[8] A. Heidarian and M. J. Dinneen, "A hybrid geometric approach for measuring similarity level among documents and document clustering," in *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, March 2016, pp. 142–151.

[9] L. Waltman, N. J. van Eck, and E. C. M. Noyons, "A unified approach to mapping and clustering of bibliometric networks," *CoRR*, vol. abs/1006.1032, 2010.

[10] R. Socher, D. Chen, C. D. Manning, and A. Ng, "Reasoning with neural tensor networks for knowledge base completion," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 926–934.

ИЗВЛЕЧЕНИЕ И ВИЗУАЛИЗАЦИЯ ПАТТЕРНОВ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ ПАТЕНТНЫХ ЛАНДШАФТОВ И ВЫЯВЛЕНИЯ НОВЫХ ТЕХНОЛОГИЧЕСКИХ ТРЕНДОВ В ОБЛАСТИ "ИНТЕРНЕТ ВЕЩЕЙ"

Николай Стулов
ФУПМ МФТИ

Извлечение сложных отношений из неструктурированных данных — это сложная задача в любой области, а в особенности в быстрорастущих областях, таких как Интернет вещей. В этой работе исследуются различные методы извлечения и визуализации этих отношений. В результате предлагается набор инструментов для обработки текстов и патентов, а также подход, который может быть использован для построения интеллектуальной системы поддержки принятия решений.

Hetero-Associative Memory Technology for Development of Intelligent Control Systems of Autonomous Mobile Robots

Ryhor Prakapovich
Robotic Systems Laboratory
United Institute of Informatics Problems of NASB
Minsk, Belarus
rprakapovich@robotics.by

***Abstract*—The problems of creating artificial intelligence technologies in the control of service robots are considered. Based on the synthesis of artificial neural networks and associative memory, a neural network model of iterative hetero-associative memory has been developed. This model is designed for recording, storing and processing sensory and control information. A neural network model of the hierarchical intellectual control system of autonomous mobile robots is proposed. This model, based on the available a priori knowledge, is able to function and adapt to changes in external conditions.**

***Keywords*—intelligence control system, autonomous mobile robot, artificial neural network, associative memory, knowledge base, OSTIS**

I. Introduction

Modern service and domestic robots must function in a dynamic human environment characterized by the uncertainty of external disturbances. This requires the involvement of non-traditional approaches to their control systems. These approaches should be based on the novel methods of knowledge bases, new types of feedbacks, modern information and telecommunication technologies. At this stage, these approaches can be combined under the common name of artificial intelligence [1].

One of the promising methods for developing intelligent control systems (ICS) can be bio-inspired. It is based on the modeling of obtaining and recognizing information processes in the nervous systems of life forms [2, 3]. Bio-inspired methods are based on the investigation of the fundamental principles of the natural control systems functioning and their evolution mechanisms. These principles determine life forms behavior, and not certain aspects of their demonstrations. In other words, bio-inspired control methods are based not so much on reference patterns classification algorithms, as on search algorithms that allow to adapt to another environmental conditions. As a rule, control systems of the life forms do not have analytical models. In addition, the necessary knowledge must be accumulate empirically during process of the control system interaction with the environment and with itself [1, 2].

Among the known methods of artificial intelligence, expert systems, artificial neural networks (ANN), associative memory and fuzzy logic can be especially highlighted [2,3]. However, to our point of view, technologies such as expert systems and fuzzy logic require the involvement of an operator (expert) in the control loop to a greater extent than the other two. Thus, ANN and associative memory are more suitable for the creation of ICS by autonomous robotic devices.

II. Artificial Neural Networks as a Units of Intelligent Control System

One of the leading directions in the field of artificial intelligence is associated with the modeling of ANNs. ANNs are able to solve a wide range of problems of image recognition, identification, prediction, optimization and control of complex objects [1, 2]. ANNs are creates on the principles of organization and functioning of their biological analogs.

Advantages of ANNs over traditional methods of artificial intelligence (production systems, decision tables, genetic algorithms, fuzzy logic, etc.) are manifested in solving problems that are characterized by problems in the allocation of rules for the functioning of systems. As a rule, such problems have a large number of possible solutions, but there is the possibility of learning on a lot of examples [4].

ANNs are also successfully used in dynamic object control systems [3,4]. ANNs have a number of unique properties that make them a powerful tool for creating control systems: the ability to learn by example, the generalization of data, the ability to adapt to changing the properties of the control object and the external environment, the suitability for the synthesis of nonlinear regulators, high resistance to damage to its elements (in force originally embedded in the ANNs parallelism) [1-5].

Thus, ANNs are convenient to use in those applications where there is no possibility of obtaining training samples from experts in a particular field. In this case, there is a process of self-organization, in which the ANNs changes its functional structure without special intervention from outside (see Figure 1). Since the process of self-organization is manifested in the fact that the ANNs independently selects samples from the input data stream for training, the choice of the ANNs model must take into account the specificity of incoming information, measuring and control means.

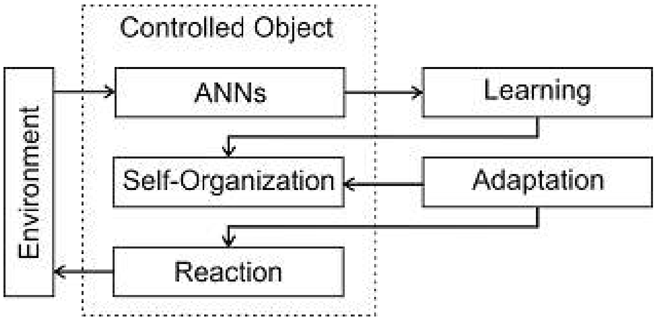


Figure 1. The adaptation process of the robot to environmental conditions through the self-organization of ANN.

Nevertheless, ANNs also have its limitations due to the lack of a universal architecture and unified approaches for use in control systems, as well as problems of additional learning of new information. At the moment in the developed control systems ANNs are used in the form of separate blocks and, mainly, only for the implementation of functions of classification or data interpolation, which significantly narrows their potential.

In special control tasks for autonomous mobile robots (AMR), ANNs are often criticized [6]. Firstly, this is due to the fact that most ANNs are not control systems, but merely imitate the recognition systems. In other words, the function "input-output" is search from some variables, and the values of the objective function of the system are specified in other variables. Secondly, most classical learning algorithms such as backward propagation errors work only in supervisory mode, that is, they are not self-learning [6]. And, finally, thirdly, similar ANNs do not work in conditions of post-training, when in the process of adding a new pattern to the trained network, the results of previous training are destroyed or changed.

III. Neural Network Model Hetero-Associative Memory for Recording, Storing and Processing of Information

On the basis of S.Grossberg adaptive resonance theory [7-10] and the ANN be-directional associative memory (BAM) [11-13] is proposed an adaptive neural network classifier ART-BAM with the retrain function. The ART-BAM classifier can be used in the autonomous mobile robot control system. With the help of the original non-iterative algorithm for new patterns detecting and the available of reference patterns in the long-term memory, the classifier ART-BAM is able to register new patterns in the sensory data stream [14]. The obtained information is used by the robot control system to generate control signals. Another important property of the developed classifier is the implementation of the learning process on the teacher mode, which allows the user to combine in one class different patterns. A model of iterative hetero-associative memory is proposed, on the basis of which the modified BAM network (ModBAM) is able to record, store and process sensory, information and control signal sequences [15]. In comparison with existing analogues, the proposed model of memory allows processing information data sequences, the length of which considerably exceeds the size of its inputs. It also has a more efficient mechanism for adding new information blocks [10 A] to memory. Based on this memory model, the following devices were proposed: a device for recording, storing and extracting binary sequences and a device for processing sensory data [15].

IV. Neural Network Model of Intelligent Hierarchical Control System of Autonomous Mobile Robots

Based on the methodology of the Pospelov's theory of natural and artificial systems behavior [16,17] and Zhdanov's method of autonomous adaptive control [6], the author proposed an original neural network model of the hierarchical ICS by autonomous mobile robots [18]. This ICS consists of conditional units of the processor and memory (Figure 2, filled blocks), which themselves consist of distributed sub-units implemented on the basis of hetero-associative ANNs. And thin arrows indicate information flows, and bold - control signals.

On the board of Controlled Object (CO) are the executive (Effectors - E) and sensory (Receptors - R) units and ICS, in the control loop which includes the operator unit. In this case, the object of control is AMR. As a rule, the Robot Operator must interact with the ICS only before the beginning of the operation of the AMR (at the time of entering the targets and the necessary initial data) or in critical situations. The ICS receives information from the external and internal environments by means of the Receptor unit and acts on the external environment by the Effector unit.

From the proposed scheme (Figure 2), it can be seen that the processor and memory units, like natural ICSs, have a distributed structure. One of the most important parts of the processor is the data preprocessing unit - the pre-processor, represented by the Pattern Formation and Recognition Units (PFR), in which information necessary for other units is extracted from the incoming data. The next part of the processor is represented by the decision making unit (DM), which controls the hierarchy of sequences of sub-goals execution.

In the proposed ICS model distributed memory is represented by a database (DB), a knowledge base (NB), a reflective unit (RE) and an Evaluation and Prediction unit (EP). The evaluation and prediction unit is responsible for cognitive functions of the ICS. The tasks of the database are not only the collection and transfer of sensory data in the Pattern Formation and Recognition unit, but also their registration (writing in parallel the values of the control signals from the effectors). This is required in order for the operator to perform a full analysis of the control decisions taken by the ICS after the AMR job is completed.

The sign $\Leftrightarrow$ indicates the process of bidirectional information exchange that occur in the 2nd and 3rd loops as a result of the iterative way of extracting information from the knowledge base and the feedback of the memory adjustment - the process of the after-training [18].

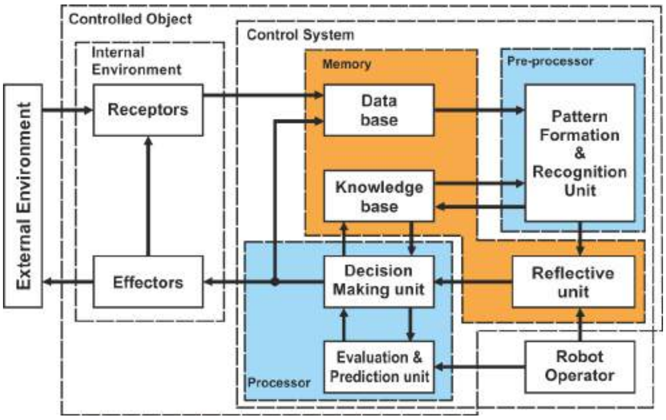


Figure 2. The main functional units of the proposed bio-inspired control systems.

Thus, three types of control loops are implemented in the proposed ICS:

1) $R \to PFR \to RE \to DM \to E$, - reactive;
2) $R \to PFR \to NB \Leftrightarrow DM \to E$, - extremum-seeking;
3) $R \to PFR \to NB \Leftrightarrow EP \to DM \to E$, - cognitive.

Unconditioned and conditioned reflexes involved in the first circuit of the proposed ICS are implemented in the RE unit in the form of two different hetero-associative ANNs, which, upon receipt of a certain stimulus, are attract a strictly fixed reaction. The reactive control loop can be implemented on the basis of the adaptive neural network controller proposed in [14], implemented as a system of rules. Thanks to the presence of hetero-associative ANNs, this controller allows to quickly extract the reference patterns from memory and process the input information, as well as add new ones to the memory.

Thus, unconditioned reflexes (i.e. "stimulus-response" learning pairs) should be recorded by the operator in the ANNs before the start of the AMR operation [14]. Conditioned reflexes should be remembered in the process of periodic external influences on the control object through the training of motoneurons ANN . This provides a mechanism for individual adaptation of the control object to small changes in the environment. Moreover, if the synaptic connections of the first ANN contain information on the critical values indicators of the energy and transport systems and must be strictly fixed, then the second - synaptic connections should have the ability to memorize the learning outcomes only for a finite time (ie, they can be retrained) . Output signals of the hetero-associative ANNs of unconditioned reflexes are passive – they affect only the control solutions of the PR unit. But signals from ANNs conditioned reflexes (motoneurons) are active, that is, they directly control the motion of the effectors. The second circuit (extremum-seeking control) is designed to implement a more complex type of behavior of the control object, which is able to recognize previously trained situations and respond to them with a corresponding sequence of actions of the effectors. In the extremum-seeking control loop, a control system must be implemented, which, with the help of experience-oriented actions, seeks to minimize certain objective functions of the control object. In other words, this circuit must implement the execution of some predefined search algorithms with branching [15].

V. Application Methodology of the Hetero-Associative Memory Technology

The neural network models, algorithms and methods proposed in this paper can find wide application in robotics and various branches of automated production. The developed model of a neural network hierarchical control system with distributed information processing [15,18] is used in the on-board control system of small-sized autonomous mobile robots. Based on the proposed neural network models and algorithms, software-hardware controls for group robots designed to solve research and research problems in conditions that are inaccessible to humans Thanks to the use and testing of the proposed neural network models for the management of mobile robots, we have gained enough practical experience. Based on this experience, TABLE 1 was created, which is intended to assist in choosing a management system to solve a particular practical problem.

Table I
Control System Selection

#	External Environ-ment	Search Algo-rithm	Control Loop	Teaching Mode	Knowledge Base
1	Simple or complex obstacles, known mapping	Random walking	Reactive	Manual control before start	Unconditi-oned reflexes
2	Simple obstacles, unknown mapping	Random walk / Objective search		Automatic / Interactive manual-automatic control	Conditi-oned reflexes
3	Complex obstacles, unknown mapping	Objective search	Extremum-seeking	Manual control before start / Interactive manual-automatic control	Branching algorithms

Table 2 shows what a positive effect the user should get when using the proposed neural network technology.

VI. Conclusions

Was realized an attempt to combine the technology of classical ANN and associative memory model based on a deterministic chaotic system into a single neural network model of memory, the ability to model not only partial manifestations of associative brain functions, but also some of the processes of exchange and processing of information, as well as the search for hidden links between existing standards in memory.

Among the main positive properties of the obtained memory model, we can distinguish the realization of information search

Table II
EFFICIENCY OF DEVELOPED TECHNOLOGY APPLICATION

#	Features	Structure	Entity	Outcome
1	Content-addressed storage	ART-BAM	Reduction of computing and energy costs	Increasing of decision making speed
2	Programming by example	ART-BAM / ModBAM	Does't require the specialist (programmer) participation	Increasing of programming speed
3	Patterns Adding / Removing	ART-BAM	Stimulus / response principles	Adaptive decision making
		ModBAM	Information block	Increasing of programming speed

by its content and storage of information in the form of structured binary sequences. It can also be noted that the process of retrieving information is a dynamic process, similar to processes occurring in various physical systems of continuous operation.

A modular neural network model of the intellectual hierarchical control system of autonomous mobile robots is developed, the modules of which are separate control loops: reactive, extremum-seeking and cognitive. On the one hand, this system has a strict hierarchical structure, consisting of three control loops. On the other hand, all calculations are performed in parallel and distributed in the corresponding functional units. The units are implemented using the proposed hetero-associative ANNs. At the moment, in the model environment, separate contours of the developed control system have been created and tested for various practical applications: the manipulator position control system, the neural network model for realizing the search motions of the autonomous mobile robot and the reactive contour of unconditioned and conditioned reflexes.

The next direction of our research will be integration of the proposed neural network models and semantic technologies of OSTIS. It is suggested that due to the universal method of representation of knowledge, OSTIS technology will effectively implement not only classical formalized knowledge bases, but also unformalized knowledge bases based on the ANNs.

Because OSTIS implemented on a remote server and can be accessed via the Internet, it is possible to control a group of autonomous mobile robots.

The work was carried out with partial financial support of the BRFFR F16P-146.

REFERENCES

[1] J. Buchli "Mobile Robots: Moving Intelligence," Pro Literatur Verlag, Germany / ARS, Austria, 2006. 586 p.

[2] S.J. Russell, P. Norvig "Artificial Intelligence: A Modern Approach, 3rd ed.," Prentice Hall, 2009. 1152 p.

[3] S.O. Haykin "Neural Networks and Learning Machines (3rd Edition) ," Prentice Hall, 2009, 906 p.

[4] M.A. Arbib "The Handbook of Brain Theory and Neural Networks, 2nd ed.,"The MIT Press, 2003. 1290 p.

[5] S. Omatu, M.B. Khalid, R. Yusof "Neuro-Control and its Applications. Advances in Industrial Control,“ Springer, 1996. 262 p.

[6] A.A. Zhdanov "Avtonomnyy iskusstvennyy intellekt [Autonomous Artificial Intelligence]; 2nd ed.," Moskva (Moscow) : BINOM. Laboratoriya znaniy, 2012. 359 p.

[7] S. Grossberg "Nonlinear neural networks: Principles, mechanisms, and architectures," Neural Networks, 1988, Vol. 1, pp. 17–61.

[8] G.A. Carpenter, S. Grossberg "The ART of adaptive pattern recognition by a self-organizing neural network," Computer, 1988, Vol. 21, pp. 77–88.

[9] G.A. Carpenter, S. Grossberg "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," Neural Networks, 1991, Vol. 4, pp. 759-771.

[10] G.A. Carpenter, S. Grossberg, J.H. Reynolds "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," Neural Networks, 1991, Vol. 4, pp. 565–588.

[11] B. Kosko "Competitive bidirectional associative memories," Frst Int. Conf. on Neural Networks, San Diego, CA, 1987, pp. 759-766.

[12] B. Kosko "Constructing an associative memory," Byte, 1987, Vol. 12, pp. 137-144.

[13] B. Kosko "Bi-directional associative memories," IEEE Transactions on Systems, Man and Cybernetics. 1988, Vol. 18, No 1, pp. 49-60.

[14] R. Prakapovich "Neyrosetevoy blok pamyati dlya adaptivnoy raboty slozhnykh tekhnicheskikh sistem v dinamicheskoy srede [Neural network memory unit for adaptive operation of complex technical systems in a dynamic environment]," Informatika (Informatics), 2010, Vol. 2(26), pp. 54-65.

[15] R. Prakapovich "Primeneniye geteroassotsiativnykh neyronnykh setey dlya zapisi i vosstanovleniya informatsii [Application of heteroassociative neural networks for recording and restoring information]," Informatika (Informatics), 2012, Vol. 2(34), pp. 38-49.

[16] D.A. Pospelov "Situatsionnoye upravleniye: teoriya i praktika [Situational control: theory and practice]," Maskva (Moscow), Nauka, 1986. 288 p.

[17] M.G. Gaaze-Rapoport, D.A. Pospelov "Ot ameby do robota: modeli povedeniya [From amoeba to robot: behavior patterns]," Maskva (Moscow), LIBROKOM, 2011. 296 p.

[18] R. Prakapovich "Neyrosetevaya model' dlya realizatsii poiskovykh dvizheniy mobil'nogo robota [Neural Network Model for Mobile Robot Search Movements]," Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2013, pp. 483–488.

ТЕХНОЛОГИЯ ГЕТЕРО-АССОЦИАТИВНОЙ ПАМЯТИ ДЛЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ УПРАВЛЕНИЯ АВТОНОМНЫМИ МОБИЛЬНЫМИ РОБОТАМИ

Прокопович Г.А.
Лаборатория робототехнических систем
ОИПИ НАН Белалруси
rprakapovich@robotics.by

Рассмотрены проблемы создания технологий искусственного интеллекта в задачах управления робототехническими аппаратами. На основе синтеза двух перспективных технологий – искусственных нейронных сетей и ассоциативной памяти – разработана нейросетевая модель итерационной гетеро-ассоциативной памяти, предназначенная записи, хранения и обработки сенсорной и управляющей информации. Предложена нейросетевая модель иерархической интеллектуальной системы управления робототехническими аппаратами, которая способна на основе имеющихся априорных знаний автономно функционировать и адаптироваться к изменениям внешних условий.

Semantics-Based Control of the Group of Intelligent Robots

Ivan Podmazov
Laboratory of Robotic Systems
United Institute of Informatics Problems of the NAS of Belarus
Minsk, Belarus
podmazov@gmail.com

***Abstract*—In this paper we offer architecture of a software part of an intelligent robot group designed for work in complex dynamic enviroment conditions. The approach is based on application of the Open Semantic Technology for Intelligent Systems (OSTIS). That will allow to accelerate the intelligent control system development process, at the same time preserving its capability and quality thanks to the advantages of the architecture and the development sequence. Our goal is to construct a group of robots with limited cognitive abilities that are able to operate knowledge in formalized form, to deduce new knowledge and also to have an access to existing external intelligent systems and knowledge bases. The intelligent robot group software is built hierarchically: overlying components are implemented on the basis of underlying ones. The hierarchy includes three abstraction layers: programming language extension level, intelligent system level, robot group control level. It seems that control system built in accordance with the suggested architecture possesses the qualities of versatility and extensibility, and that will provide the breadth of supported functionality.**

***Keywords*—intelligent system, robotics, robot group, control system, semantic technology, OSTIS**

I. Introduction

Agent group control in a complex unpredictable environment is an actively explored and developing field of artificial intelligence. A concept of agent in the group artificial intellect as an independent science is abstract since the greatest interest of the latter is the properties related to interactions and not the properties of active units. Nevertheless, individual properties are important and can influence significantly on the efficiency of the whole group. It is known that increase in agent's cognitive abilities leads to complication and enrichment of behavioral diversity improving by that its adaptive capability in current conditions of rapidly changing environment [1], [2]. For instance, in nature species possessing more developed brain have an evolutional advantage and take a higher place in the food chain. Eventually, when a species gets the ability to think i.e. to find hidden, directly unrecognizable regularities of the world's phenomena, its possibilities come to a new level [3]. The most obvious example is homo sapiens, a species not distinguished by physical strength and dexterity learned to subdue the forces of nature and to use them for own purposes through its intellect, and thereby became dominant in the scales of the Earth's biosphere. In practical applications, real control objects are identified with the agents. Particularly, in robotics such objects are group robots.

Invention of artificial "servants" for the sake of elimination of the need for manual labor is the one of the human's oldest dreams. The robotization and automatization processes originate from the earlier and more common process of construction of artificial tools with the required properties. This process have led to the appearance of the devices that "boost" the human body physical qualities, allowing to carry out working operations faster and more precisely. Total rejection of the human involvement in production and service cycle is simultaneously the goal and the outcome of automatization and robotization [4].

Works have been started to create "thinking" devices that could not just become a high-grade substitute of a human workers, but also to be networked into much more effective and powerful systems, though the significant difficulties showed up. It turned out that only the formalizable, simply algorythmized procedures can be easily automated, while the automatization of unformalizable procedures face some difficult problems. In despite of series of considerable and useful results obtained in this scope, a human is without a rival in solving creative and other non-formalizable problems at the moment. [5], [6]. Thereby presently the basic practical robotics problem is to construct collectives of robots with limited, but sufficient cognitive abilities for specific purposes such as formalized knowledge processing (including inference of new knowledge) and access to vast array of information accumulated by humanity etc.

II. Robot group architecture

This article reviews the problem of construction of a versatile intelligent control system for a group of intelligent robots acting in complex dynamic environment. Schematic representation of the group architecture is shown in the fig. 1.

The group consists of N in general different robots, each of them contains:

- general-purpose computing unit (controller) – blocks RS, KB, CS;
- wireless communication module – block T;
- sensor set (camera, microphone, magnetometer, obstacle sensors etc.) – block S;
- actuator set (chassis, manipulators etc.) – block A.

Except robots, the group includes a server used for centralized synchronization of the knowledge bases and for inter-

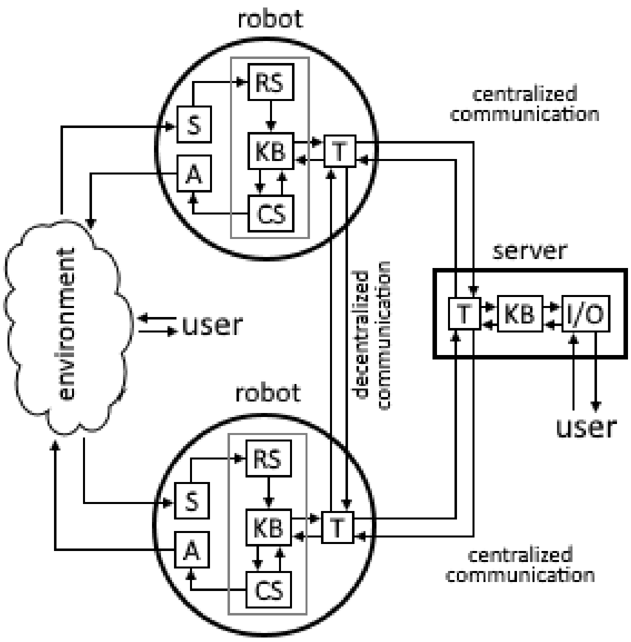


Figure 1. Scheme of a group (collective) of intelligent robots (S = sensors, RS = recognition system, KB = knowledge base, CS = control system, A = actuators, T = transmitter, I/O = input-output system).

action with the user. Interaction with the user is carried out through the intelligent input-output system (block I/O).

The robots and server run software with all functionality required for semantic calculations and processing of knowledge in formalized form. Each robot is able to recognize events and contexts in an input data stream (from sensors, block S) and generate its semantic description in a certain universal fact-presentation language (by means of the recognition system, block RS). Besides, the robot accepts individual and group commands which are extracted from sensor data the same way. Newly generated descriptions are put in a local knowledge base (block KB), maintaining its consistency. In fact, this base represents a semantic map of environment that can be used to complete the task whether by a robot itself or by some group subset. In order to synchronize the base, its fragments are distributed within the group on request – directly or through the server containing the global database (block KB). Robot control is carried out directly by the intelligent control system (block CS), which does the semantic analysis of the current situation and makes a decision on further actions in obedience to the user input (hereafter carried out by actuators, block A).

To give an example of the use of such a group, we consider the problem of mapping of premises performed by a group of robots that are able to answer questions like:

- how many floors there are in the building, how many rooms there are on the floor, how many doors there are in the room etc.;
- how many robots in the group are active, where are they located;
- what does a particular robot see or hear, etc.;

as well as are able to carry out assignments like:

- to find a certain object or person (to set location);
- to get, to deliver an object or to escort a person to a given point;
- to identify/modify an object state (for example, to find out if all the doors on the floor are closed) etc.

Besides, the input may be supplied vocally to all group members. Each request requires recognition of a complex verbal command, knowledge database operations, semantic calculations and interaction with the other robots. The Laboratory of Robotic Systems of the United Institute of Informatics Problems has already got some practical results in the mentioned direction: a computer vision system for calculating a point of perspective on picture frame received from onboard camera has been designed. That lets a robot to move on a sidewalk, in the corridors of the building and even to find the right door [7].

Thus, the problem of construction of a control system for a group of robots that corresponds to the specified description, involves the solution of the following subtasks:

1) implementation of the recognition system (block RS);
2) developing of the system of knowledge processing, synchronization and storage (blocks KB, CS);
3) developing of the "human-robot" interface that supports the use of anthropomorphic language (block I/O).

It is supposed that each of them can be solved by using and integration of existing intelligent and other program systems (which will require implementation of specialized program interfaces).

III. Software architecture

The intelligent robot software is built in accordance with the principles of the OSTIS (Open Semantic Technology for Intelligent Systems) [8]. The technology is called "semantic" since its essence is in applying semantic networks of special unified kind for the representation of any structured information and constructing of special information processing systems. Thereby a number of problems is planned to be solved, among which one can mention simplification of the integration of different information representation models, shortening of intelligent system development terms and complexity as well as improvement of its quality.

According to the OSTIS principles, the intelligent robot group software of the considered type is supposed to be hierarchical as shown on the fig. 2.

The hierarchy consists of three levels of abstraction:

1) programming language extension level;
2) intelligent system level;
3) robot group control level.

First two levels are universal and are present in one or another form in every intelligent system implemented in accordance with the OSTIS principles. The third level is a quintessence of intelligent system and corresponds some

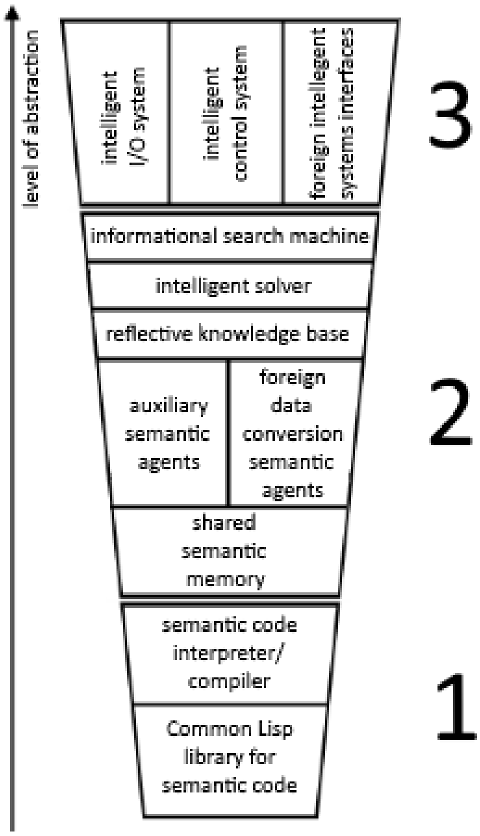


Figure 2. Structure of the robot group software.

particular problem, contains means and description of the considered problem solution.

The first abstraction level contains means that expand the programming language features concerning semantic networks processing. The basis of this is the library implementing a domain-specific language (DSL) in which all overlying layers are written. It is known [9] that one of the most convenient languages for DSL implementation is Common Lisp. Due to the powerful macro system allowing to transform language syntax drastically, introducing by that the means that make the code systematic, brief and laconic, speed and quality of development increase at times in comparison with popular general-purpose programming languages [10]. Above the specified library lies a meta application intended for semantic code execution by means of interpretation or compilation and its subsequent launch. In both cases the output is a proper specialized program containing instructions for semantic agents of a certain type that is executed as a separate process in an operating system.

The second level of abstraction represents an implementation of the intelligent system basis – sc-machine, supplemented by a knowledge base that contains information about the system organization and usage. The fig. 3 shows a general scheme of internal organization interactions of the semantic machine.

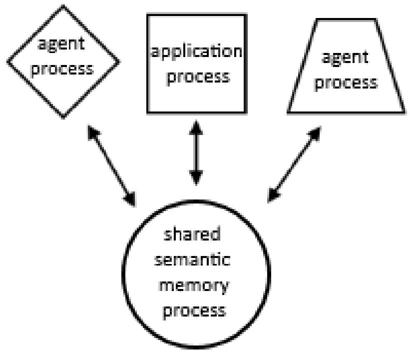


Figure 3. Robot's multiagent system of a semantic machine (sc-machine).

Sc-machine consists of a multiplicity of agents (operating system processes) of different complexity and functionality that interact with the shared semantic memory (which is also a process). All inter-agent interactions are conducted exclusively through the shared memory. High-level intelligent subsystems – intelligent solver, information retrieval machine and user interface core are built on this basis. In order to render interfacing of semantic technology with the outer systems possible, there are agents that perform convertion of semantic networks to external formats and vice versa. The knowledge base which is a part of the second level contains detailed description of the whole system including basic concepts and their interrelation. Besides, the base includes information on rights, access to resources and interaction of all the entities participating in system operation, both program agents and users.

On the third, robot group control level, there is a complex of means that use the previous levels functionality to solve the assigned problem of robot group control. This complex includes an intelligent user interface (which contains server web-interface as well as robot's voice command recognition subsystem), the control system actually and the collection of interfaces for external intelligent and other program systems. The latter is intended to exploit external systems in order to extract missing information that can not be obtained in the system (for example, when a user requests a fact from external sources), as well as to have an impact on them. Intelligent group control system performs parsing of input which is presented as semantic network, calculates the expected behaviour and ensures its accomplishment by outputting the response through the user interface either by sending appropriate signals to actuators which lead to the desired result. As a matter of fact, the control system does not contain any hard-coded parts, since the whole functionality is provided through well-established interaction of underlying subsystems – the information retrieval machine and the prearranged knowledge base.

IV. ELEMENTS OF SEMANTICS-BASED CONTROL

Semantics-based control of the groups of intelligent robots is understood as the transfer of the maximum propotion of computation to an informational search machine and an intelligent solver from of the hard-coded algorithms. Group interaction and control are provided by two factors: the availability of knowledge base synchronization and the possibility to utilize other robots' resources in order to solve the specific robot problem through the support of corresponding terminal elements and rules. Synchronization is the exchange of information between robots. The information presented in the form of a semantic network, the volume and content of which is determined by the interaction graph corresponding to the task (it may be raw/processed sensor data, calculations results, etc). Received information updates the state of the robot and alters its behavior. Deduction of the robot behavior is carried out by informational search machine and intelligent solver on the basis of logic deriving process. This process uses the acquired facts about the state of the environment as well as the prearranged rules of behavior from the knowledge base.

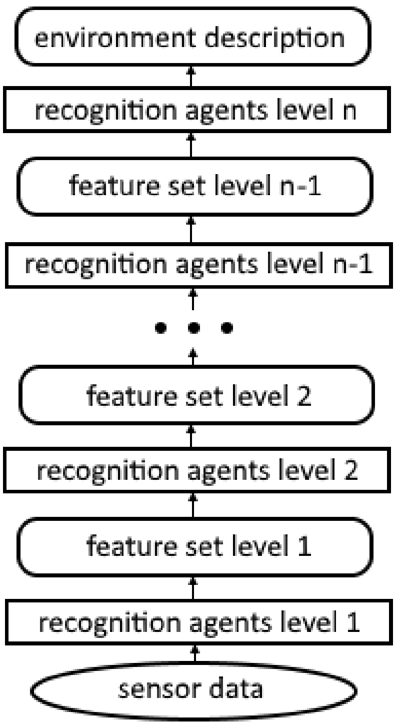


Figure 4. Recognition agents hierarchy and input data conversion sequence.

The specified functionality is implemented through the availability of agents of the following types.

- General-purpose (universal) agents provide functioning of the entire system as a whole, including auxiliary transformations and algorithms that are inherent in any intelligent system built according to OSTIS principles. They also include data representation transformation agents and all agents that are not specific to the robot group control system.
- Agents of recognition of events and patterns in the input data stream, as well as the agents for parsing verbal input form a hierarchical subsystem consisting of several levels. Each next level takes results of the previous level as the input; the first level performs the primary processing of data from the sensors (fig. 4). Such a structure allows to significantly compress the input stream by discarding non-essential information and forming a compact semantic description that can later be transmitted to the rest of the robots and "understood" by them without any additional calculation (example: frame from the on-board camera => number, types, orientation of the objects captured). Combining the descriptions of the environment at different times and from different sources makes it possible to construct a semantic map of the environment.
- Synchronization agents of knowledge bases in the group select, send, receive and integrate semantic information. The mode of operation of this type of agent is not fixed and is also subject to adjustment by the joint activities organizing mechanism for a group of robots.
- Agents for output generation and actuator control, like agents for input data recognition, form a hierarchical system that converts a high-level description of the action decisions made into hardware implementation level commands (for example: "output required" => sound/text, "motion required" => signals fed to the drives).

Simplified algorithm for the operation of a robot in semantics-controlled group is as follows:

1) collect data from sensors and generate their semantic descriptions, add it to the knowledge base;
2) receive semantic information from the other robots and the user and add it to the knowledge base;
3) revise (repeat derivation of description) of the current behavior as a result of the running of informational search machine and intelligent solver over the updated knowledge base;
4) transfer the required semantic information to other robots and the user;
5) execute the accepted behavior.

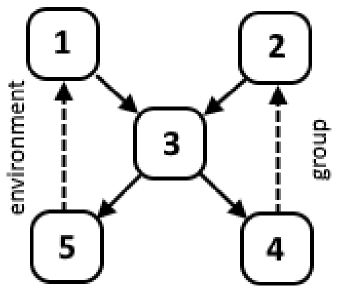


Figure 5. Dependence graph for the robot functioning algorithm (solid arrow – direct internal dependence, dashed arrow – existence of external feedback, feedback channel is noted nearby)

Figure 5 depicts the dependency graph of the algorithm steps. It is understood that the steps are performed asynchronously. It is guaranteed that the integrity of the database and the control system is always maintained and can not be violated by impact on the robot through the provided communication channels.

Thus, the structure of the system is flexible and allows to make changes to it with minimal costs, even in on-line mode. If there are agents that monitor the correctness of the state of the knowledge base, accidentally making an error becomes less likely. On the grounds of the taken architectural solutions, it seems that the intelligent control system for a group of robots has a significant potential regarding expansion the provided functionality.

We suppose that the OSTIS has significant power and is able to offer a solution that could be a leap forward over the existing ones in various fields of computer science. In particular, it can help in improving of intelligent technologies of the group robotics on account of integration with the third-party intelligent technologies.

V. Conclusion

In this paper we offer architecture of a software part of an intelligent robot group designed for work in complex dynamic enviroment conditions. The approach is based on application of the Open Semantic Technology for Intelligent Systems (OSTIS). That will allow to accelerate the intelligent control system development process, at the same time preserving its capability and quality thanks to the advantages of the architecture and the development sequence.

Our goal is to construct a group of robots with limited cognitive abilities that are designed without taking an example from animals or humans, since we desist from attempts of solving non-formalizable problems. The robots must be able to operate knowledge in formalized form, to deduce new knowledge and also to have an access to existing external intelligent systems and knowledge bases.

The considered group consists of a server and a few in general different robots, each of which contains a general-purpose computing unit and wireless communication module that is necessary for network connection, besides sensors and actuators. The server targets two purposes: centralized synchronization of robot knowledge bases and providing a platform on which the user web-interface is run. Besides the server, interaction with user can be directly effected verbally with robots (verbal command/questions – verbal answer).

The intelligent robot group software is built hierarchically: overlying components are implemented on the basis of underlying ones. The hierarchy includes three abstraction layers: programming language extension level, intelligent system level, robot group control level. The first abstraction level contains means that expand the programming language features concerning semantic networks processing: domain-specific language and semantic code interpreter/compiler. The second level of abstraction represents an implementation of the intelligent system basis – sc-machine, supplemented by a knowledge base that contains information about the system organization and usage, as well as intelligent solver and information retrieval machine built on this basis. The third level includes an intelligent user interface, a collection of interfaces for external intelligent and other program systems and actually a robot group control system. On the third, robot group control level, there is a complex of means that use the previous levels functionality to solve the assigned problem of robot group control. The third level includes an intelligent user interface, the control system actually and the collection of interfaces for external intelligent and other program systems.

Semantics-based control of the groups of intelligent robots is understood as the transfer of the maximum propotion of computation to an informational search machine and an intelligent solver from of the hard-coded algorithms. Group interaction and control are provided by two factors: the availability of knowledge base synchronization and the possibility to utilize other robots' resources in order to solve the specific robot problem through the support of corresponding terminal elements and rules.

It seems that control system built in accordance with the suggested architecture possesses the qualities of versatility and extensibility, and that will provide the breadth of supported functionality.

Acknowledgment

The author would like to thank Natalia Podmazova for help in translating this article into English.

References

[1] I. A. Kalyaev, A. R. Gaiduk, and S. G. Kapustyan, *Modeli i algoritmy kollektivnogo upravleniya v gruppakh robotov*. Moscow: Fizmatlit, 2009.

[2] G. A. Prokopovich, "Bionicheskaya struktura ierarkhicheskoi raspredelennoi sistemy upravleniya avtonomnymi mobil'nymi robotami," in *Iskusstvennyi intellekt*, vol. 1, 2013, pp. 181–190.

[3] A. A. Zhdanov, *Avtonomnyi iskusstvennyi intellekt*, ser. Adaptivnye i intellektual'nye sistemy. Moscow: Binom, 2012.

[4] D. Acemoglu and P. Restrepo, "Robots and jobs: Evidence from us labor markets," National Bureau of Economic Research, Working Paper 23285, March 2017. [Online]. Available: http://www.nber.org/papers/w23285

[5] J. Searle, "Minds, brains, and programs," *Behavioral and brain sciences*, vol. 1, no. 3, pp. 417–457, September 1980.

[6] J. McCarthy, "From here to human-level AI," *Artificial Intelligence*, vol. 171, no. 18, pp. 1174–1182, 2007, special Review Issue. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0004370207001476

[7] G. A. Prokopovich, "Razrabotka sistemy tekhnicheskogo zreniya dlya servisnogo mobil'nogo robota," in *III Vserossiiskii nauchno-prakticheskii seminar «Bespilotnye transportnye sredstva s elementami iskusstvennogo intellekta»: trudy seminara*. Innopolis, Tatarstan Republic, Russia: Pero, 2016, pp. 127–136.

[8] V. V. Golenkov and N. A. Guliakina, "Graphodynamical models of parallel knowledge processing," in *Proceedings of the II International Scientific and Technical Conference*, Minsk, 2012.

[9] P. Graham, *On LISP: Advanced Techniques for Common LISP*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[10] I. V. Podmazov, "Vybor yazyka programmirovaniya dlya realizatsii prototipa programmnogo interpretatora sc-mashiny," in *XIV Mezhdunarodnaya nauchnaya konferentsiya «Molodezh' v nauke – 2017»*. Minsk: in press, 2017.

УПРАВЛЕНИЕ ГРУППОЙ ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОВ НА ОСНОВЕ СЕМАНТИЧЕСКОЙ ТЕХНОЛОГИИ

Иван Подмазов
Лаборатория робототехнических систем
Объединённый институт проблем информатики НАН Беларуси
Минск, Беларусь
podmazov@gmail.com

В данной статье предлагается архитектура программной части группы интеллектуальных роботов, предназначенной для работы в условиях сложной динамической среды. Подход основывается на применении открытой семантической технологии проектирования интеллектуальных систем (OSTIS). Это позволит ускорить процесс разработки, но при этом не уступить в мощи либо качестве благодаря преимуществам архитектуры и последовательности построения интеллектуальной системы управления группой. Нашей целью является создание группы роботов с ограниченными когнитивными способностями, способных оперировать знаниями, представленными в формализованном виде, выводить новые знания, а также иметь доступ к существующим внешним интеллектуальным системам и базам знаний. Представляется, что система управления, построенная согласно предложенной иерархической архитектуре, будет обладать качествами гибкости и расширяемости, что позволит обеспечить широту поддерживаемого функционала.

Design of batch manufacturing enterprises in the context of Industry 4.0

Valery Taberko
and Dzmitry Ivaniuk
and Valery Kasyanik
JSC «Savushkin Product»
Brest, Republic of Belarus
id@pda.savushkin.by,
val@pda.savushkin.by

Vladimir Golovko
Brest State Technical University
Brest, Republic of Belarus
gva@bstu.by

Natalia Guliakina
and Kirill Rusetski
and Daniil Shunkevich
and Alexander Boriskin
and Natalia Grakova
Belarusian State University
of Informatics and Radioelectronics
Minsk, Republic of Belarus
{guliakina,rusetski,grakova}@bsuir.by,
shunkevichdv@gmail.com,
coloss_000@mail.ru

***Abstract*—This paper presents an evolution of an ontology-based approach to designing batch manufacturing enterprises. According to Industry 4.0 approach, instead of isolated view of a manufacturing process inside a single enterprise this new approach encompasses related business entities as well - raw material suppliers (e.g. dairy farms) and large-scale consumers (e.g. stores or retail chains). Special attention is paid to logistics processes: a short description of fundamental logistics processes of cottage cheese production is provided, as well as subject domain structure of logistics and an example of formal specification of emergency logistics situation. It is shown that multiagent industrial control system with agents interacting via shared memory is compliant with design principles of Industry 4.0 approach. Standards formalization topic is touched upon as well. PFC, a graphical procedural model specification language, formalization is discussed. PFC is specified in ISA-88.02 standard. Graphical language formalization allows industrial control system users to communicate to it in a unified manner using diagrams that are widely understood by engineering specialists. This paper also outlines an agent-oriented approach to robot interaction within industrial robotic complexes based on shared semantic memory interaction mechanism.**

***Keywords*—integrated industrial control automation, logistics process, ontology-based enterprise model, Industry 4.0, cyber-physical system, ontology, knowledge base, OSTIS technology.**

I. Introduction

This article considers the further automation development of the batch manufacturing enterprises by the example of JSC "Savushkin Product". It consists in the transition from consideration of manufacturing processes, occurring within a specific enterprise, to the consideration of the full product cycle – from a store application to its execution (satisfaction of the consumer's request). This process, apart from the enterprise, as subjects include dairies, logistic services, shops and others. Also, even within the framework of one enterprise, disparate systems are used today (several SCADA systems, a transport management system, a warehouse management system, etc.). Coupling of such systems and maintenance of information consistency in them is carried out manually (or omitted). Now there is a need to automate information coordination and to provide, where appropriate, the interaction of devices at different stages of the product cycle. The existence of such requirement has led to the forthcoming Industry 4.0 in Germany and its analogs in other countries.

The objective of this paper is therefore an evolution of ontology-based model of batch manufacturing enterprise which widens the scope of production stages to include production processes that are happening outside of an enterprise, in accordance with the Industry 4.0 initiative. External logistics processes, such as milk delivery from dairy farm to dairy factory, or final product delivery to customers, can be used as examples.

A. *Industry 4.0 and cyberphysical systems*

The concept Industry 4.0 was formulated in Germany in 2011. It means creating and implementing production-ready cyberphysical systems, as well as using Internet of Things (IoT) and Internet of Services (IoS) in manufacturing processes[1]. Note that this term is specific to Germany, and is rarely used outside. There are concepts similar to Industry 4.0 in other countries: Smart factory in Netherlands, Future Factory in Spain, Industrial Internet (of Things) in the USA.

Industry 4.0 design principles are outlined below [2]:

- **Interoperability.** CPS and humans are connected over the Internet of Things (IoT), Internet of People (IoP) and the Internet of Services (IoS).
- **Virtualization.** CPS can monitor physical processes. Sensor data are linked to virtual plant models, which include the current state of all CPS. In case of failure a human can be notified and provided with all necessary information and, hereby, supported in handling the decision-making.
- **Decentralization.** The rising demand for small batches of custom-ordered products makes it increasingly difficult to control manufacturing systems centrally. CPS can have

computers embedded in them to enable them to make decisions on their own, and only delegate tasks to a higher-level equipment in case of failure. Nevertheless, for quality assurance and traceability it is necessary to have centralized control of the system. For example, RFID tags can “tell” machines which working steps are necessary, thus eliminating the need for the centralized control of this part of the small or individual batch production process.
- **Real-Time Capability.** For organizational tasks it is necessary that sensor data is collected and analyzed in realtime. In case of an equipment failure its task can be rerouted to another piece of equipment.
- **Service orientation.** The services of companies, CPS, and humans are available over the IoS and can be utilized by other participants. Services can be offered both inside and outside the company. CPS can offer their functionality, for example, as a set of web services. It allows for composition of production process from smaller operations according to a customer specification encoded on an RFID tags, for example.
- **Modularity.** CPS have to be flexible to easily adapt to changing requirements (e.g., seasonal fluctuations or changed product or production environment characteristics) Adaptation can be done by replacing or expanding individual modules of the system. Module compatibility requires standardized software and hardware interfaces, so that new modules are identified automatically and can be utilized immediately via the IoS.

CPS is a collection of intelligent, easily integrated physical components with built-in computational resources that closely cooperate and monitor changes in their environment [3].

To build a CPS it is necessary to integrate computational resources and technical processes. Sensors, manipulators, information and control systems should interoperate at all stages of the production including those outside of a particular factory [4], [5], [6], [7], [8]. It needs to be said, though, that implementation of a new industrial control system, including intelligent ones, should be based on resources that the enterprise already possesses [9].

Need for integrated automation of complex processes that require coordinated interaction of multiple services and technical equipment drives creation of such systems. From now on we will discuss enhancement of the industrial control level of the batch manufacturing enterprise in terms of designing a CPS responsible for producing cottage cheese "Khutorok" at JSC "Savushkin product".

Such CPS should offer informational support and industrial control automation throughout the entire cottage cheese "Khutorok" production process – from milking a cow to delivering final product to the store. This process can be divided into the following main stages:

1) gathering milk at the farm;
2) milk delivery from the dairy farm to the dairy factory;
3) milk processing at the factory, cottage cheese production;
4) forming and packing;
5) final product delivery to the customer (shops, retail chains, etc.).

B. Problems in integrated enterprise control systems development and the proposed approach to solving them

Main problem with integrated enterprise control systems development lies in integration of its various components and facilitation of their interoperability. It can be solved the traditional way, by developing communication layers between heterogenous components of the system (interfaces, protocols, etc.). On the one hand it leads to considerable overhead required to develop them, on the other hand it complicates system architecture, which leads to increased costs of its maintenance and further development. Continuous evolution of production technologies at various stages and expansion of production itself requires industrial control system to be flexible, i.e. able to be easily extended by various components. Existing components should be modifiable when possible or required. To solve this problem we propose to extend the original ontology-based approach to design of batch manufacturing enterprises [10]. Enterprise is viewed as an integrated multiagent system, within which:

- all information is integrated within unified informational space (enterprise knowledge base stored in the semantic memory);
- all participants (people, robots, various integrated production systems etc.) are interpreted as agents that are working with this shared knowledge base; It means that (a) they are monitoring the knowledge base for the situations they can handle and (b) they specify results of their work in the knowledge base, so that this information is available for other agents to analyze. This approach reduces production process management to proper specifications of tasks in the shared knowledge base – with time-frames, priorities, assignees, etc.
- knowledge base has hierarchical structure, i.e. is a hierarchy of subject domains and corresponding ontologies.
- multiagent system is a hierarchical system in itself - agents can form infinitely nested collectives, since particular collective as a whole can be a member of another collective. For example, a group of robotic systems can be logically (or even physically) joined to form an integrated robotics system which can solve certain class of problems.

Multiagent system over a shared knowledge base implemented using ***OSTIS technology*** is therefore Industry 4.0-compliant and can be interpreted as a CPS:

- **shared knowledge base** implements **interoperability** of people, sensors, and equipment, serving as an intermediary of an interoperation, **virtualization** – knowledge base contains a representation of an enterprise model and environment with necessary level of details, **service orientation** – every participant of a production process (agent) is specified within the shared knowledge base

including its functionality (services provided), **modularity** – it hosts a library of reusable and interoperable components.
- **multiagent approach** implements principles of **decentralization** by the definition of a multiagent system [11], **realtime analysis and reaction** – agents monitor enterprise knowledge base state and activate in response to certain situations (including emergencies).

This approach offers several advantages, such as:
- there is no need to develop interoperability tools for system components (human-robotic interaction, human-human interaction tools, etc.) due to their interaction via shared memory
- since all agents interact via shared memory, in general, physical implementation of the agent does not matter to the system. Therefore, gradual replacement of manual labor by automated systems or the improvement of such systems does not require changes to the industrial control system in general;
- due to the use of a shared knowledge base and associative search in such knowledge base, any production process participant has access to all the information at any time, as needed, not only to a limited number of predetermined fragments, increasing the number of which may incur additional overhead costs. Thus, monitoring various processes becomes easier, and the answers to user questions can be found faster. User requests can be elaborated in numerous ways;
- information stored in the knowledge base can be rendered differently for various categories of users, and while the information itself remains unchanged, only the mechanism of displaying it will change. Therefore, there is no need to duplicate information;
- since all production processes are specified and managed via the knowledge base, making changes to such processes generally boils down to making changes in the knowledge base and replacing the corresponding equipment, if necessary. At the same time, the overhead costs for reprogramming the components of the system, and for facilitating the interaction between them, and for facilitating the interaction between them are substantially reduced;
- specification of all production processes in a shared knowledge base provides diverse options for their automatic analysis, including continuous monitoring of current processes, automatic detection and elimination of emergency situations, optimization of current processes, automatic planning of future processes, etc.

C. Architecture of the proposed system

Proposed system is based on the ***OSTIS Technology*** and according to it consists of a knowledge base, knowledge processing machine, and user interface In general, industrial control system knowledge base consists of [10]:
- ontologies of the industry-specific standards, such as ISA-88 [12]
- enterprise models based on these ontologies (e.g., physical, procedural and process models for ISA88)
- enterprise improvement ontologies, that formalize principles of improving and adapting an enterprise to changing conditions
- tools of collaborative development of enterprise knowledge bases and knowledge processing machines
- industrial control systems user interface models
- information service model for various user classes
- enterprise knowledge representation models that allow to specify it in all of the necessary aspects:
  - enterprise knowledge management model [13];
  - ontology-based enterprise model [14];
  - multiagent enterprise model [11];
  - enterprise situational control model [15];
  - business process re-engineering model [16].

Previous paper [10] discussed the formalization of standards, in particular ISA-88, in the form of a family of ontologies. Several fragments of enterprise models that were formalized using these ontologies, were shown. This paper also touches upon the ISA-88 standard, but focuses on the enterprise procedural model specification language – PFC, which is described in Part 2, Chapter 6 of this document [17]. In addition, the article shows evolution of an ontology-based enterprise model for a formal description of processes occurring outside the enterprise, in particular, logistical ones. Enterprise CPS knowledge base should contain following models, among others, to adequately describe manufacturing process:
- models of process cell description languages
- process cell models described using these languages
- logistics process models
- enterprise robot interaction model

The model of logistics processes is necessary at all stages of production to describe the internal (inter-shop logistics, warehouse logistics) and external logistics processes (cooperation with milk supply farms and retail partners). Models of process cell description languages and models of cells described with them are used at the third and fourth stages mentioned in the introduction, to formalize the production process and the structure of the equipment used for this. The interaction model of industrial robots specifies the physical model of the enterprise. The concept of the robot has wider interpretation in this context, which includes equipment modules and their complexes that perform their tasks with minimal human intervention, if any.

II. Implementing models, that are used to describe the enterprise and its processes

A. Language model for process cell description languages

For the convenience of the enterprise personnel operation with the system of complex automation, it is necessary to ensure, on the one hand, the ability of the system to interact with users in convenient ways (including using various graphic languages, limited natural language and voice messages), on the other hand – to provide the possibility to add new language means to the system, for example, new graphic languages.

Each intelligent system operates with a knowledge base in the internal language, and the dialog is implemented as a message exchange between the user and the system. To facilitate such dialog, it is necessary to convert certain knowledge base fragments into their external representation. Such representation can either be universal or specialized.

The universal external language for message exchange we will call the external language for message exchange, which allows to describe knowledge of any kind. Such is Semantic Computer Code (SC-code) and all its representations.

Translation from the internal language to external and back is organized in such a way that the translation mechanisms do not depend on the external language. In order to implement a new specialized language, in this case it will be necessary only to describe its syntax and semantics, while the universal translation model will not depend on this description.

Every language is characterized using three primary aspects. Each one of these aspects is described in the corresponding ontology.

- *Ontology of language semantics*
- *Ontology of language texts*
- *Ontology of language rules*

Language semantics ontology implies the choice of a set of uniquely defined entities that are understandable at the associative level and which carry a certain meaning.

Language texts ontology researches the syntactic structures that are images, symbolic representations of the language entities. Number of these kinds of images (symbols) is not limited and depends on the context being used.

The ontology of language rules is directed to the consideration of the rules of the language specifies unambiguous correspondences between the set of entities (the alphabet of the language) and the set of images (file signs) used to translate texts into and out of the intelligent system memory and also to visualize these texts.

The mechanism of translation is provided due to the presence in the system of a set of receptor and effector agents [18] in the mode of permanent exchange of messages between the user and the system. This message exchange mechanism is as follows:

1) The user writes some information with the editor of one of the specialized external languages.
2) Receptor agents fix the fact of the translation start of the written syntactic structure.
3) Internal agents use a set of rules to transform the syntactic structure into a sequence of sc-elements that constitute a fragment of a semantically connected sc-text that is unambiguously interpreted in the system's memory.
4) If the user makes changes in the resulting sc-text, then the reverse process occurs: the correspondence between the entity signs and their images is established, as a result of which the syntactic structure in the selected language is displayed.

The semantics of any language implies the introduction of a set of strictly defined entities sufficient for writing texts that represent a sense for the user or machine (system). Texts of a language are understood as syntactic structures that are images of the language entities. Finally, the rules of the language specify unambiguous correspondences between the set of entities (the alphabet of the language) and the set of images (file signs) used to translate texts into and out the intelligent system and to visualize these texts.

In the process of each language description, it is possible to identify certain aspects that are common for all languages or a particular family of languages. Research in this area are aimed at justifying a certain metalanguage, which defines the structure for describing the majority of existing languages. This meta-language will give impetus to the development of natural-language interfaces and will allow to introduce algorithmic precision into the linguistic aspects of any language.

The technology of cottage cheese production can be described in accordance with the ISA-88 standard. In the context of automated production and the ISA-88 standard, the following specialized external languages are distinguished: the procedural model description language (PFC) and the physical model description language (P&ID). With the use of the PFC language, a fragment of the production cell for the production of "Khutorok" cottage cheese will be described, which will be considered below.

The PFC language is defined in Chapter 6 of ISA-88.00.02 and is intended to describe recipes with complex procedures, involving parallel steps and conditional branching. PFC diagrams represent procedural logic using a set of icons connected by directional connections indicating the order in which procedural elements are executed.

The alphabet of the PFC language includes the following elements:

- procedural elements – elements of a procedural hierarchy (phases, operations, etc.);
- additional elements – elements responsible for allocation, synchronization and transfer of the resources within procedure (allocation element, synchronization element, etc.);
- structures represented in the form of classes of temporal entities that specify order.

B. Model of production cell of cottage cheese "Khutorok"

As an example of the procedural model use, a production cell of the "Khutorok" cottage cheese produced by "Savushkin Product" enterprise will be used. This cell reflects the stage of the milk processing in the plant using the example of a specific product manufacturing. The structure of the "Khutorki" project is presented on Listing 1:

Project "Khutorki"
=> *inclusion**:
    master recipe
    => *inclusion**:
        • *recipe procedure*

- *equipment procedure*
  => *inclusion**:
    - *"Milk whey separation" operation*
    - *"Milk whey pumpdown" operation*

Listing 1. Procedural hierarchy specification for "Khutorki" project

Master recipe describes the process from processing milk mixture to curd mass packing. The recipe and equipment procedures focus on the production of curd mass as is the master recipe. Finally, hardware procedure focuses on the operation of separating and pumping out the milk whey.

Fig. 1 shows the PFC representation of the procedure (PFC alphabet was discussed earlier).

Procedural model fragment that corresponds to the PFC chart in Fig. 1 is shown in Fig. 2. Entities and relations used in this structure will be explained in listings 2-3 and accompanying text.

PFC element
=> *inclusion**:
- *structural element*
- *procedural element*

Listing 2. PFC element classification

Structural element is a PFC element which in conjunction with several procedural

Procedural element is an element of procedural hierarchy which includes phases and operations.

execution order*
<= *subdividing**:
{
- *implicit transition**
- *explicit transition**
}

Listing 3. PFC transition element classification

Implicit transition* is a binary relation, the first component of which is a procedural element, after which execution of the procedural element, which is the second component, will begin.

Explicit transition* is a binary relation, the first component of which is a procedural element, after which execution of the procedural element, which is the second component, will begin, after certain condition evaluates to True.

Transition condition* is a binary relation, the first component of which is the instance of an explicit transition* relation, the second is the structure containing the expected result of the procedure.

Objects that are studied in the Ontology of PFC texts and an example of a corresponding syntactic structure will be shown in listing 4 and Fig. 3, accordingly.

PFC element image
=> *inclusion**:
- *resource allocation element image*
- *synchronization element image*
- *procedural element image*
- *procedure nesting indicator image*
- *procedure execution element image*
- *directed link image*

Listing 4. Specification of PFC element images

Resource allocation element image – the image of an oval whose ***identifier**** is the resource specification.

Synchronization element image – image of a rectangle adjacent to the image of a linear primitive, which is the height of the rectangular primitive of the procedural element.

Procedural element image – image of a rectangular primitive, the number of selected right angles indicates the position of the procedural element in the procedural hierarchy.

Procedure nesting indicator image – a plus sign inside the right-hand, separated from the observer, the right angle of the image of the procedural element and touching the boundary of this selected corner.

Procedure execution element image – image of a graphic primitive associated with the phase of executing a procedural element.

Directed link image – image of a linear primitive incident to images of PFC language elements.

Image caption* – a binary relation the first component of which is the sign of the image of the procedural element, and the second is the sign of the file containing some textual explanation to the image of the procedural element.

Any rule in the **Ontology of PFC language rules** is a *correspondence* defined on *atomic formulas*. Semantic and syntactic aspects of entity identification rule are shown in Fig. 4.

C. Ontological model of the cottage cheese production logistic chain

Any production task can be considered as a complex logistical task. However, logistical processes are not limited to the scope of production shops and even the enterprise – they also cover delivery services and interaction with suppliers of raw materials and stores.

Logistical chain of cottage cheese "Khutorok" production includes the following stages:

- Dairy farm
  - Cow
  - Tank
- Dairy plant [production site]
  - Milk truck
  - Finished product shop [production cell]
    - Acceptance post
    - Acceptance tank
  - Soft cheese and cottage cheese production shop [production cell]
    - Milk storage tank
    - Coagulator
    - PFU
    - Cooler
  - Finished product shop [production cell]

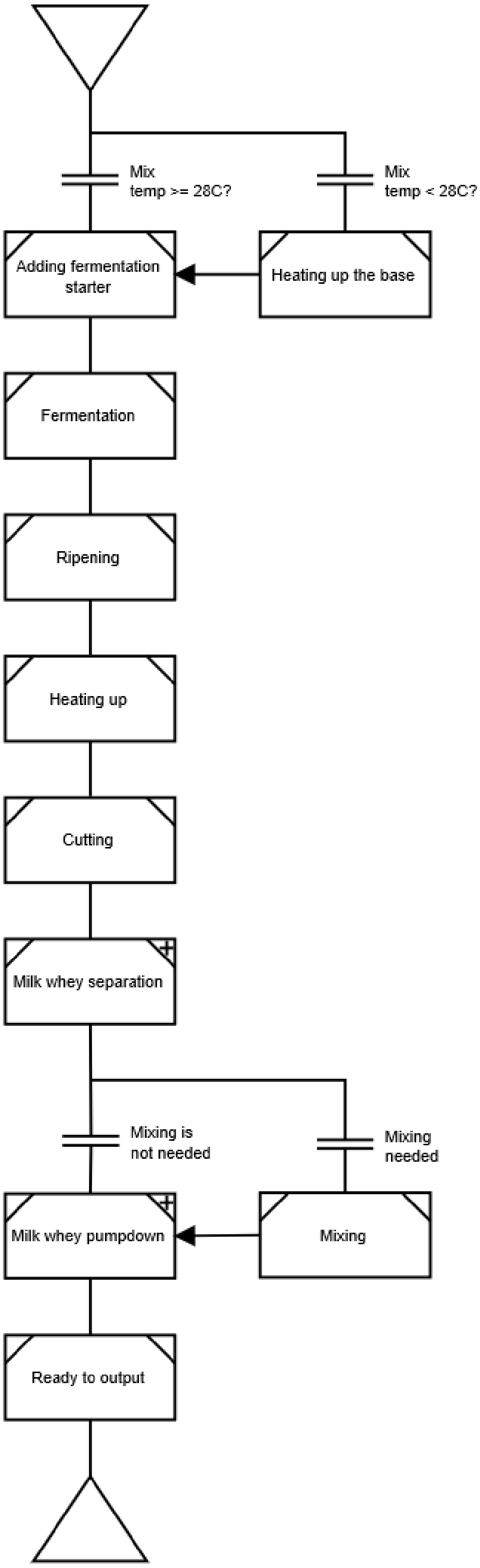


Figure 1. Equipment procedure for producing cottage cheese

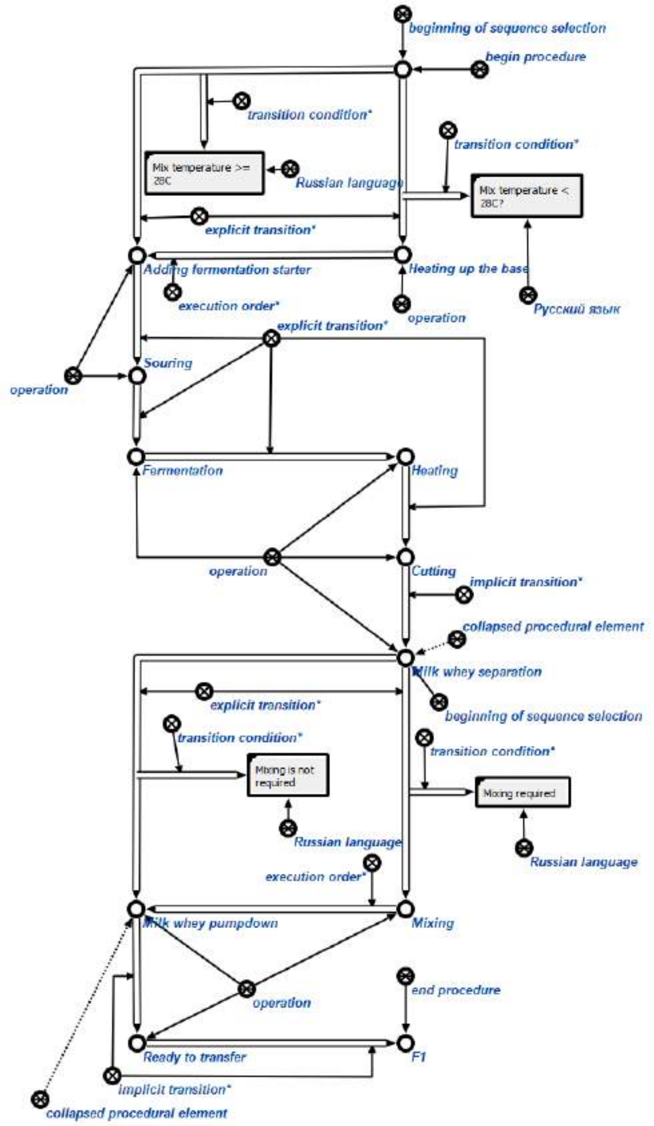


Figure 2. SCg representation of cottage cheese production equipment procedure

  * Box
  * Pallet
  * Storehouse
  * Set-up area Truck
- Store
  - Store warehouse
  - Store shelf
  - Customer

Fig. 5 schematically represents the logistics chain.

Briefly consider how today the main tasks within this chain are being solved.

Milk after milking on a dairy farm is collected in a special container (tank) and cooled before pumping into milk truck. The laboratory on the farm conducts organoleptic, chemical and other tests to determine the milk quality. Information about the samples is recorded in the company's accounting system, for example, 1C Enterprise.

Further, milk is pumped into milk truck, which carries it directly to the dairy plant. In order to minimize delays in the way and to avoid damage to milk, the movement of the milk truck is monitored at the plant. To solve this problem, vehicle position monitoring tools are used – TMS-systems, OpenStreetMap maps, information from the GPS-navigator of the car and the drivers mobile phone. In case of inaccessibility of GPS datalogist can directly contact the driver by phone to clarify it position.

After arrival at the milk plant, the milk again subjected to laboratory tests, the results of which are entered into the enterprise accounting system. Then it is determined in which milk storage tank and from which post milk to be pumped. Milk truck is sent to the appropriate post, where the acceptance operator initiates the milk transfer using the SCADA-system for process control and enters into the enterprise accounting system the necessary data for input raw material accounting.

After pumping milk in a tank, the milk is cooled. The acceptance operator, in coordination with the opera-tor of the hardware shop and the operator of the curd shop, prepares the mixture (using the pasteurizer) and feeds it to the desired curd shop coagulator (several SCADA systems, each operator uses its own project). The masters of the corresponding shops also keep records about intershop movement of mate-rial values using enterprise accounting system.

From the mixture in the curd shop coagulator, the operator prepares the curd mass, controlling the process by means of the SCADA system, and then supplies it for forming to the PFU. Operators of the filling line or robots shift the formed product into a consumer packaging – polyethylene packaging. Packed cottage cheese is labeled, cooled and the operators (or robots) fit in boxes, the boxes are stacked on pallets and through the conveyor get to the automatic warehouse, that is managed by a WMS (warehouse management system).

The pallet from the automatic warehouse along the conveyor is delivered to the set-up zone that is also managed by the WMS. There storekeepers or robots carry out loading of machines that deliver products to customers. At this time, the masters record the shipment of finished products in the enterprise accounting system.

Machines deliver products to specific customers, such as shops, retail chains, etc. Logistics monitor for delivery to the buyer product, using the same monitoring tools as in the milk delivery from farm to factory – TMS-system, OpenStreetMap maps, information from the vehicle GPS-navigator, GPS-navigator from driver's mobile phone, checking calls to the driver, and so on.

Thus, the logistics process for the "Khutorok" cottage cheese is rather complicated – many services and specific people are involved, about a dozen different software tools are used, the consistency of information in which is often supported manually by operators calling, manually entering information into accounting systems, etc.

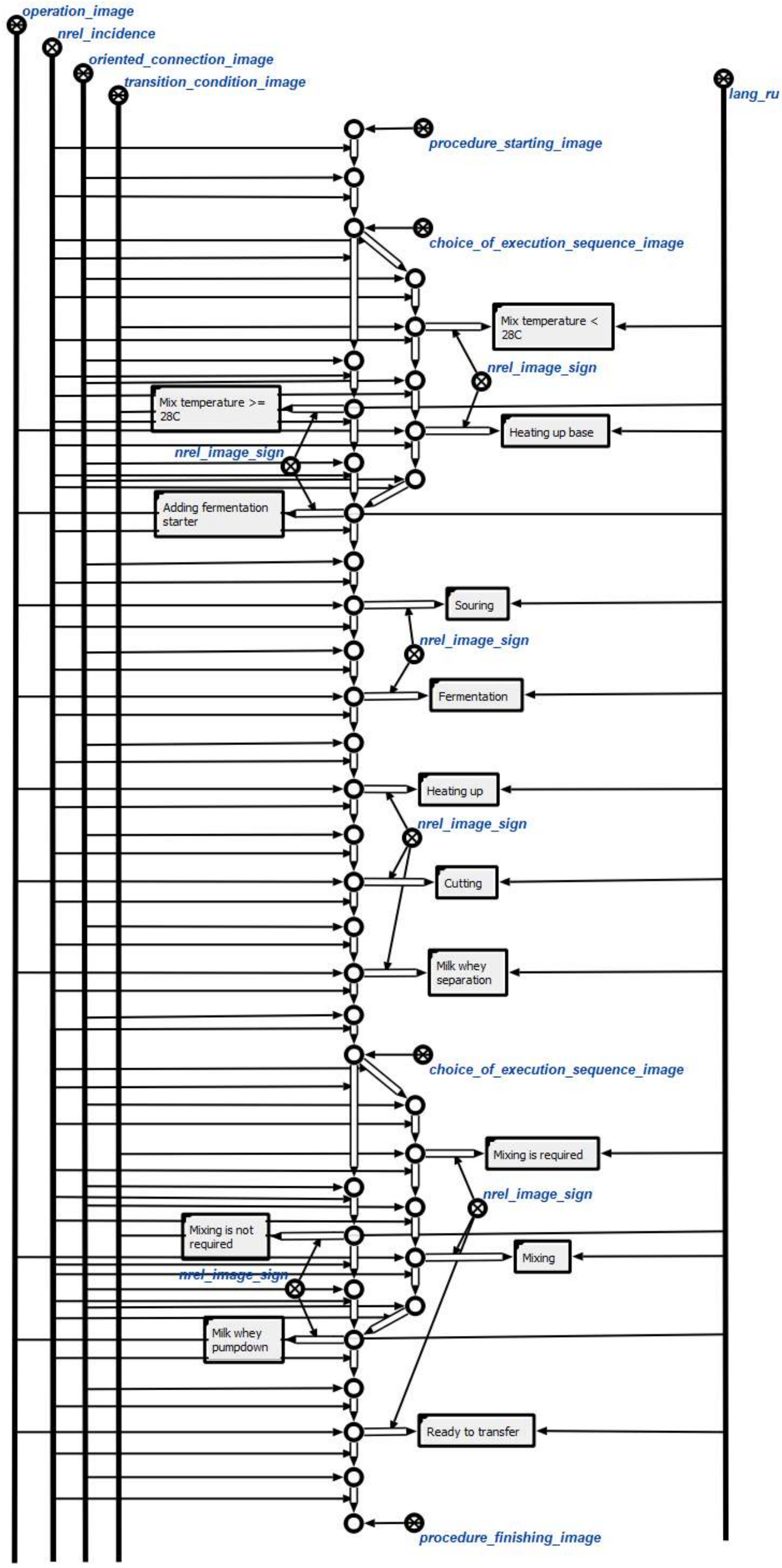


Figure 3. Syntax structure of the PFC diagram of the equipment procedure of the cottage cheese

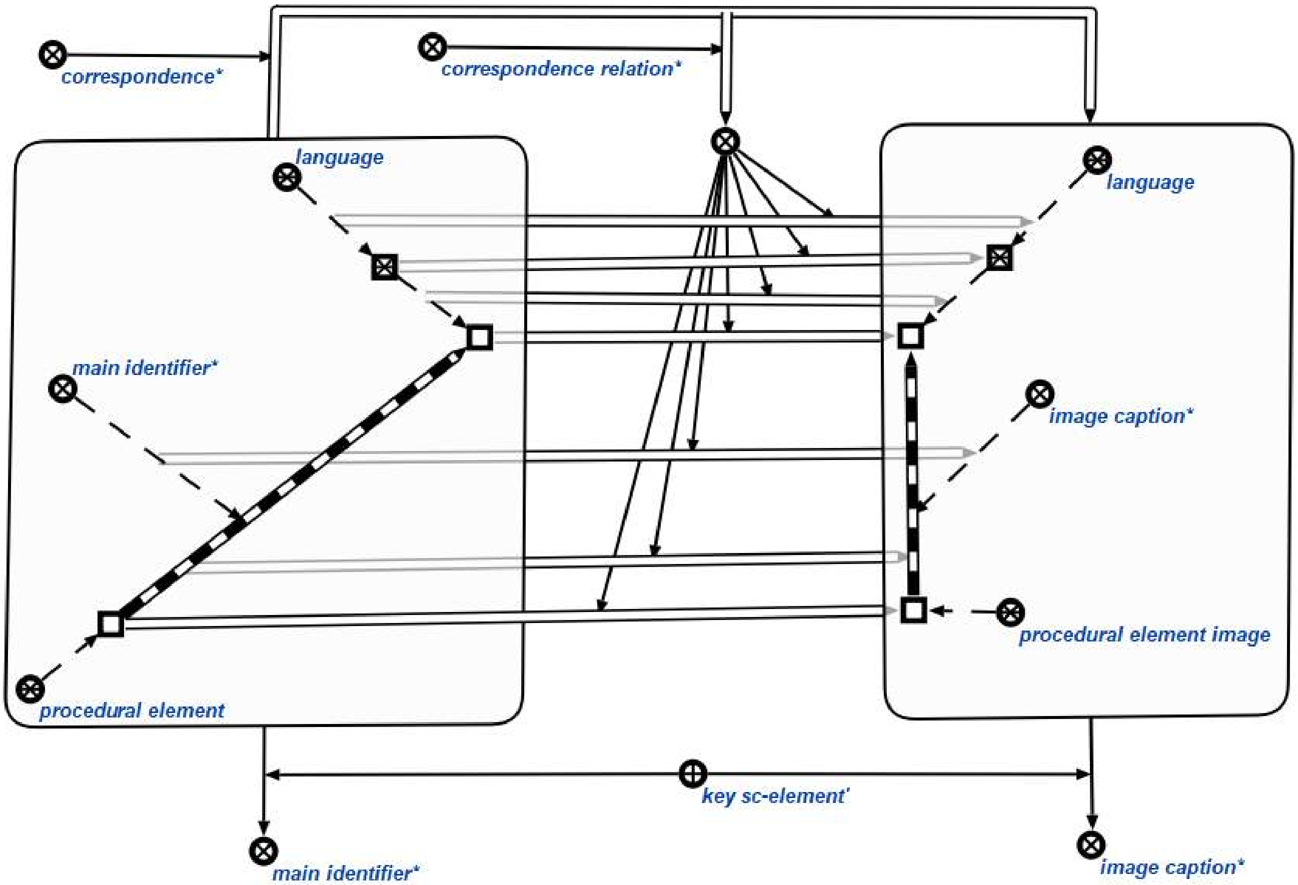


Figure 4. Translation rule example

So, even some parts of the logistics process, for example, warehouse management, include several subprocesses (business processes of the second level) [19]. You also need to take into account the restrictions on the production process duration (a few hours) and selling (a few days) of the manufactured products. The above description omits planning processes for short, which involve additional personnel and software.

In accordance with the ontological approach to the design of the enterprise, it is necessary to form a hierarchical system of subject domains and their ontologies to describe the logistic aspect of the enterprise's activity. The structure of the relevant sections of the knowledge base is shown in Listing 5.

Section. Subject domain of logistics
*<= section decomposition**:
{
- *Section. Subject domain of logistics processes*
- *Section. Subject domain of routes*
- *Section. Subject domain of logistics process participants*
  *<= section decomposition**:
  {
  - *Section. Subject domain of customers*
  - *Section. Subject domain of suppliers*
  - *Section. Subject domain of personnel*
  - *Section. Subject domain of transport*
  - *Section. Subject domain of orders*
  - *Section. Subject domain of logistics documents*

  }

}

Listing 5. Section structure of the logistics knowledge base

Listing 6 shows the examples of structural specifications of some subject domains.

Subject domain of logistics processes
∈ *key sc-element'*:
Section. Subject domain of logistics processes
∋ *maximum class of research objects'*:
logistics process
∋ *researched relation'*:
- *working hours**
- *start time'*
- *end time'*
- *execution time**
- *working days**

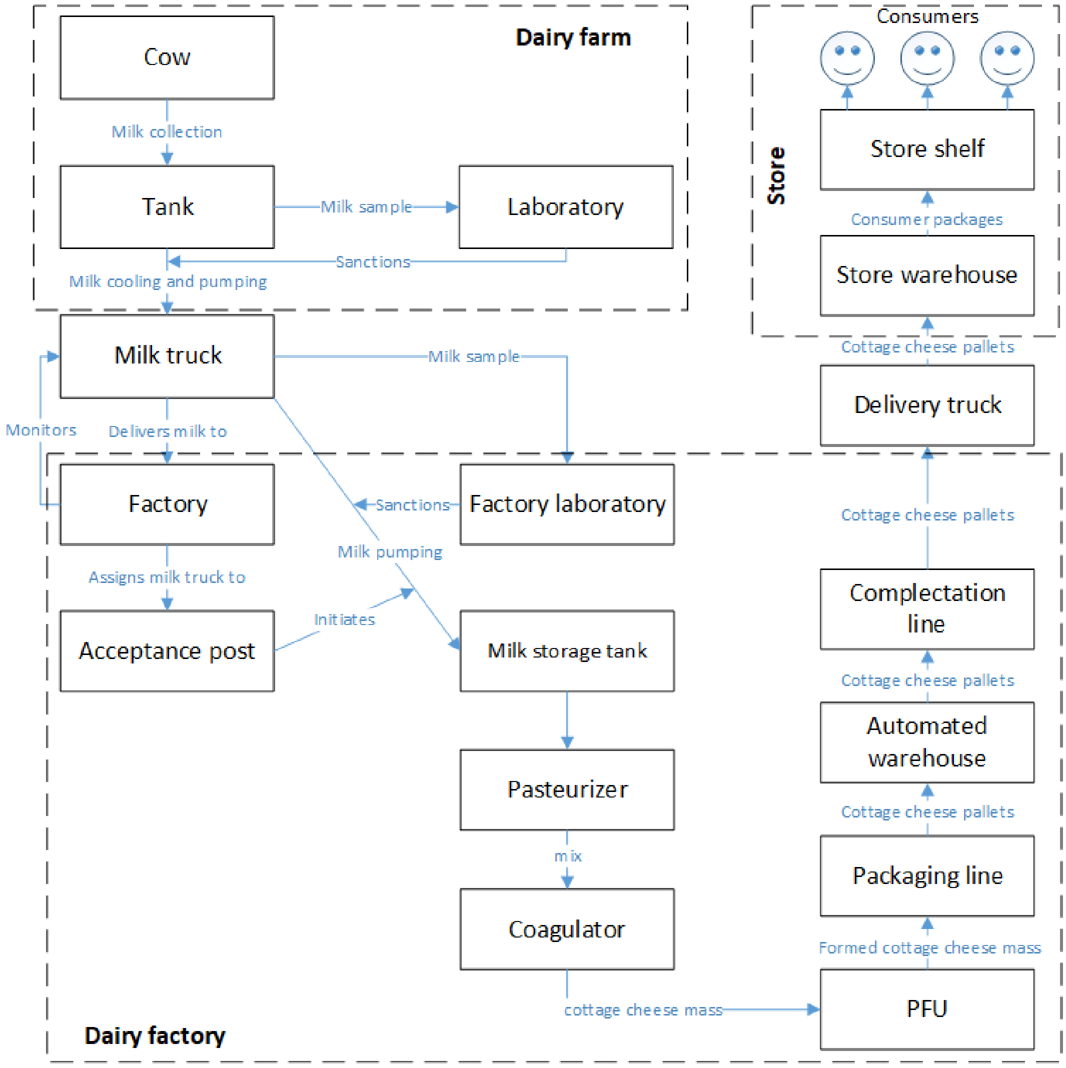


Figure 5. Logistics chain of cottage cheese production

∋ *non-maximum class of research objects'*:
- *truck arrives at a warehouse*
- *processing of the customer's request*
- *planning of shipments*
- *receiving customer orders*
- *route calculation*

Subject domain of routes
∈ *key sc-element'*:
Section. Subject domain of routes
∋ *maximum class of research objects'*:
route
∋ *researched relation'*:
- *destination point'*
- *departure point'*

Listing 6. Subject domain specification examples

Based on the developed system of subject domains, an example of a logistical situation involving two vehicles was formalized, one of which was involved in an accident.

Listings 7 and 8 depict formal description of the situation.

Truck 1
=> *coordinates*:*
[52.19206, 25.266405]
∈ *participant'*:
- *Car accident*
- *Glaze*

∋ *departure point'*:
"Ruzhany-agro" farm
∋ *destination point'*:
JSC Savushkin Product
∈ *truck*

Truck 2
=> *coordinates*:*

[52.265865, 23.967364]
∋ *departure point'*:
JSC Savushkin Product
∋ *destination point'*:
"Vasilishki" farm
∈ *truck*

Listing 7. Example of logistics situation participants specification

Listing 7 shows specification of the participants of the logistics situation. Note that situation participant specification uses two role relations – departure point' and destination point' which are researched within Subject domain of routes.

The above specification says that there are two trucks on the way, the first one is from the Ruzhany-Agro farm to the plant, the second from the plant to the Vasilishki farm, for each of them the GPS coordinates taken with vehicle sensors. First track also participates in "Car crash" and "Glaze" situations. Car crash situation specification is shown on Listing 8.

Car accident
=> *goods loss**:
3 liters
=> *delay**:
2 hours 30 minutes
=> *accident**:
...
=> *closest place**:
Ruzhany
=> *consequence**:
• *truck repair*
• *driver treatment*
=> *material damage**:
• *300 BYN*
<= *cause-effect**:
• *Glaze*
∋ *participant'*:
Truck 1
∈ *car crash*

Listing 8. Example of car accident situation specification

The knowledge base fragment, shown in Listing 8, indicates that the vehicle was involved in an accident near Ruzhany due to the icy conditions. Because of this accident, 3 liters of transported milk were lost, and the delivery of the remaining milk was delayed by 2.5 hours; car repair and driver's treatment are required. The total damage from the accident is 300 Belarusian rubles.

To visualize the logistics situation on the geographic map, a demonstration prototype of the corresponding user interface component was developed. The map is provided by open geoinformation web service OpenStreetMap [20]. Trucks on the map are placed in accordance with the latest GPS-coordinates fixed in the system (they are related by the ratio of the coordinates* to the corresponding vehicle). The damaged truck is highlighted in color. Logistics situation described in listings 7 and 8 is represented as a map in Fig. 6.

D. Intellectual integration of robots into production complexes

Robotics integration into industrial plants from year to year is increasing, and it requires a reduction of labor costs for the design, development and installation of robotic cells. The robot is an universal machine, but it still requires special preparation of the environment for its work, setting up of the software and algorithm for the certain tasks according to its placement Thus, the universality of the robot as a hardware device is decreased by the specialization and uniqueness of the software that manages it in a particular task. Classical programming of industrial robots based on generation control system for end of tool needs a lot of time and labor resources. Existing systems of off line programming allow to create 3D models of the production line, load robot models and design a robot control algorithm in a virtual environment, and then transfer the control system to a real robot. This approach make possibility to reduce the time and complexity of introducing robotic cells into the production process. However, such a solution is only an automation of the problem of designing and programming an industrial robot. The problem of universal control algorithms development in these software products is not supported, although a significant contribution is the possibility of programming in high-level languages (Java, Python) in contrast to the platform-dependent languages of industrial robots.

Like authors noted before, one of the key production tasks is the task of logistics.

Let us consider an example of a specific production problem in field of logistics solved at the batch manufacturing enterprise with the use of robots in the context of the production of "Khutorok" cottage cheese – finished products packaging.

The using of robots in the task of production logistics is dictated by the high productivity of machines, the need for continuous and accurate processing of goods. The integration of a robot into such a process requires setup of the control system for a specific product, line parameters, etc., which reduces flexibility and production possibilities. The Intelligent industrial robot control system that can be independently reconfigured depending on the type of product, line parameters and production process gives the required universal to the industrial robot cell.

The second class of tasks are organization of few working lines equipped with several robots (see Fig. 7), which interact with each other by processing products going along the line. So on the line for the production of cottage cheese is meant the use of up to 6 robots of various kinematic schemes.

As it was said before, within the proposed approach each robot (or robot complex) is treated as an agent over a shared semantic memory, reacting to events occurring in this memory and specifying all its actions in it.

We can say that the mechanical part of the robot acts as a hardware interpreter of a certain class of programs, and the program part of the robot (programmable controller, etc.) – as a compiler of the robot program stored in semantic memory, into a set of signals understandable by the mechanical part of the robot.

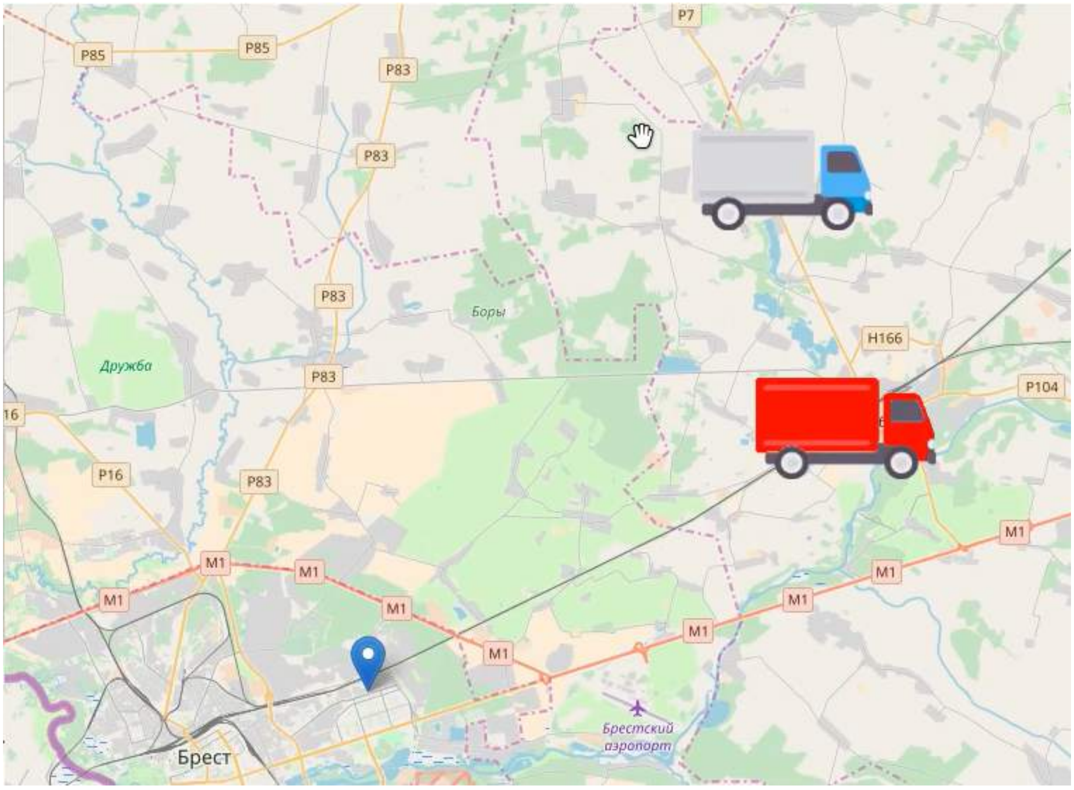


Figure 6. Image of the logistics situation on the map

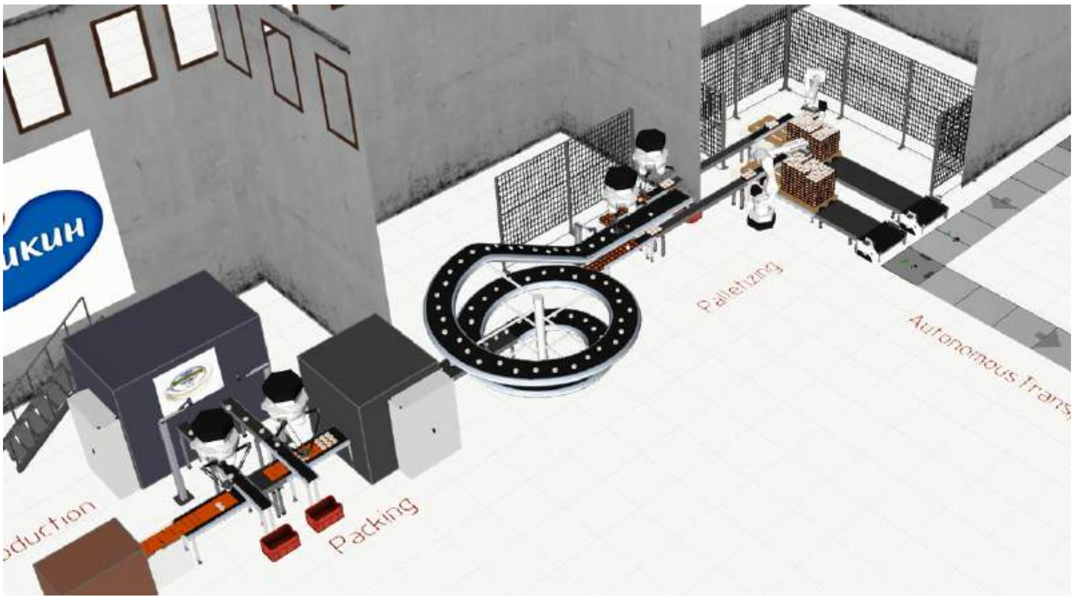


Figure 7. Simulation model of the curd shop with 6 robots

In this case, the addition of a new robot or robot complex is reduced to:

- the development of an ontology of actions that the robot or robot complex is able to interpret, i.e. the description of the denotational semantics of the programs they interpret. In most cases, only classes of actions will differ, other formal means that specify, for example, the sequence of actions or the arguments (operands) of these actions will stay the same.
- the development of tools that allow to transform the

actions specified in semantic memory into signals understandable to the robot.
- the development of the robot activity program itself for the solution of the current class of tasks using the specified ontology of actions.

This approach has the following advantages:

- the robot programming with this approach is clearly divided into logical level (level of performed actions) and hardware level (the level of commands or signals intelligible directly to the robot). This fact provides the following advantages:
  - programming of the robot to perform a specific production task is carried out at the level of operations that are understandable to a specialist in the field of production and does not require special knowledge about the robot's structure, its internal commands and the languages on which programs are compiled at the controller directly;
  - the complexity of the robot reprogramming significantly reduces since reprogramming comes to changing the specification of some actions in the semantic memory and does not require the introduction of changes to a lower level; because the robot program becomes understandable for the automation system itself, i.e. control of the robot's actions and its reprogramming can be carried out in automatic mode, i.e. to be regulated by the system itself;
  - A program of robot actions stored in semantic memory can be visualized in numerous ways, including, in user-friendly graphical languages, which further simplifies the process of manual reprogramming of the robot.
- the presence of shared memory ensures simultaneous consideration of all available robots as a single complex, coordinate their work depending on the needs of the production process, distribute the tasks they solve without having to interact with specific robots in places of physical location. Thus, the management of production processes, monitoring of their implementation can be carried out centrally and remotely;
- as was said before, the approach to communication of system components by means of shared memory provides the flexibility of the system, i.e. allows you to gradually replace manual labor with automatic or introduce more advanced versions of automatic systems without making any changes to the basic automation system. In addition, access to various knowledge stored in the knowledge base allows robots to independently make certain decisions that consider the product nomenclature, their characteristics, knowledge of the product types and ways of products stacking, the location of specific batches, etc.;
- in addition, the specification in the knowledge base of all robot actions provides the possibility of self-learning of the robot based on its own activity, the use of accumulated knowledge in solving typical problems, optimizing its own activity.

Thus, the proposed approach to the intellectualization of production robotic complexes allows us to build a flexible self-adjusting system, which increases the utilization of the robot, shortens the payback period.

III. Conclusion

The article considers the development of the approach to ontology-based design of control systems for batch manufacturing enterprise considering the principles formulated within the Industry 4.0. In addition, the formalization of the ISA88 standard was started in the [10] in terms of the specification in the form of a family of ontologies of syntax and semantics of the graphical language describing the procedural model of the enterprise PFC.

Key points of this paper:

To further increase the manufacturing automation level, it is necessary to consider them from Industry 4.0 — as distributed complexes of control systems, devices, people and services, covering not only production shops, but also warehouses, and interaction with raw material suppliers, wholesale customers and much more.

Considering an enterprise as a multi-agent system over a shared knowledge base built using ***OSTIS technology*** is fully in line with the main principles of the Industry 4.0 concept: interaction, virtualization, decentralization, analysis and response, service orientation, modularity.

A similar approach can be used to organize the interaction of industrial robots within production robotic complexes, which will simplify adding new robots or changing their composition and functionality even to the newbie in the hardware-software platform of a certain robot.

Within Industry 4.0, much attention is paid to human-machine interaction in production automation systems, especially the visual one, which is reflected in the related visual computation concept of Visual Computing [21].

Therefore, the work on formalization of syntax and semantics of the graphic language of the Procedural Function Chart, aimed at the formation of a unified approach to the design of user interfaces of automation systems, also get in the scope of Industry 4.0 direction.

Acknowledgment

The authors would like to thank the staff and administration of the JSC "Savushkin Product" for their support of the work.

References

[1] H. Kagermann, J. Helbig, A. Hellinger, and W. Wahlster, *Recommendations for Implementing the Strategic Initiative Industrie 4.0: Securing the Future of German Manufacturing Industry*, ser. Final Report of the Industrie 4.0 Working Group, 2013.

[2] M. Hermann, T. Pentek, and B. Otto, "Design principles for industrie 4.0 scenarios," in *InSystem Sciences (HICSS), 2016 49th Hawaii International Conference on 2016 Jan 5. IEEE.*, 2016, pp. 3928–3937.

[3] *Applied Cyber-Physical Systems. In S.C.Suh et al. (Eds.).* Heidelberg: Springer-Verlag, 2014.

[4] V. Tarasov, “Strategicheskii inzhiniring predpriyatii budushchego: massovoe sotrudnichestvo, internet veshchei, initsiativa «industriya 4.0», chto dal’she? [Strategic engineering of the enterprises of the future: mass cooperation, internet of things, "industry 4.0" initiative, what’s next?],” in *Inzhiniring predpriyatii i upravlenie znaniyami (IP&UZ-2016): sb. nauch. tr. XIX nauch.-prakt. konf. Moskva, FGBOU VO «REU im. G.V. Plekhanova»*, 2016, pp. 57–68, (in Russian).

[5] Y. Tel’nov, A. Danilov, and V. Kazakov, “Setevaya model’ sotrudnichestva i kooperatsii predpriyatii [Network model of enterprises colaboration and cooperation],” in *Inzhiniring predpriyatii i upravlenie znaniyami (IP&UZ-2016): sb. nauch. tr. XIX nauch.-prakt. konf. Moskva, FGBOU VO «REU im. G.V. Plekhanova»*, 2016, pp. 68–72, (in Russian).

[6] V. Tarasov, “Intellektual’nye sredy: tekhnologicheskaya osnova novoi ekonomiki v smart-gorodakh [Intelligent environments: the technological basis for a new economy in smart cities],” in *Reinzhiniring biznes-protsessov na osnove sovremennykh informatsionnykh tekhnologii. Sistemy upravleniya znaniyami. Sbornik nauchnykh trudov XV-i nauchno-prakticheskoi konferentsii (RBP-SUZ-2012, Moskva, MESI, 26-27 aprelya 2012 g.)*, 2012, pp. 180–194, (in Russian).

[7] V. Kupriyanovskii, D. Namiot, V. Drozhzhinov, Y. Kupriyanovskaya, and M. Ivanov, “Internet veshchei na promyshlennykh predpriyatiyakh [Internet of things in industrial enterprises],” *International Journal of Open Information Technologies*, vol. 4, no. 12, pp. 64–67, 2016, (in Russian).

[8] V. Kupriyanovskii, D. Namiot, and S. Sinyagov, “Kiber-fizicheskie sistemy kak osnova tsifrovoi ekonomiki [Cyber-physical systems as the basis of the digital economy],” *International Journal of Open Information Technologies*, vol. 4, no. 2, pp. 19–25, 2016, (in Russian).

[9] (2017, Nov.) Garbrecht S. The three rules of industrial Operations Management and Industrial IoT Applications. [Online]. Available: https://www.linkedin.com/pulse/three-rules-industrial-operations-management-iot-steven-garbrecht/

[10] V. Taberko, D. Ivanyuk, V. Golenkov, K. Rusetskii, D. Shunkevich, I. Davydenko, V. Zakharov, V. Ivashenko, and D. Koronchik, “Proektirovanie predpriyatii retsepturnogo proizvodstva na osnove ontologii [Designing enterprises of batch production on the basis of ontologies],” *Ontologiya proektirovaniya [Ontology of design]*, no. 2, pp. 123–144, 2017, (in Russian).

[11] V. Tarasov, *Ot mnogoagentnykh sistem k intellektual’nym organizatsiyam [From multi-agent systems to intelligent organizations].* M.: Editorial URSS, 2002, (in Russian).

[12] (2017, Nov.) ISA88, Batch Control. [Online]. Available: https://www.isa.org/isa88

[13] T. Gavrilova, D. Kudryavtsev, and D. Muromtsev, *Inzheneriya znanii. Modeli i metody: Uchebnik [Knowledge engineering. Models and methods: Textbook].* SPb.: Izdatel’stvo «Lan’», 2016, (in Russian).

[14] B. Shvedin, *Ontologiya predpriyatiya: ekspirientologicheskii podkhod: Tekhnologiya postroeniya ontologicheskoi modeli predpriyatiya [Ontology of the enterprise: the experimental approach: The technology of constructing the ontological model of the enterprise].* M: Lenand, 2010, (in Russian).

[15] D. Pospelov, *Situatsionnoe upravlenie. Teoriya i praktika [Situational management. Theory and practice].* M.: Nauka, 1986, (in Russian).

[16] M. Robson and P. Ullah, *A practical guide to business process reengineering.* Gower Publishing, Ltd., 1996.

[17] *Instrumentation, Systems, and Automation Society, 2001, ANSI/ISA-88.02-2001 Batch Control Part 2: Data Structures and Guidelines for Languages.* ISA, Research Triangle Park, USA, 2011.

[18] D. V. Shunkevich, “Ontology-based design of knowledge processing machines,” pp. 73–94, 2017.

[19] Y. Kolchmakhin, “Skladskie biznes-protsessy [Warehouse business processes],” *Skladskoi kompleks*, no. 1, pp. 64–67, 2014, (in Russian).

[20] (2017, Nov.) OpenStreetMap. [Online]. Available: https://www.openstreetmap.org/

[21] A. Stork, “Visual computing challenges of advanced manufacturing and industrie 4.0,” *IEEE computer graphics and applications*, vol. 35, no. 2, pp. 21–25, 2015.

ПРОЕКТИРОВАНИЕ ПРЕДПРИЯТИЯ РЕЦЕПТУРНОГО ПРОИЗВОДСТВА В КОНТЕКСТЕ НАПРАВЛЕНИЯ INDUSTRY 4.0

Таберко В.В., Иванюк Д.С., Касьяник В.В. (ОАО «Савушкин продукт», г. Брест, Республика Беларусь)
Головко В.А. (Брестский государственный технический университет, г. Брест, Республика Беларусь)
Гулякина Н.А., Русецкий К.В., Шункевич Д.В., Борискин А.С., Гракова Н.В. (Белорусский государственный университет информатики и радиоэлектроники, г. Минск, Республика Беларусь)

В данной работе предлагается развитие онтологического подхода к проектированию предприятий рецептурного производства, заключающееся в переходе от рассмотрения производственных процессов в рамках одного предприятия, к рассмотрению процессов, охватывающих, в соответствии с концепцией Industry 4.0, и смежные предприятия – поставщиков сырья (молочные фермы) и оптовых потребителей продукции (магазины, торговые сети). Особое внимание уделяется логистическим процессам – приводится краткое описание основных логистических процессов, касающихся производства творога, структура предметной области логистики и пример формализации нештатной логистической ситуации. Обосновывается соответствие многоагентной системы предприятия со взаимодействием агентов через общую память основным принципам Industry 4.0. В рамках формализации стандартов рассматривается формализация внешнего языка спецификации процедурных моделей Procedure Function Chart, определенного во второй части стандарта ISA-88. Формализация внешнего языка позволяет организовать взаимодействие с пользователями системы автоматизации предприятия на основе унифицированного подхода с использованием понятного инженерному персоналу языка диаграмм. Рассматривается также агентно-ориентированный подход к организации взаимодействия роботов в рамках роботизированных производственных комплексов, основанный на взаимодействии через общую семантическую память.

Railway traffic management system with the intellectual control

Erofeev A.A.
Belarusian State University of Transport
Gomel, Belarus
alerof@tut.by

***Abstract*—The article substantiates the necessity of applying methods of intellectual control in the railway traffic management system. The list of operational tasks that can be solved by intellectual methods is offered. The structure of an intelligent management system for the transportation process is proposed. The stages of the system creation are determined. The experience of implementing intelligent management on the Belarusian Railroad is described.**

***Keywords*—railway transport, transportation process, intelligent management, system structure, operational task**

I. Introduction

Railway traffic management system is complicated but holistic. Intellectualization of any element (goal, subsystem and so on) should be explored in connection with all subsystems and tiers of entire system. System properties are consequence of it complexity:

- the large number of interrelated and interdependent elements, with a change in the performance characteristics of any of the elements reflected in the functioning of the other elements and the whole system;
- the system and the elements included in it are mostly multifunctional;
- the system elements interaction occurs both through the realization of the physical process, and through information exchange, i.e. using transport channels and information channels;
- there is a common goal of the system, despite the diversity of local goals of the functioning of its elements;
- the elements interaction within including system and with adjacent systems (for example, with neighboring roads) is of an indeterminate extent stochastic;
- the human factor, both in making a decision and in implementing it, increases the uncertainty of the system states [1].

II. Theoretical prerequisites of intellectual control

Decision making in the railway traffic management system (RTMS) involves a high level of uncertainty of information about the environment and the object of operation, as well as the need to minimize the risks from making inefficient decisions. For these conditions, the development of control decision (CD) is proposed to be viewed from the position of situational management theory [2], based on the management of complex technical and organizational systems and logical-linguistic models for current situations. The process of situational management assumes the presence of the following elements (Fig 1):

- **analyzer** generates information on a specific operational situation and throws a message about the need for intervention in the management process;
- **classifier** arranges the information about the current situation to one or several classes to which the implemented control must correspond;
- **correlator** receives all information from the classifier and produces a control solution if a single solution comes from the extrapolator, and transfers information to the random selection block if the extrapolator forms several rules;
- **extrapolator** contains all the logical and transformational rules (LTR), determines what LTR should be used;
- **selection block** determine a rule from the LTR proposed by correlator.

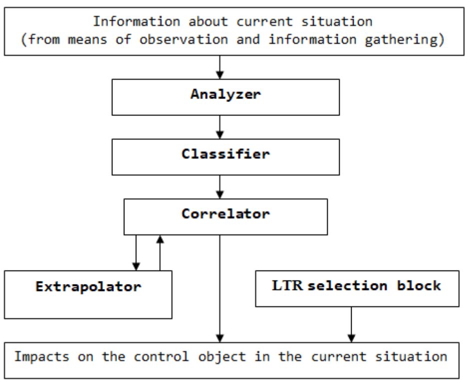


Figure 1. Generalized scheme for the implementation of situational management in RTMS.

Management of the transportation process using methods and models inherent in intelligent systems involves the solution of the following operational tasks:

- organization of railroad car traffic, design and real-time adjustment of the train picking plan;
- train schedule development
- locomotive work management
- real-time planning for work with trains and loading

- empty railroad car traffic adjustment
- real-time management for trains on districts and lines.
- real-time management of station complexes
- real-time management of the local cargo delivery in the local work areas.

The list of tasks to be solved will expand while the intellectual control application area in the railway traffic management system is growing. Within the scope of these tasks, individual subtasks may appear. For example, in the framework of the task "train picking plan design", subtasks can be singled out: design of a plan for sender routing; one-group trains, multi-group trains, local work, etc. This approach does not contradict the methodology of intellectual management, but only confirms the need for designing RTMS as a scalable, self-learning and dynamically developing system.

Regardless of the tier (network, road, linear) and the detail degree of the operational task, for all of them there are unified approaches to describing the problem environment and the search for rational SD. Such approaches are:

- consideration of the operational task solution functional environment as a multi-agent system in which each of the management objects may have local target criteria and, as a consequence, conflicts of interest;
- obligatoriness to consider the problem as multi-criteria. At the same time, both the set of criteria and the level of their significance can change at different times;
- solving operational tasks in conditions of incomplete and uncertain initial data;
- necessity to improve the methods of solving operational tasks through the organization of "learning". The learning of the system involves both the use of more effective methods of searching for rational CD, and the change in the optimality criteria and their significance, the definition (clarification) of the missing initial ones proceeding from the values of the parameters of the objects of the functioning environment;
- the task solution involves not only a quantitative assessment of the parameters characterizing an effective control decision, but also the possible difficulties forecasting associated with the CD implementation. Those the search for rational CD provides for a risk analysis of its implementation and possible consequences.

The use of unified approaches to the description of the problematic environment will provide the following advantages in comparison with the "traditional" methods of solution:

- initial data uncertainty Decrease due to use of the harmonized initial data;
- decrease in entropy in solving operational tasks by using the results of solving some problems as initial data for solving others;
- the ability to search for global extremes in assessing the efficiency of the transportation process, rather than local for each individual task;
- ensuring an objective assessment of the influence of results from solving one operational problem to the decision about control solutions in another task.

III. Experience in the design and implementation of a railway traffic management system on the Belarusian Railways

The Belarusian State University of Transport has developed technical documentation for the development of the Integrated Train Management System at the Belarusian Railways (ITMS-BRW) (Fig. 2).

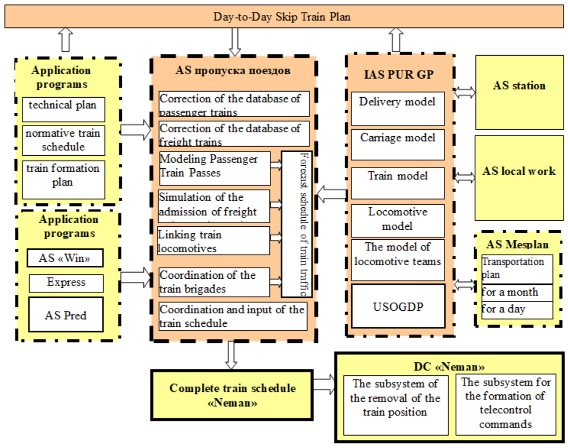


Figure 2. Functional structure of the Integrated Train Management System at the Belarusian Railway.

Creation of the ITMSBRW involves the following stages:

1) Creation of transportation process information and mathematical models on the basis of a unified road network for data transmission, information and reference systems development and implementation, a phased transition to modern microprocessor-based dispatching centralization systems in the sections of the Belarusian Railways.
2) Intellectualized information and planning systems development and implementation, which oriented to the TMC (Transportation Management Center) real-time dispatch center, restructuring of the transportation management system at the Belarusian Railways departments.
3) Transition to intelligent forecasting, planning, management and decision support systems, implementation of measures for additional centralization and concentration of road management, restructuring of the linear transportation management system, development of promising measures, development of necessary technical, technological and regulatory documentation.

At the moment, the tasks of the first stage have been largely solved at the Belarusian Railway: monitoring, displaying and control the signaling devices state, monitoring the train situation, automating the trains' routes, maintaining the executed traffic schedule and applications to it, analyzing the schedule executing, railway cars and locomotive models handling. The release of this tasks set allowed to significantly reduce the load of train dispatchers. As a result, the required number of dispatching circles was reduced from 33 to 21.

The tasks of the second stage are: construction of the forecast train traffic schedule, planning of train building at the stations, locomotives and locomotive crews provision planning for the completed trains (ready for departure) on the basis of the train handling simulation. Currently, most of these systems are in industrial and pilot maintenance in the TMC. Specialists of the Belarusian Railways took part in their development with the direct participation of the BelGUT staff. The automated system for the collection of requests and planning to provide technological gaps (AS "Okna") and the system of accounting and alerting “warnings” for trains (AS PRED) railways have been put into commercial use. The system of real-time train building planning (USOGDP) and the system of the forecast trains schedule automatic construction (AS PGDP) are in trial operation. An end-to-end system of shift-daily loading / unloading planning, which includes all tiers of management, operates on all objects of the Belarusian Railways.

In parallel, BelGUT carried out work on automating the development of normative and variant train traffic schedules, and individual tasks of train formation plan automated calculation were solved. The release of the second stage tasks made it possible to ensure the rhythm of the train traffic process, to reduce the idle time at the station while waiting for departure, and to reduce the required fleet of train locomotives.

Thus, the prerequisites for the third stage intelligent transport systems introduction are created. Such systems include automatic preparation of train routes, offering recommendations on the trains including into the traffic schedule, optimal trains crossing and outrunning, adjusting measures to prevent or eliminate difficulties in the trains handling.

IV. The problem environment of intellectual management in railway transport

Intellectual railway traffic management system (IRTMS) should have qualitatively new "intellectual" properties [3]:

1) the ability to conduct purposefully in any operational environment;
2) the ability to adapt to changes in environmental conditions;
3) the ability to learn and build knowledge bases on the interaction of the environment and the IRTMS;
4) the ability to apply the knowledge acquired to make a decision and organize its execution in changing environmental conditions.

To implement these properties, it is proposed to consider IRTMS as a multi-agent system. A multi-agent system has the following properties [3]:

1) Autonomy and reactivity. The agent is independent and self-governing. The agent receives data from the external environment, responds to its changes and does not require the user to take any additional steps to start the work.
2) Decentralization. There is no agent managing other agents in the system. In this case, one agent may exist in the system, which will generate other agents and set goals for them, but the internal behavior of agents is determined only by their own rules, goals and intentions.
3) Communicative. The agent must communicate with other agents using some agreed language of communication. In the process of communication agents can exchange knowledge or set other agents new targets for implementation.
4) Purposefulness. Each agent must have a specific goal that the agent is trying to accomplish. The behavior of this agent, its interaction with other agents and the external environment must be subordinated to the fulfillment of this goal.

The key aspect for the formation of an effective multi-agent system is the correct description of the problem environment and the subsequent ontological design [4]

When describing the problematic environment of the Intellectual railway traffic management system (IRTMS) for each agent, it is required to establish:

- performance indicators;
- the operating environment;
- executive mechanisms;
- sensors.

Description fragment of IRTMS problematic environment see in table 1.

In the control system of the transportation process, it is often necessary to solve multicriteria tasks, the performance indicators for various agents can vary. In this case, the same indicator in some cases can be considered as determining, in others - as a limitation in the solution of the problem, in the third - as a component of the integrated indicator.

Depending on the state of the problem environment of the system and the macro tasks assigned to the IRTMS, various purposes of functioning may be set. Depending on this, with the same initial data, various optimality criteria, methods and algorithms for solving problems can be used, and, as a result, different ”right” results can be obtained. Therefore, before solving the problem (agent actions), it is necessary to formulate a description of the problem situation.

While formulating the following criteria should be discovered:

- criteria for optimality of decisions;
- key indicators of the quality of decisions;
- features of the external environment of functioning.

In determining the problem situation it is necessary to solve the following issues:

- formulation of the problem and its classification in accordance with the characteristics established in the system;
- discover the newness of the problem
- discover the prerequisites, environmental conditions, macrosystem factors that led to the emergence of a new problem situation;
- description of the relationship of the problem situation to other tasks solved within the system (for example, the need for a planned repair of the track may require an operative correction of the train formation plan or not

Table I
DESCRIPTION FRAGMENT OF IRTMS PROBLEMATIC ENVIRONMENT

The problem environment parameter	**Description**
System of shift and daily planning of loading-unloading	
Benchmarks (Performance metrics)	Compliance with the delivery and the performance of cargo operations, the minimum need for loading resources, the minimum run of wagons in the empty state
Environment	Railway stations open for freight operations wagons, cargo, railway network
Executive method	Showing on the user workspace of plans and tasks for carrying out cargo operations
Sensors (input)	IAS SMD FT[a] messages about operations with wagons, transportation requests from the AS "Mesplan", showing data on the users workspace
Train formation planning system	
Benchmarks (Performance metrics)	Compliance with delivery deadlines, minimum time of wagons in stations
Environment	Technical stations, wagons at stations and in trains
Executive method	Showing train formation plan and train departure schedule on workstations
Sensors (input)	IAS SMD FT[a] messages about operations with wagons, executed train schedule, station tracks model
Train schedule development system	
Benchmarks (Performance metrics)	Provision of the specified throughput (maximum trains number), service speed
Environment	All trains in the main tracks on hauls and stations
Executive method	Showing developed train schedule on the workstations and in the executed schedule
Sensors (input)	Traffic controller workstation
Route preparation system	
Benchmarks (Performance metrics)	The minimum route preparation time, enforcement of the specified time intervals, the maximum level of reliability, provision of traffic safety conditions
Environment	All trains in the main tracks on hauls and stations
Executive method	Traffic light, actuators of switches
Sensors (input)	The hauls and station tracks block-sections controllers of occupancy, state controllers of the switches

[a]Information-Analytical System of Supports Management Decisions for Freight Transportation.

require, affect the unloading parameters of the station stations or will not affect, etc.);

- discover the completeness and degree of reliability of initial data on a problem situation;
- discover the methodology and methods for solving the problem, including on interrelated tasks.

It should be noted that a problematic situation is understood as any technological task being solved, and not only that which is connected with deviation from established modes of functioning.

Whale formulate a description of a problem situation, it is necessary to compose the **solving problem situations library**. I.e. a set of methods and techniques that were used to solve a similar problem in previous periods.

V. CONCLUSION

The involvement of intelligent management in the railway traffic management system in full will allow to increase the productivity of train locomotives by at least 5%, reduce the wagon turnover by 2.5%, increase the productivity of operative staff by 30%.

REFERENCES

[1] Erofeev A.A. Semioticheskaya model' perevozochnogo protsessa i ee ispol'zovanie pri proektirovanii intellektual'nykh sistem [Semiotic model of the transportation process and its use in the design of intelligent systems]. *Intellektual'nye sistemy upravleniya na zheleznodorozhnom transporte. Komp'yuternoe i matematicheskoe modelirovanie [Intelligent control systems in railway transport. Computer and Mathematical Modeling]*, 2017, pp. 24-26.

[2] Erofeev A.A. Intellektual'noe upravlenie perevozochnym protsessom: ot operativnogo k planovomu [Intelligent control of the transportation process: from operational to planned]. *Zheleznodorozhnyi transport [Railway transport]*, No 4, 2017, pp.74-77.

[3] Ryabykh N.P. Reshenie zadach adaptivnogo planirovaniya perevozochnogo protsessa v real'nom masshtabe vremeni v usloviyakh ogranicheniya resursov [Solving the problems of adaptive planning of the transportation process in real time in the conditions of resource limitation] *Intellektual'nye sistemy upravleniya na zheleznodorozhnom transporte. Komp'yuternoe i matematicheskoe modelirovanie [Intelligent control systems in railway transport. Computer and Mathematical Modeling]*, 2012, pp.57-59.

[4] Golenkov V.V. Ontology-based Design of Intelligent System. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for the design of intelligent systems]*, 2017. pp. 37-56.

[5] Directive 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport Text with EEA relevance. Official Journal L 207 , 06/08/2010 P. 0001 – 0013

[6] Gase-Rapoport M.G., Pospelov D.A. *Ot ameby do robota: modeli povedeniya [From amoeba to robot: behavior patterns-* Edition 2], Moskow, Editorial URSS, 2004, 296 p.

[7] Proceedings of the Eastern Asia Society for Transportation Studies, Vol. 5, pp. 1424 - 1432, 2005.

[8] Matyukhin V.G. Intellektual'nye sistemy dlya zheleznodorozhnogo transporta. Opyt i perspektivy [Intelligent systems for railway transport. Experience and prospects]. *Intellektual'nye sistemy upravleniya na zheleznodorozhnom transporte. Komp'yuternoe i matematicheskoe modelirovanie [Intelligent control systems in railway transport. Computer and Mathematical Modeling]*, 2015, pp.3-5

[9] Sinyagov S, Digital Railroad - create digital assets. Sergey Sinyagov, Vasily Kupriyanovsky, German Sukonnikov, Sergey Bulancha, Dmitry Namiot Julia Kupriyanovsky International Journal of Open Information Technologies ISSN: 2307-8162 vol. 4, no. 10, 2016 P. 43-53

[10] Nikiforov V.O., Slita O.V., Ushakov A.V. *tIntellektual'noe upravlenie v usloviyakh neopredelennosti: uchebnoe posobie. [Intellectual management in conditions of uncertainty: a manual]*, Sankt Petersburg, SPbSU ITMO, 2011. 226 p.

ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ В СИСТЕМЕ ОРГАНИЗАЦИИ ПЕРЕВОЗОК НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ

Ерофеев А.А.

В статье обоснована необходимость применения методов интеллектуального управления в системе организации перевозок на железнодорожном транспорте. Сформирован перечень эксплуатационных задач, которые могут решаться интеллектуальными методами. Предложена структура интеллектуальной системы управления перевозочным процессом. Определены этапы создания системы. Описан опыт реализации интеллектуального управления на Белорусской железной дороге.

The Principle of Systems for Information and Analytical Activities Support Building Using OSTIS Technology

Aleksey Kaeshko
Belorussian State University
of Informatics and Radioelectronics
Minsk, Belarus
ondister@gmail.com

Aleksandr Kazakov
Institute of the Border Service
of The Republic of Belarus
Minsk, Belarus
kezanew@yandex.by

***Abstract*—This paper describes principles of systems for information and analytical activities support building using OSTIS Technology. Also the main stages of the creation of such a systems as ostis-systems defined and described.**

***Keywords*—ostis, sc-code, semantic network, decision support system, knowledge base**

I. Introduction

The activities of any organization and its effective evolution are related to a purposeful and continuous analysis of the environment. The quality of management and the continuity of the system functioning depend on the quality of the information received (representativeness, content-richness, sufficiency, availability, relevance, timeliness, accuracy, reliability and stability).

The process associated with the search, collection (extraction), processing and presentation of information is usually called information-analytical activity [1]. The tasks of information and analytical activities in different subject domains may differ, but in whole they are similar:

- Collection of information;
- Processing of information in accordance with the goal;
- Transfer of information to the consumer (the person making the decision);
- Quality control of decision making.

The information systems used to solve this class of problems are Decision Support Systems (DSS).

Among the many possible criteria for DSS classifying [2], [3], [4], [5], the main ones are:

- by interaction with the user:
  - passive - a system that helps the decision-making process, but can not make a proposal, what decision to take;
  - active - directly involved in the right solution development;
  - cooperative - allows the user to modify, supplement or improve the solutions offered by the system.
- by the way of support:
  - model-oriented DSS, which use access to statistical, financial or other models;
  - DSS based on communications, which supports multi-user operation;
  - data-oriented DSS, have access to the organization's time series. In work they use not only internal, but also external data;
  - document-oriented DSS, which manipulates unstructured information, stored in various electronic formats;
  - knowledge-oriented DSS, which provide specialized fact-based solutions to problems.
- by technical level:
  - DSS of the whole organization - the system is connected to large information repositories and serves many managers of the organization;
  - desktop DSS, which is a system that serves only one user's computer.
- by functional filling [6]:
  - Execution Information System (EIS) - has a simplified interface, a basic set of proposed features, fixed forms of information representation;
  - The Desicion Support System (DSS) itself - is a full-featured data analysis and research system designed for trained users, having knowledge in the subject domain of the research.

In all the above variants of the classification, with the exception of the latter, one can trace the assumption of a continuous information exchange between the decision-maker person (DM) and the analyst, carrying out informational and analytical support of the decision. However, within a large organization, the head does not have the ability to carry out permanent monitoring of all information flows. This task is assigned to an analysts, and in this case it would be appropriate to talk about the system for information and analytical activities support (SIAAS), i.e. about the cooperative DSS system of the organization.

Thus, SIAAS operates in close interaction with an analyst (an expert in the field of knowledge). Accordingly, the results of such a system operation depend on the expert's knowledge. Undoubtedly, with this approach, new information should be created by the system without the use of "black box" models. This approach is especially justified in the formation of solutions in weakly formalized subject domains. Currently, such systems are created on the basis of subject domain ontologies, that is, are knowledge based [7]. As a rule, such systems are based on semantic networks and semantic technologies.

II. Problems of existing approaches

Despite the fact that there are approaches to the complex support of the intelligent DSS development process in the weakly formalized subject areas [8], a number of problems are traced in the works.

The first of them is the principle of ontologies creating. Ontologies can be created on the basis of semantic networks using the UML (Unified Modeling Language) [10], RDF (Resource Description Framework) [8] or conceptual maps [11] in a specialized editor only by knowledge base engineers, experts (analysts, users) are not able to do that. Ontologies are created manually, automatic (automated) integration of knowledge from heterogeneous sources is not considered. As a result, the knowledge base of DSS can act as a reference system or be used as a productions system for external programs. New knowledge can appear in the database as a result of knowledge-engineering or as the result of the work of production based programs. In some cases, simple factographic data received from external sources can be placed in the knowledge base [10].

The second problem is the principles of stored knowledge processing. As a rule, for each system, a specific implementation of the knowledge processing machine [12], which exists separately from the knowledge base, is developed. Such machine is a set of programs that interact with each other and the knowledge base directly, which implies a number of problems with the coordination of access to individual elements of the semantic network. In the OWL 2 format, it is possible to describe in the ontology so-called semantic reasoners, to test the ability of a class to contain some individuals, but most of them are capable of processing information only on the basis of direct logical inference.

The third problem is the principles of an interface building. Multiple editors of ontologies assume a universal interface. However, for a person who does not specialize in knowledge engineering, such an interface is not suitable. Moreover, different interfaces are needed to solve different tasks: one need a map to work with geographic information, and a formulas editor for mathematical problems. Thus, a unified knowledge representation does not always imply a unified interface.

And, finally, the process of intelligent DSS development based on semantic technologies is practically never described. It is completely incomprehensible how to parallelize the development of such systems: how to jointly develop the knowledge base and administer it, create knowledge processing agents, and protect information.

As a result of these shortcoming, at present there is no applicable technology to design intelligent systems for information and analytical activities support.

III. Proposed solution

As a technological basis to build a formalized system for information and analytical activities support, it is proposed to use Open Semantic Technology for Intelligent Systems (OSTIS).

OSTIS technology is aimed at development of computer systems, managed by the knowledge [13]. Computer systems of this class, developed by OSTIS Technology, are called ostis-systems. This is a complex technology that defines the theoretical basis for knowledge representation and processing, and contains a software implementation of semantic network storage (sc-storage) for Linux and Windows (64 bit). There are software adapters to working with sc-storage in C, C ++, Java, Python, .Net [14], [15]. The base code for knowledge representation in form of semantic networks with set-theory representation within the OSTIS is called SC-code (Semantic Computer code).

The OSTIS specification defines for any system a platform-independent unified logic-semantic model of this system (sc-model of a computer system) and a platform for such models interpreting. In turn, each sc-model of the computer system may contain a sc-model of the knowledge base, a sc-model of the knowledge base processing machine, an interface sc-model and an abstract sc-memory in which the SC-code constructions are stored. It is possible to describe semantic code constructs in both textual (SCs, Figure 5) and graphical (SCg, Figures 3 ,4) notation. In addition, there is a platform-independent graphical procedural language SCP (Semantic Code Programming), whose program texts are also written using SC-code.

Thus, using OSTIS as a technological platform, we will consider the principles of formalized system for information and analytical activity support design.

IV. Stages of an ostis-system for information and analytical activities support design

The design of the ostis-system for information and analytical activities support should start with the choice of the platform and the deployment of OSTIS [16] technology components. Then it is necessary to develop ontologies of top-level subject domains. To do this, it is enough to complement the sc-models core, which is the most important component of the reusable OSTIS components library, since it already contains ontologies of the most common subject domains, for example, the subject domain of numbers and numeric structures, the subject domain of sets, etc. At the same time, the intelligent metasystem IMS [17] provides consulting services and support for ostis-systems developers. It accumulates libraries of reusable OSTIS components. It is very important at this stage to ensure the IMS and the projected system's

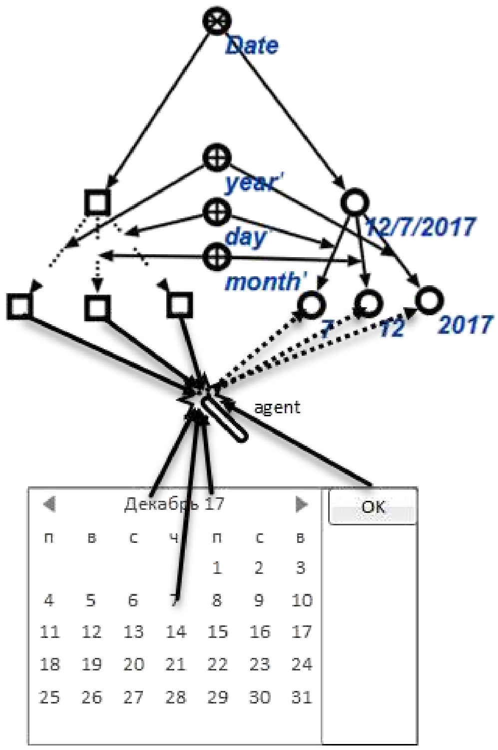


Figure 1. Agent-generated sc-construction based on template

compatibility at the level of entities and sc-models. You need to make sure that the new concept of a child ostis-system is correctly handled by IMS agents, does not contradict the already existing concepts in the knowledge base of IMS. For example, in Figure 3 one can see that the *time in transit in minutes** tag has a data type incompatible with IMS.

At the second stage, a system of subject domains and their ontologies are developed, in which a system for information and analytical activities support operates. First of all it is necessary to implement the ontology of classes of system tasks. As mentioned earlier, only knowledge engineers can create ontologies in the knowledge base. For users (analysts) this process is inaccessible, since it requires special knowledge and skills, as well as for security reasons. Also, there are difficulties with the integration of data from various sources.

To solve these problems, it is necessary to develop a so-called basic patterns of ontology models. Then on their basis, using agents, appropriate ontologies will be created. This approach became possible due to the presence of elements of a variable type in OSTIS. A schematic diagram of such agents operation using the example of the date generation agent is shown in Figure 1.

This approach makes it possible to create the so-called dynamic ontologies that the expert (the analyst) creates himself. At the same time, the creation of structures takes place through a user-friendly interface. So the user may not even suspect that he is working with the knowledge base, not with the regular program. The basic patterns of ontology models allow you to create the same constructions with different timestamps and for different users, which allows you to describe temporary data and solve security problems.

Thus, in ostis-systems, domain ontologies can be created by knowledge base engineers, by agent-adapters from external sources and by users themselves on the basis of patterns of ontology models using the familiar and user-friendly interface.

In parallel with the creation of domain ontologies and basic patterns of ontology models, it is necessary to create software agents for their processing. This can be done in any of the available programming languages for a particular implementation platform or using SCP as a unified cross-platform language for knowledge processing. Several developers can do this at the same time, in conditions that the ontologies and the base model templates are as independent as possible. It is also necessary to implement agents for the production of new knowledge. Thus agents can use any approach to knowledge processing. The core of the ostis-systems knowledge processing machines (sc-machines) already includes a minimal set of domain-independent information search sc-agents, necessary for the ostis-system knowledge base navigation.

The main problem in the software agents development is a performance. The SCP language is cross-platform, the agents, based on it, are part of the knowledge base. However, the execution of programs in this language can be inefficient in terms of performance compared to agents implemented in C++ or on the .Net platform. But the use of agents not implemented on SCP makes the ostis-system platform-dependent.

The third step is a system interface development. The intelligent OSTIS metasystem (IMS) already contains a basic user interface. Its main task is to translate the message from the user, received in some external language, to the internal language of the system (SC-code), and also translate the system response to some external language, understandable to the user and display this answer [18]. The graphical interface is based on the use of SCg-code (Semantic Code graphical), which is one of the possible ways of visual representation of SC-code texts. It is assumed that any ostis-system should use the technology of such interfaces design. However, this does not mean that the ostis-system for information and analysis activities support can not use its own interface designed for it, convenient for the end user and suitable for solving a certain class of tasks. Figure 2 shows the interface of the geo-information subsystem.

In this example, the interface displays text and graphic information about the route. The agent collects the data to display from the knowledge base. Route information is also available for viewing and editing via the unified IMS interface (Figures 3, 4). However, to work in such an interface, you need to know the SCg-code. Thus, it is advisable to use the unified interface for system developers and knowledge base engineers. For ordinary users (analysts) it is better to use a friendly interface developed for them.

Thus, the technology of ostis-systems design allows to create and use any necessary interface, not only unified. For, example some ostis-systems use a speech interface.

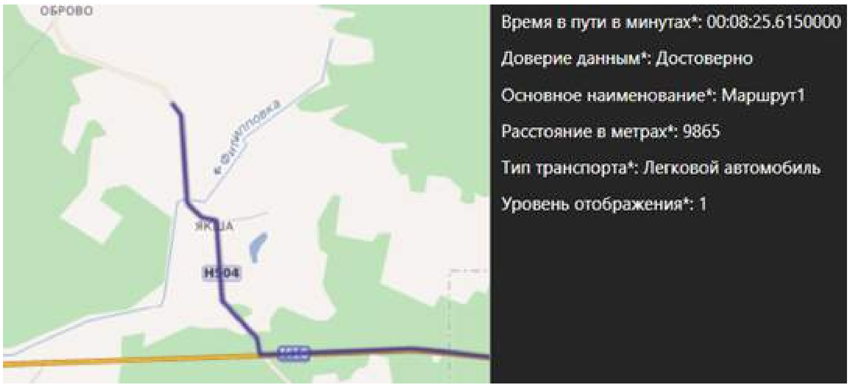


Figure 2. Information about the route in the GIS.

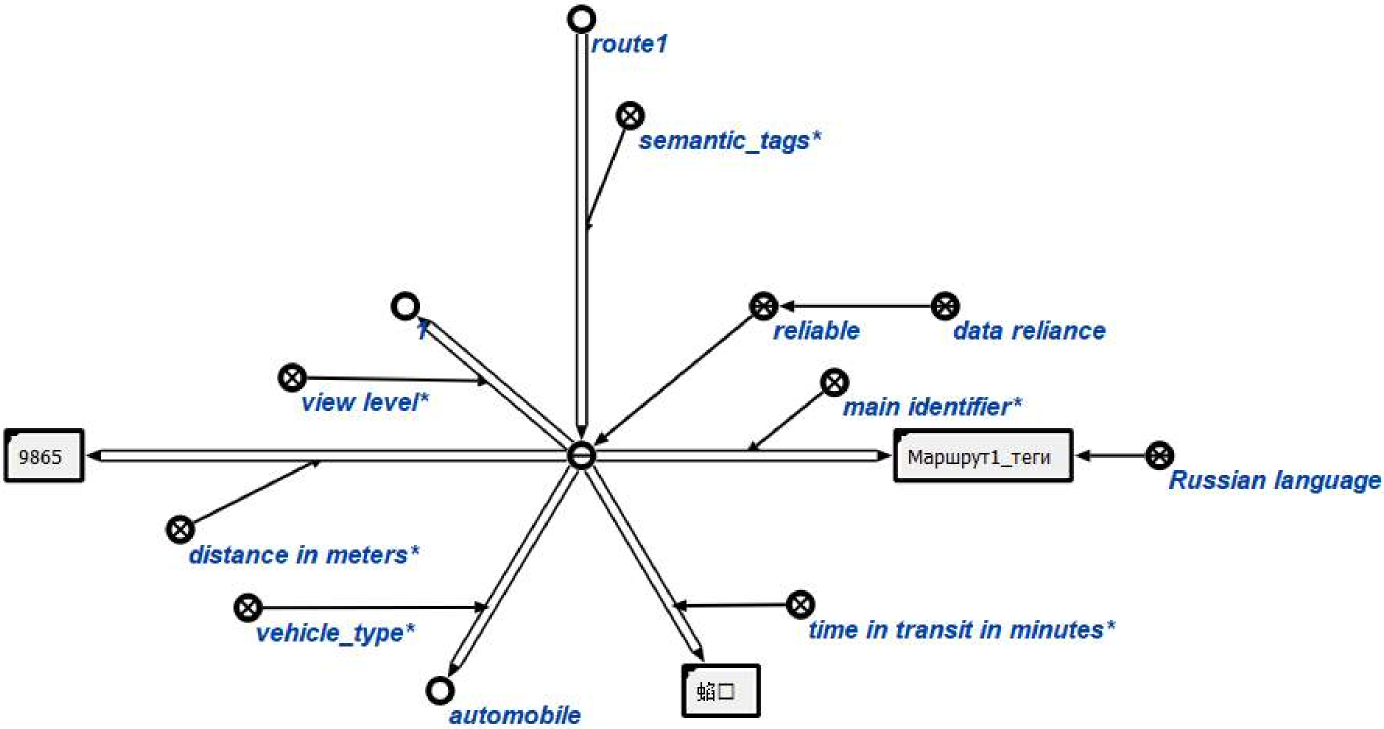


Figure 3. The route tags in knowledge base

The fourth stage is the integration of knowledge. It is necessary to develop appropriate ontologies of sources (recipients) of data, their models. It is also necessary to create adapter models. After the development of collection (data transfer) and information processing agents, it is necessary to take care of its verification. It is assumed that integration with external information sources means avoiding cross-platform. However, at present time due to several technical problems, in practice, any ostis-system with the developed domain-dependent interface already acquires dependence on the platform for which this interface is intended. Integration of information leads to the creation of a semantic network, where the entities from the various components are represented in the form of an interconnected graph.

To improve performance, you can organize the storage of individual elements of the knowledge base in specialized databases, and store only references to them in the knowledge base. So, for example, it is expedient to store video materials, indexes of geodata, indexes of full-text search.

ACKNOWLEDGMENT

To test the proposed principle of system design, we have developed a system for information and analytical activities support in the field of border management, based on the Common Integrated Risk Analysis Model (CIRAM) FRONTEX [20].

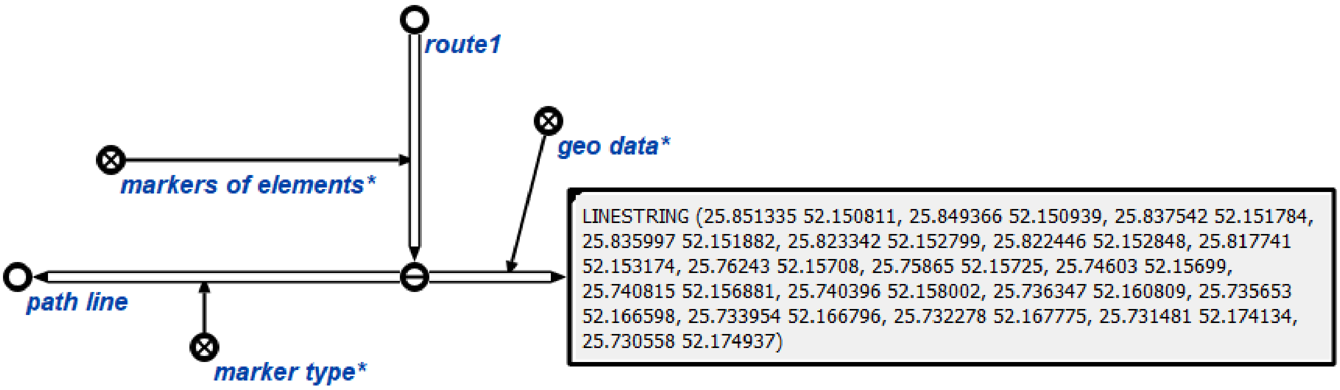


Figure 4. The route marker in knowledge base

```
class_sbs_vulnerability_route<-sc_node_not_relation;;
class_sbs_vulnerability_route => nrel_main_idtf:[Abstract route] (*  <- lang_ru;; *);;
class_sbs_vulnerability_route => nrel_service_description:[Abstract route for analisys]
(*  <- lang_ru;; *);;


_temp_sbs_vulnerability_route<-class_sbs_vulnerability_route;;

sc_node_not_binary_tuple ->_tuple_sbs_vulnerability_route;;
_temp_sbs_vulnerability_route _=>nrel_service_description_tags:_tuple_sbs_vulnerability_route
(*
 _=> nrel_main_idtf: _temp_link_ru;;
 _=> nrel_service_gis_zindex: _temp_service_gis_zindex;;
 _=> nrel_service_gis_vehicle_type: _temp_service_gis_vehicle_type;;
 _=> nrel_service_gis_route_time: _temp_link_time_interval;;
 _=> nrel_service_gis_distance: _temp_link_num_long;;
 _=> nrel_service_info_type: _temp_service_info_type;;
*);;
```

Figure 5. The route template in SCs code

The information system is designed to work in OS Windows 7-10, the platform-independent part is implemented on .Net. During the implementation of the first stage, a small expansion of the Sc-models core was required: several data types (TimeSpan) and an ontology for working with geoinformation content were added. At the second stage, the basic templates of ontology models of countries, vehicles, people, resources were developed. To fill the ontologies, the interface for working with map and other data is implemented. A platform-dependent agent for similar images search is developed. Some simple production inference agents have been implemented, for example, an agent of adding a car model to the models set of a car brand.

The ontology of weather conditions has been developed, software adapters are being developed to obtain weather information from third-party sites. Currently, the development of the system is at the 2nd - 3rd stage of design.

Developed ontologies and basic models templates are maximally independent, which allows to distribute the development among several employees. Thus, OSTIS Technology is an appropriate basis for the design of systems for information and analysis support.

References

[1] Kuznetsov I.N., Informatsiya: sbor, zashchita, analiz Uchebnik po informatsionno-analiticheskoi rabote [Information: collection, protection, analysis. Textbook on information and analytical work] Moskow., Jausa, 2001

[2] Power D.J. A Brief History of Decision Support Systems. Available at: http://DSSResources.COM/history/dsshistory.html (accessed 2017, Dec)

[3] Terelianski P.V., Sistemy podderzhki prinyatiya reshenii. Opyt proektirovaniya : monografiya [Decision support systems. Experience of building: monograph], Volgograd, VolgGTU, 2009. 127 p.
[4] Edwards J.S. Expert Systems in Management and Administration — Are they really different from Decision Support Systems // European Journal of Operational Research, 1992. Vol. 61. pp. 114—121.
[5] Haettenschwiler P. Neues anwenderfreundliches Konzept der Entscheidungs-unterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft. Zurich: Hochschulverlag AG, 1999. pp. 189—208.
[6] dss. Available at: https://studfiles.net/preview/6231972/page:2 (accessed 2017, Dec)
[7] Varshavski P.R., Eremeev A.P., Modelirovanie rassuzhdenii na osnove pretsedentov v intellektual'nykh sistemakh podderzhki prinyatiya reshenii [Modeling reasoning based on precedents in intelligent decision support systems], Iskusstvennyi intellekt i prinyatie reshenii [Artificial Intelligence and Decision Making], 2009. No 2. – pp. 45–57.
[8] Zagorulko G.B., Zagorulko Yu.A. Podkhod k organizatsii kompleksnoi podderzhki protsessa razrabotki intellektual'nykh SPPR v slaboformalizovannykh predmetnykh oblastyakh [An approach to organization of integrated support of the development of intelligent DSS in weakly formalized domains]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2016, pp. 61-64. (in Russian)
[9] Zagorulko Yu.A. Tekhnologii razrabotki intellektual'nykh sistem, osnovannye na integrirovannoi modeli predstavleniya znanii [Technologies for the development of intelligent systems based on an integrated knowledge representation model]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2013, pp. 31-42. (in Russian)
[10] Chernyakhovskaya L.R., Shkundina R.A., Nugaeva K. Ontologicheskii podkhod k razrabotke sistem podderzhki prinyatiya reshenii [Ontological approach to the development of decision support systems] // Vestnik Ufimskogo Gosudarstvennogo Aviatsionnogo Tekhnicheskogo Universiteta [Bulletin of Ufa State Aviation Technical University], 2006. Vol. 4. pp. 68-77.
[11] Desai M., Ghosh A., Das S. Semantic Network Based Modelling to Design a Knowledge-Driven Decision Support Tool for Floral Diversity // Journal of Medical and Bioengineering Vol. 5, No. 1, February 2016. pp.53-57.
[12] Shunkevich D.V. Ontology-based Design of Knowledge Processing Machines // Open semantic technologies for intelligent systems, 2017, pp. 73-94.
[13] Golenkov V.V., Gulyakina N.A., Proekt otkrytoi semanticheskoi tekhnologii komponentnogo proektirovaniya intellektual'nykh sistem. Chast' 1: Printsipy sozdaniya [Project of open semantic technology of component designing of intelligent systems. Part 1: Principles of creation] Ontologiya proektirovaniya [Ontology of design] 2014, vol. 1, pp.42-64
[14] sc-machine. Available at: https://github.com/deniskoronchik/sc-machine (accessed 2017, Dec)
[15] sc-net adapter, Available at: https://github.com/ ondister /ScMemoryNet (accessed 2017, Dec)
[16] Shunkevich D.V. Davydenko I.T., Koronchik D.N., Zhukov I.I., Parkalov A.V. Sredstva podderzhki komponentnogo proektirovaniya sistem, upravlyaemykh znaniyami [Tools for component design of knowledge-driven systems support] Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2015, pp. 125-136 (in Russian)
[17] The IMS.OSTIS website. Available at: http://ims.ostis.net/ (accessed 2017, Dec)
[18] Koronchik D.N. Semanticheskaya tekhnologiya komponentnogo proektirovaniya pol'zovatel'skikh interfeisov intellektual'nykh system [Semantic technology of component design of user interfaces of intelligent systems]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2011, pp. 293-304 (in Russian)
[19] Koronchik D.N. «Umnyi Dom», Upravlyaemyi Znaniyami ["Smart House", Managed by Knowledge]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2016, pp. 155-158 (in Russian)
[20] FRONTEX, Common integrated risk analysis model // FRONTEX Risk analysis unit. Warsaw, 2012. 52 p.

ПРИНЦИПЫ ПРОЕКТИРОВАНИЯ СИСТЕМ ПОДДЕРЖКИ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ НА ОСНОВЕ ОТКРЫТОЙ СЕМАНТИЧЕСКОЙ ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Алексей Каешко, Белорусский государственный университет информатики и радиоэлектроники.
Александр Казаков, Институт пограничной службы Республики Беларусь

В настоящее время в информационно-аналитической деятельности широко используются информационные системы, относящиеся к классу систем поддержки принятия решений (СППР). Некоторые из таких систем основаны на базах знаний на основе семантических сетей. При этом, авторы считают, что системы поддержки принятия решений, основанные на знаниях и использующие опыт аналитика необходимо отнести к отдельному классу - систем поддержки информационно-аналитической деятельности.

Не смотря на то, что существуют подходы комплексной поддержки процесса разработки интеллектуальных СППР в слабо формализованных предметных областях, в настоящее время отсутствует технология создания таких систем. Авторами предложены принципы разработки таких систем с использованием технологии OSTIS. В статье определены этапы проектирования таких систем и содержание каждого из этатов. Рассмотрен подход к созданию т.н. динамической онтологии на основе шаблонов. Приводится пример платформозависимого интерфейса для работы с картографической информацией.

Для апробации предложенного принципа проектирования систем нами разработана система поддержки информационно-аналитической деятельности в области охраны границ, основанной на методологии CIRAM (Common Integrated Risk Analysis Model) FRONTEX.

Reproduction of managerial activity in the system of simulation of military operations

Rulko Yauheny
Military academy of the Republic of Belarus
Minsk, Belarus
achtung-2@yandex.ru

Buloichik Vasily
Military academy of the Republic of Belarus
Minsk, Belarus

***Abstract*—The main requirement for modeling is adequacy. Obviously, in order to obtain adequate results in the simulation system, it is necessary to organize control over the behavior of agents in certain conditions of the emerging situation on the basis of the logic of common sense and the requirements of the statutory documents. For the purpose of forming algorithms for the actions of the opposing sides, it is proposed to use a visual designer that allows you to set sequences for performing various tasks of objects, synchronize them, describe events and set reactions to them. In addition, an approach based on the use of the fuzzy inference algorithm is proposed, which allows automatic selection of one of the alternatives of behavior at an arbitrary point of decision making during simulation. Using the hierarchy analysis method allows you to operate with high-level categories that are understandable to the military specialist when building the base of decisive rules.**

***Keywords*—military operations, imitating modeling, decision-making, fuzzy logic.**

I. Introduction

Currently, the developed countries of the world are actively developing systems for the simulation of military operations (MIS). Examples of this are the Joint Warfare System (USA) [1, 2], the product line of JSC "NPO RusBitekh" (RF) [3], the simulation system for assessing the effectiveness of the air force and air defense troops "Svisloch-1" (RB) [4], etc. The expediency of such developments has long been proven, significant financial resources are allocated to their creation. In the research center for modeling military operations, the bottom-up approach was chosen as the basis for constructing the modeling system being developed, involving the initial creation of models at the lower level of the composition and their subsequent aggregation into larger ones [5]. So, the models of individual modules - the chassis, gun, reconnaissance means, etc. - form a set of base objects, from which autonomous model objects can later be configured - a tank, an infantry fighting vehicle, an archer. Each object is modeled separately, with a high degree of detail. From the model objects are formed group objects corresponding to the organizational and staff structure of military units - Fig.1.

To implement the control of the behavior of objects in the system, a mechanism is implemented for the formation of tasks - functional actions performed by model or group objects, for example, "routeing", "attacking the enemy", "defense of designated lines and positions", "fixed barrage" of artillery and

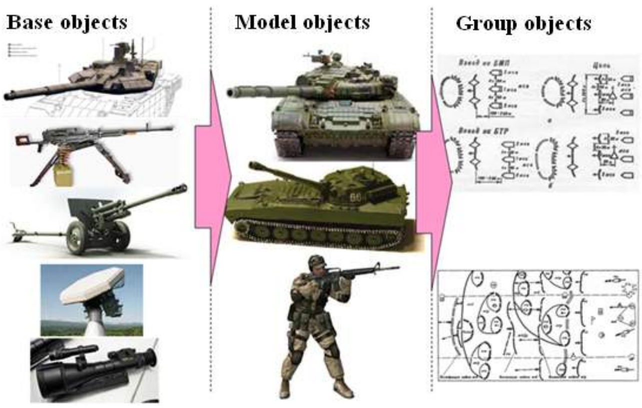


Figure 1. The mechanism of structural composition of objects in the system of simulation of military operations.

The tasks for model objects are set by the user. Major tasks, such as "attacking the enemy," are built on the basis of smaller ones, such as "moving with the exit to the line", "unfolding in line", etc. During the implementation of the task of the group model object, the tasks of model objects included in the its composition. Tasks in most cases require the user to input the original data. So, when the company comes, it is necessary to indicate the line of transition to the line of platoon columns. However, the logic of performing such tasks is set fairly rigidly in the code of the modeling system and can not be changed by the user. Within the framework of simulation modeling, elements of a complex model can act according to a rigid, preset scenario or adaptively, in accordance with a changing environment. Obviously, the second option is more preferable from the point of view of ensuring adequate reproduction of the behavior of real objects, which makes it necessary to develop appropriate mechanisms.

II. Visual constructor

In the course of the research, the following technological approaches to creating tools for managing model objects were tested and tested in practice, allowing them to flexibly change their behavior during the simulation process: the use of an external object-oriented programming language (DSL) [6]. use of scenario (scripting) programming languages, the source code of which can be changed during the work of the modeling system without recompilation; Use the mechanism of compiling source code in C Sharp, which describes the required behavior of objects at runtime; Use of visual programming languages, in which instead of writing code in textual form, manipulation of graphic objects is carried out. Based on the results of the research, it was decided to develop its own visual design tool - a visual designer for controlling the behavior of objects - Fig.2.

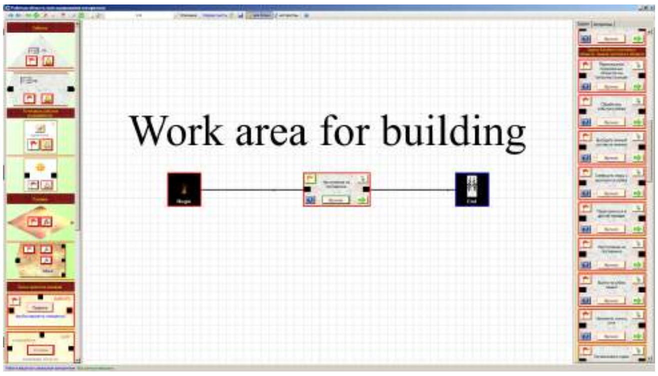


Figure 2. Work area.

Each task in the designer window has its own visual representation in the form of a graphic block having an input and an output - Fig.3.

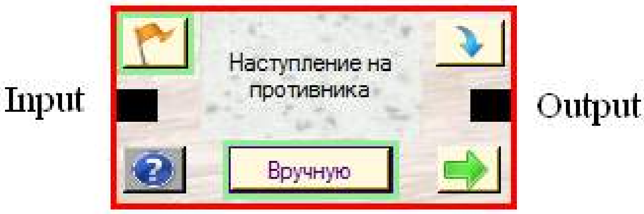


Figure 3. Visual representation of the task block.

Further, the visual scheme, compiled with the help of the designer and describing the behavior of the object, will be called an algorithm. For each task, a specific executing object must be assigned, and the required initial data is entered. Blocks can be connected to each other by directional control transmission lines. The principle of the designer is to sequentially execute the blocks in accordance with the established relationships. To this end, implemented a special program pipeline execution of blocks. Each of the blocks allows you to visualize the current execution status. Due to the hierarchical structure of model objects, at each level of management objects will perform tasks corresponding to their level. Thus, as part of the company's mission to attack the enemy, each of the platoons will carry out its task of attack with its routes of nomination, milestones, etc., but on the whole this can be regarded as the company's actions. In this regard, in addition to the task blocks, block algorithms are used that encapsulate the set of blocks entered by the user, into which, in turn, other algorithm blocks can be included, which results in a hierarchy of subdivision operations - Fig.4.

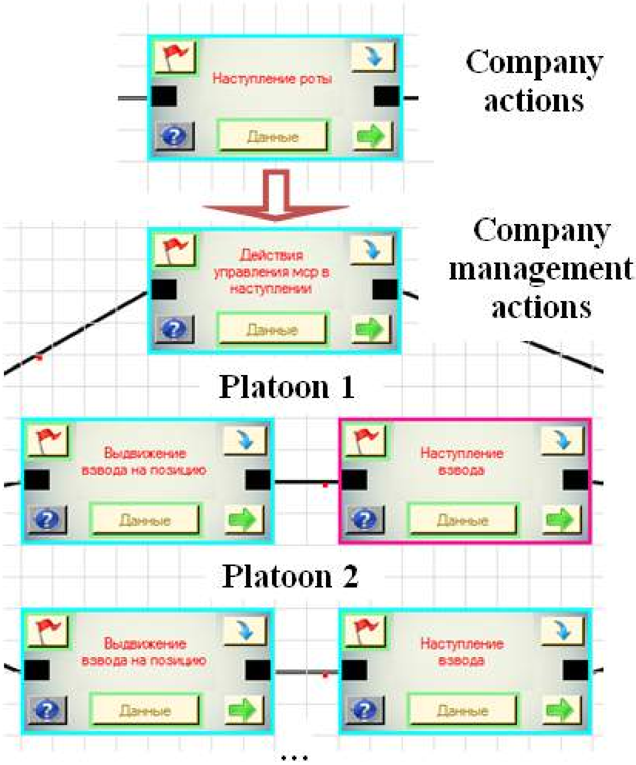


Figure 4. Nested action structure.

In addition to the task blocks and algorithm blocks, there are a number of control blocks: event blocks, logical condition blocks, time "Ch" generation and control blocks, time delay block and others - Fig.5.

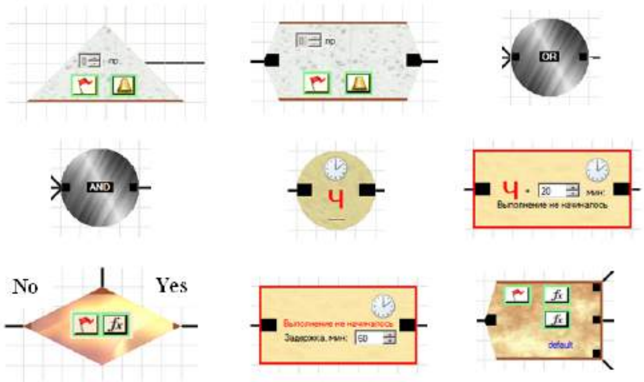


Figure 5. Control Units.

So, for example, the condition block "AND" sends control of the blocks connected to its output once, if control comes from each of the blocks connected to its input. This allows you to perform some actions to comply with the system of conditions. Thus, an offensive against the enemy can be carried out by the unit only after it leaves the concentration area, takes appropriate lines, and the artillery preparation of the prospective area of enemy forces is completed - Fig.6.

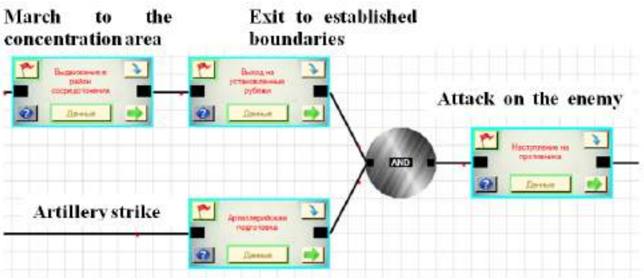


Figure 6. Using the condition block "AND".

With the help of the designer, it is possible to create potential algorithms - templates that define a general order of operations, in which there are no specific initial data. These algorithms can be used in other algorithms of a higher level. In the process of constructing the algorithm, the user is provided with a list of already created potential algorithms for the object of the given and lower levels. There is a special kind of algorithm - the scheme for managing the experiment, within the framework of which control is carried out not by any of the objects, but by the whole set of available objects on the map for one side of the troops (RED or BLUE). Each of the parties can have only one updated, that is, the current management scheme. Likewise, there are potential control schemes that represent a template for conducting the entire experiment for one side or the other. The control schemes can be prepared - they are not executed, but they contain all the data necessary for execution and can be updated at any time. The updated and prepared control schemes are stored together with an example of the situation.

III. Decision Making

At the same time, the question of organization of rational choice of behavior alternatives by model objects in the process of simulation experiment at given points of decision-making remains open. In fact, this creates the need to develop a mechanism that allows to formalize the managerial experience of commanders of appropriate levels and automatically choose the preferred behavior from pre-prepared on the basis of the parameters of the current situation and pre-defined decision rules. As a rule, in the decision-making process, the commander operates with difficultly formalized information at the level of complex categories [7]. In contrast, the state of the object in the modeling system is described as a set of values of specific parameters. To convert a set of object parameters to higher-level parameters, it is suggested to use the hierarchy analysis method [8, 9]. For example, the parameter "Unit losses" can be represented in the form of ranked systems of other parameters - Fig.7.

The second task, which requires its solution, is the synthesis of the apparatus of decision-making on the basis of the

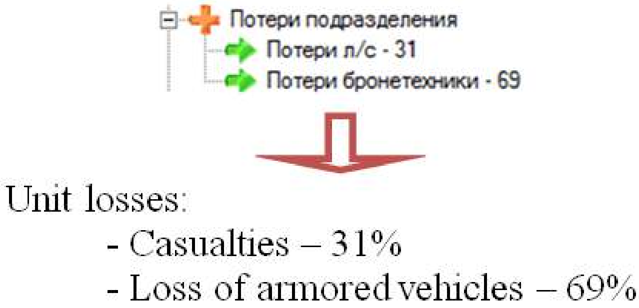


Figure 7. Complex parameter "Unit losses".

complex concepts obtained. For this, the use of the fuzzy inference mechanism is proposed [10, 11]. In this case, the membership functions of fuzzy sets are constructed on the axes of previously formed complex parameters. Each of the membership functions corresponds to a certain logical variable that the military specialist understands. So on the axis of the "unit loss" parameter, fuzzy sets with linguistic variables "low", "tangible", "high", "critical" can be located. Further, using the data of linguistic variables, the formation of the base of decision rules is carried out with indication of their coefficient of certainty:

1) "IF 'Unit losses = critical' THEN 'Waste' (0.8);
2) "IF 'Unit loss = high' AND 'Loss of the enemy = critical' THEN 'Defense' (0.5);
3) "IF 'Unit loss = high' AND 'Loss of the enemy = high' THEN 'Defense' (0.5);

and so on...

At the decision point, the value of the complex parameter is calculated and the values of the membership functions are determined - Fig.8.

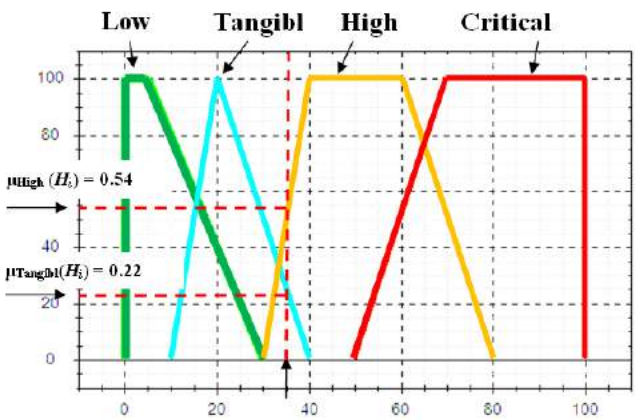


Figure 8. Determining the values of membership functions.

Further calculations assume the fulfillment of the main stages of fuzzy inference. In this case, pre-formed strategies for the behavior of model objects act as consecutive factors of the decisive rules; in this connection, the proposed algorithm does not actually have a stage of defuzzification, and for the accumulation of conclusions, the algebraic union formula is used. All this makes it possible to automatically select an alternative to the behavior at the decision point given in the form of an appropriate algorithm in the designer, by the model object itself in the simulation process without operator participation - Fig.9.

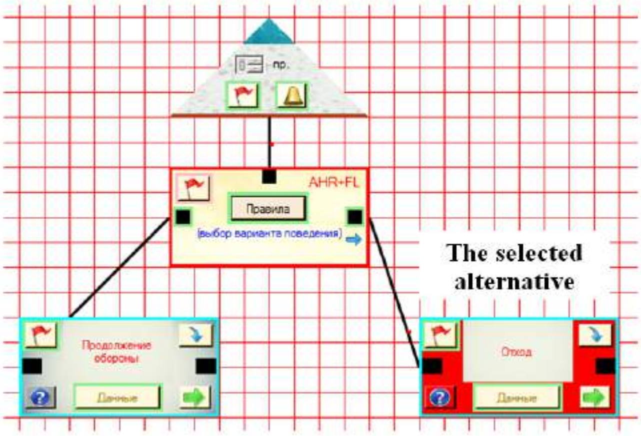


Figure 9. Choice of behavior in the modeling process.

IV. Conclusions

As a result, the proposed set of solutions allows for the formalization of the behavior of objects in the system of modeling military operations by creating event-time schemes of actions of the opposing sides, within which decision-making by objects is carried out on the basis of the fuzzy logic inference algorithm using the hierarchy analysis method. This, in fact, allows you to formulate strategies for the behavior of the parties, determining the sequence of their performance of the task and the response to emerging events.

References

[1] Rezyapov N., Chesnakov S., Inyukhin M. Imitatsionnaya sistema modelirovaniya boevykh deistvii JWARS VS USA [Simulation Modeling of Combat Operations JWARS USAF]. Zarubezhnoe voennoe obozrenie [Foreign Military Review], 2008, no 8, pp. 27-32.

[2] Oblako v pogonakh: voennye sistemy raspredelennogo modelirovaniya. Available at: http://old.computerra.ru//vision//648544// (accessed 2018, Jan).

[3] The AO «NPO RusBITech» website. Available at: http://www.rusbitech.ru// (accessed 2018, Jan).

[4] The Belfortex company website. Available at: http://www.belfortex.com// (accessed 2018, Jan).

[5] Ob utverzhdenii Kontseptsii sozdaniya sistemy modelirovaniya voennykh deistvii: prikaz Ministra oborony Respubliki Belarus [About the statement of the Concept of creation of SMVD: the order of the Minister of Defenses of the republic of Belarus], No 664, 2009.

[6] Fowler M. Predmetno-orientirovannye yazyki programmirovaniya [Domain-specific languages], Moscow, Williams, 2011. 576 p.

[7] Buloichik V.M. Razrabotka modulya prinyatiya reshenii sistemy modelirovaniya voennykh deistvii [Working out of the module of decision-making of system of modeling of military operations]. Sbornik nauchnykh statei VARB [Collection of scientific articles VARB], 2013, no 24, pp. 10-15.

[8] Saaty T. Prinyatie reshenii pri zavisimostyakh i obratnykh svyazyakh [Decision making with dependencies and feedbacks], Moscow, LKI, 2008. 360 p.

[9] Saaty T. Prinyatie reshenii. Metod analiza ierarkhii [Decision making. Analytic hierarchy process], Moscow, Radio i svyaz, 1993. 278 p.

[10] Terano T. Prikladnye nechetkie sistemy [Applied fuzzy systems], Moscow, World, 1993. 368 p.

[11] Leonenkov A.V. Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH [Fuzzy modeling in MATLAB and fuzzyTECH], St. Petersburg, BHV - Petersburg, 2005. 736 p.

ВОСПРОИЗВЕДЕНИЕ УПРАВЛЕНЧЕСКОЙ ДЕЯТЕЛЬНОСТИ В СИСТЕМЕ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ВОЕННЫХ ДЕЙСТВИЙ

Рулько Е.В., Булойчик В.М.

Главное требование к моделированию – адекватность. Очевидно, что для получения адекватных результатов в системе имитационного моделирования, необходимо организовывать управление поведением агентов на основании логики здравого смысла и требований уставных документов. С целью формирования алгоритмов действий противоборствующих сторон предлагается использование визуального конструктора, позволяющего задавать последовательности выполнения различных задач объектов, осуществлять их синхронизацию, описывать события и задавать реакции на них. Для имитации процесса принятия решений объектами предлагается использование подхода на основе использования метода анализа иерархий и алгоритма нечеткого логического вывода. Это позволят осуществлять автоматический выбор одной из альтернатив поведения в произвольной точке принятия решения в ходе имитационного моделирования. Всё это даёт возможность формировать набор стратегий поведения противоборствующих сторон и позволяет оценить каждую из возможных стратегий в сочетании не с одиночным сценарием поведения оппонента, а с системой его возможных действий, что позволяет получить комплексную и всестороннюю оценку рассматриваемой стратегии.

Cognitive ontological models of cyber security in the social networks

A.Aktayeva
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
aaktaewa@list.ru

R.Niyazova
L.N.Gumilyov Eurasian National University,
Astana,Kazakhstan
rozamgul@list.ru

E.Makatov
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
er_han89@mail.ru

A.Dautov
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
d.abeke@mail.ru

***Abstract*—The structure and basic principles of technology for increasing the probability of identifying subjects of information processes of open Internet resources based on ontology methods are considered. Based on this ontology the knowledge base intended for creation of the program systems supporting ensuring information security has been realized. The developed ontological knowledge base has been used when developing the software complex intended for identification of the user of social networks when ensuring information security, monitoring and preventing threats. This article is next in a series of articles by the authors in which they continue to monitor and analyse the current state and new tendencies in the field of information security and safety of information.**

***Keywords*—ontology, knowledge base, information security, Social network, SPARQL, identification**

I. Introduction

During rapid growth in application of information technologies, the successful decision of problems in the field of an information security and protection of the information assumes the more effective activity during a safety in all areas of ability to live of the person. Safety, reliability and privacy are also a part of cyber security. When these systems begin to have a direct physical impact, society has now become responsible for the safety of people and environments.

Historically, the term cybersecurity referred to all the technologies associated with the gathering, processing, storing, and security of information.

The rapid increase in computing and communications power has raised considerable concern about privacy both in the public and private sector. However, with the passage of time and the progress of technologies, the term has acquired different connotations. High latency and frequently the international character of such crimes raise their public danger of the world community.

The modern term, information security or cybersecurity, came into widespread use only in the late 1970s and is now used generally to embrace both computer and communication technologies and their common basis protection – microelectronic technology and all the related software technology. The term Cyberspace notionally represents the various environments that have evolved to support networked computing across the globe. Cybersecurity builds on traditional information security to deal with the evolution of Cyberspace as it grows to include very large and complex systems, mobile computing platforms, cloud computing platforms, and an array of sensors and actuators. The transnational and transboundary character of many products ICT and the international coherence of social networks are used cybersecurity with a view of fulfillment of illegal actions concerning users and the owners the Internet-resources placed in a transnational segment, as well as an AIS, cooperating with the Global network.

The situation is aggravated with the stereotypes which have taken root in a world society about impunity so-called «cybersecurity», the uselessness of accepted measures on strengthening the area of safe use ICTS (info-communication technologies and systems), the limited possibilities of a society on attraction to the responsibility of hi-tech crimes guilty of fulfillment, despite of the developed legal information institutions security in the field of social networks [18].

Cybersecurity should be considered as a sustainable state of the information sphere, ensuring its integrity and protection of ICT infrastructure facilities in the presence of adverse internal and external influences because of awareness of the society of its values, vital interests and development goals.

The neglect of cybersecurity policy when using the resources of social networks of the Internet leads to an increased risk for privacy, unauthorized use or modification of publicly available personal data, as well as disclosure of users' personal data or their transnational accessibility to criminal communities or intelligence agencies of different countries. This, in turn, causes the need to control the subjects of information processes to identify possible areas of information impact and impact on users of social networks on the Internet. Within the framework of this task, it is extremely important to identify the subjects of information processes that can legally distribute 'unreliable or contradictory' messages [15, 16]. Many Internet resources and services, such as forums, portals (social networking resources), online stores, face various manifestations of problems of manipulation and artificial formation of public opinion, by 'organizing' focused thematic dialogues in which many users have multiple account accounts. The possibility of using social portals for information dissemination and insufficient functionality of authentication and authentication mechanisms for users who leave messages determines a few directions for improving protection systems and information security monitoring systems of ICTS.

In this connection, the problem arises of increasing the probabilistic indicators of the quality of methods for identifying users of various Internet portals. One of the promising areas of research in this area is the modeling of cybersecurity systems using the principles of cognitive ontological representation of knowledge, considering the specifications of this subject area.

II. Principles of cognitive ontological representation of knowledge

Ontologies were proposed for the declarative representation of knowledge and are defined in general terms as a special kind of knowledge base or as a 'specification of conceptualization' of any subject area. This means that in the subject area, based on the classification of the basic terms, the main concepts - concepts are singled out, and the connections between them are established - conceptualization. Then the ontology can be represented graphically or described in one of the formal languages (formal ontology) is the ontology specification process.

The ontological representation of knowledge is used for the semantic integration of information resources, adequate interpretation of the content of text documents presented in natural language [15,16,19]. The basis for developing the principles of cognitive ontological representation of knowledge is a diagram reflecting the interrelationships of the basic concepts of security, given in the International Standard for Cybersecurity ISO / IEC 27032: 2012. The ontology formed on this basis reflects only the concepts described in this standard, and only partial details the various aspects that need to be considered when designing a system for ensuring cybersecurity of the social network infrastructure of the Internet [15,16,19]. Then, as can be seen from Figure 1, to ensure the safety of the social networking infrastructure of the Internet, it is necessary to take into account a variety of the factors reflecting the characteristics of all stakeholders, their resources, possible threats and take appropriate response measures against adverse internal and external impacts on ICT infrastructure facilities. Cognitive models are used to model information security threats, and event models are used to model development options for different situations (see Fig.2). And cognitive modeling of ontology is the construction of cognitive models (oriented graphs) in which vertices correspond to concepts, and arcs to connections between the factors [15, 16, 19].

Event-based modeling – the construction of behavioral models (the behavior of users of social networks), and as objects of modeling can be considered as people and technical objects. The essence of the event modeling method is to track the sequence of events on the model in the same order as they would occur in the real system [15,16,19]. The joint use of cognitive and event modeling allows obtaining a more objective assessment of the situation in social networks. For this we introduce ontologies of events used in the transition from cognitive to event models.

When formalizing the ontology problem, it is important to note that: the ontology is one of the tools needed to model the domain; ontology contains a list of key concepts of the given subject area and specification of their meaning; knowledge of the meaning of key concepts represented by the ontology should be obvious to any expert in the given subject area, knowledge bases are developed on the basis of ontologies [12]. The proposed concept supports the technology of researching the directions of the development of the social Internet infrastructure considering the information security requirements. We assume that the infrastructure of social networks of the Internet is defined as

$VSN = \{O, E, MC, MS\} \cup TSN,$

where O is the set of ontologies;
E – set of descriptions of use cases;
MC – a set of cognitive models;
MS is the set of event models;
TSN – an ICT / social networking support tool that includes a description of the knowledge presented in the form of ontologies, descriptions of precedents, cognitive and event models, and means of operating them.

The formal ontology of the subject domain $S\delta$ is the pair S and δ, where δ is the set of key notions of the domain, and S is the set of analytic sentences describing the meaning of these key concepts [2,15,16].

As a result of the analysis of the state of the subject area of the identification of users of the social Internet, it is necessary to distinguish the following: – due to the wide possibilities for providing anonymity to user's social networks on the Internet, the identification methods are of particular importance. However, this method does not take into account changes in the technical characteristics of the device - methods for determining the authorship of the text used by classical linguists show good results for large volumes of text that have undergone correction, but require thorough adaptation for the processing of short messages – to improve the quality of methods for identifying users of social It is necessary to develop a cortege of linguistic signs of a short message, allowing to take into account the peculiarities of construction and of the identifiers.

To date, the most popular methods of identification, using technical characteristics, primarily, such as:

- HTTP Cookie;
- IP-address;
- MAC-address;
- geolocation data;
- data about the operating system, browser, hardware parameters (resolution and screen size, CPU, etc.).

Identification methods using technical characteristics of IKTS are effective for searching for 'single trolls' or unscrupu-

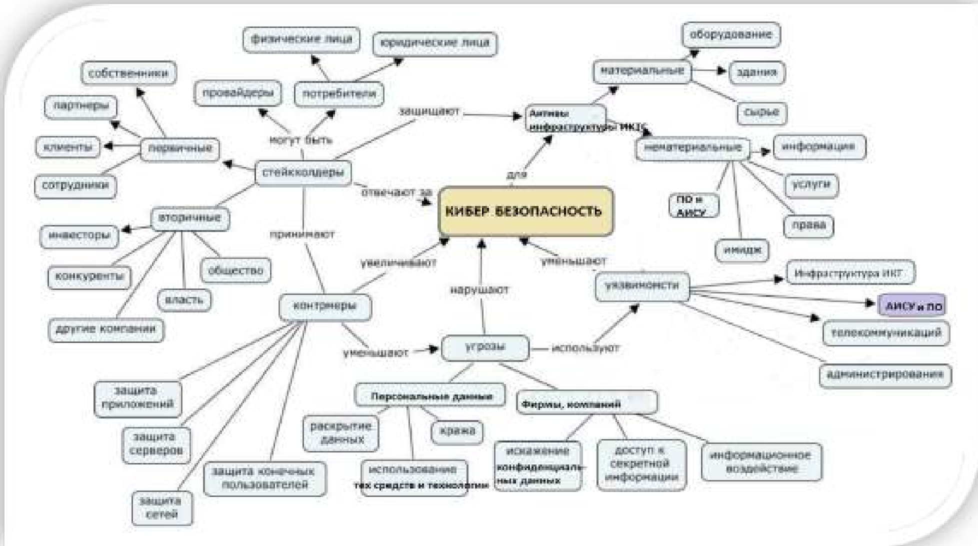


Figure 1. Ontology of Cybersecurity for ISO / IEC 27032: 2012 [15,16].

lous users, but are ineffective for combating and identifying organized astroturfing (AstroTurf) conducted by special organizations that can provide a change in these characteristics [2,4,5,12].

All these features must be considered in the identification process to improve the quality of the results obtained. Thus, based on the received model of a text message containing information about the lexical, graphmatic and syntactic components, it becomes possible to define a profile for each user of the Internet portal.

A user profile is a collection of data and settings from the user's environment. The construction of a user profile is possible on the basis of a number of technical characteristics and statistical data. This approach to creating a profile cannot always give a reliable result. The proposed user profile is especially important in cases where it is possible to replace, clone several technical characteristics of devices, i.e. almost unambiguous identification of the user is impossible [2,4,5,12].

The method of creating an Internet user profile involves the implementation of a number of steps:

- Processing of user messages within the Internet portal;
- Parsing messages by parts of speech followed by the use of templates to highlight the most common constructions.

Lexicographic analysis of the message and allocation of structures in accordance with the patterns described:

- Statistics on the use of punctuation marks and special symbols;
- Selection of lexical constructions based on words and word forms of the language, as well as the identification of thematic special words and phrases specific to a audience online.

The implementation of the proposed method for constructing the user profile of the Internet portal is aimed at solving the indicated tasks in cases where several people use the same PC, or the messages are left by users located on the same local subnet. Figure 3 shows the process of creating a user profile using linguistic characteristics [2,4,5,12].

III. Applications of the ontological representation of knowledge for the construction of the user profile in social network

The ontological approach to the representation of knowledge makes it possible to apply the existing and tested approvals of the advanced profile analytical requests for each user of the Internet portal. To use the ontological representation of knowledge to build a user profile of the Internet portal and to compile a cognitive model of the domain, it is necessary to:

1) Identify significant factors;
2) Construct a matrix of mutual influences;
3) Determine the initial trends of changing factors.

Thus, the infrastructure of the social networks of the Internet includes a knowledge space that integrates: ontological models of knowledge in the field of IS research, knowledge base about precedents in social networks and knowledge bases containing cognitive models of strategic threats of information security and event models of development and consequences of events in social networks, and also tools for describing knowledge (see Fig. 3).

Information extraction is traditionally aimed at finding information that describes a certain area of knowledge specified by the data structure. Ontologies are just a formal domain model expressed, for example, in the form of a graph of concepts and

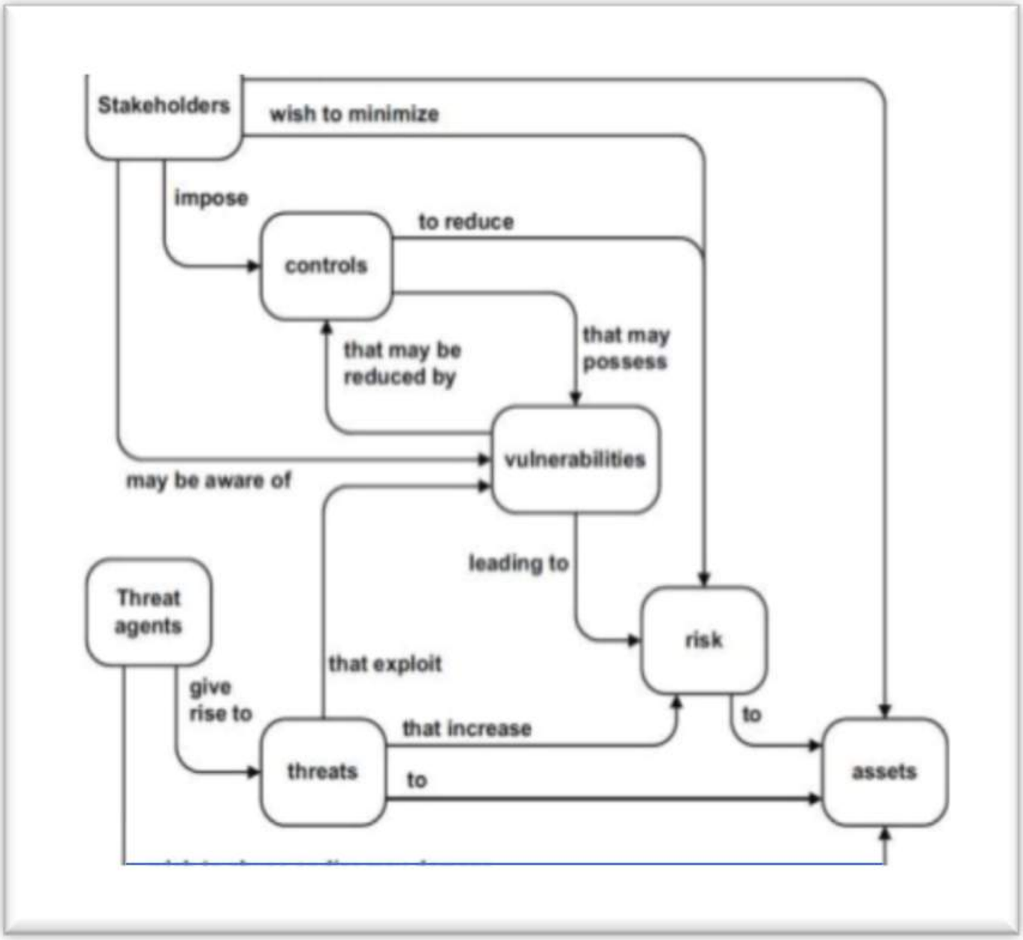


Figure 2. Basic concepts of security and the nature of the relationship between them [21].

Identification methods				
Statistical analysis		Machine learning	Linguistic analysis	
One-dimensional	*Multidimensional*	Bayes Method	Statistical	Analytical
Student's test	Entropic approach	Decision trees		
Two-sided Fisher test	Kolmogorov-Smirnov test	Genetic algorithms		
QSUM	Complexity approach	Neural networks		
χ^2 - Pearson (Pearson's agreement criterion)	χ^2 -n Irson for distributions	Reference Vector Machine		
	Statistical cluster analysis	Method to the nearest neighbors		
	Linear Discrete Analysis			
	Method of main components			
	Markov chains			

Figure 3. Basic Methods of Identifying the Authors of Messages (posts).

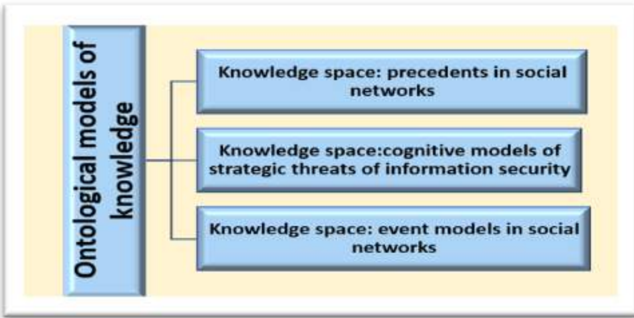


Figure 4. Ontological models of knowledge in the information security: infrastructure of social networks.

relationships, which generalizes the hierarchical data structure that is commonly used for filling in the task of extracting information and includes the steps (see Fig.4).

Indicators of the effectiveness of information extraction algorithms are divided into two classes, namely, the correctness indicators, for example, accuracy, correctness of the extracted information, completeness: the amount of information allocated relative to the volume of all available information and the measure of redundancy, as well as estimates of computing resources such as time and memory.

A query using ontologies can be performed automatically using the mechanisms of logical inference. For this we use the SPARQL language as the query language for ontologies.

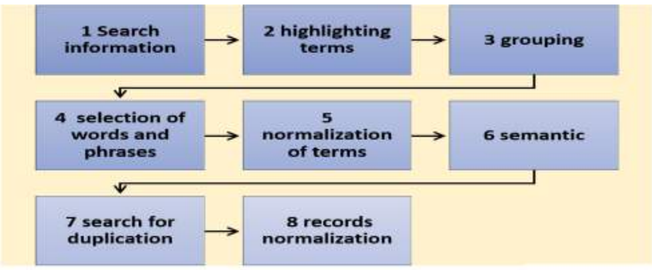


Figure 5. Ontological approach to the representation of knowledge: the allocation of information.

The choice of this language is due to the high level of its development, maturity and good potential, as confirmed by the following facts:

- SPARQL language received the status of official recommendation of the W3C2 consortium in 2008;
- SPARQL language is not tied to a specific program complex, unlike other query language ontologies;
- for SPARQL there is many software implementations and applications [4,5,12].

Below are examples of the use of the SPARQL language in scientific research:

Example 1. The list of topics of portal participants, which are actively explored within the framework of the subject matter of interest. Interpretation of the query: 'to issue all the results for the last (2016) year and sort them according to the descent of the occurrence in these results.' We formalize it in the language SPARQL. First, we will form a set of Terms containing all terms with / without repetitions), compared to the results of the activity for the year.

```
SELECT ?term
WHERE {
?term a cs:term.
?res a swrc:Result.
?res swrc:isAbout ?term . ?res swrc:year 2016. }
```

The resulting set of terms Terms, it is necessary to sort by descending the number of repetitions of each unique element. The terms at the beginning of the sorted list determine the directions that are actively exploring within the area of knowledge of interest.

Example 2.The list of users in the direction of interest. Interpretation of the query in SPARQL language is as follows: 'give a list of users whose search results are related to the terms of the given direction T t1,. . . , tn '. We formulate this query in SPARQL.

```
SELECT DISTINCT ?person
WHERE {
?person a swrc:Person.
?res a swrc:Result.
?res dc:creator ?person.
{?res swrc:isAbout t_1 }
UNION{?res swrc:isAbout t_2}...
UNION{?res swrc:isAbout t_n}.}
```

Example 3.List of publications similar to the given.Interpretation of the query in SPARQL as follows issue a list of the query associated with the terms that characterize the given search".The formal record of this query in SPARQL is presented below:

```
SELECT DISTINCT ?
WHERE {
?p a swrc:Publication . ?term a cs:term.
?p swrc:isAbout ?term . Pub swrc:is About ?term.}
```

Example 4. List of forums devoted to the direction of interest. Let's write this query as follows: 'give a list of forums related to the terms of the given direction $T\{t1, ..., tn\}$'. Interpreting the query in SPARQL:

```
SELECT DISTINCT ?forum
WHERE {
? forum a swrc: Forums.
{?forum swrc:isAbout t_1}
UNION{?forum swrc:isAbout t_2}...
UNION{?forum swrc:isAbout t_n}.}
```

The relationship between queries, the formal model of the system being developed, and the query code in SPARQL allows you to control the impact of:

- modifications of the set of requests received in the system and the ontologies used on the system code;
- modifications in the system code to the used ontologies and considered queries of the IS domain.

And creates additional opportunities for effective software verification at all stages of its life cycle [4,5,12].

IV. Conclusion

Social Networks on the Internet have become an important part of daily digital interactions for more than half billion users around the world. Unconstrained by physical spaces, the Social Networks offer to web users new interesting means to communicate, interact, and socialize. The Social Networks exhibit many of the characteristics of human societies in terms of forming relationships and how those relationships are used for personal information disclosure. However, current Social Networks lack an effective mechanism to represent social relationships of the users that leads to undesirable consequences of leakage of users' personal information to unintended audiences [20].As a result of the theoretical studies and their practical implementation, the ontology of the information security domain was developed in social networks, in particular, a method for identifying the user profile of the Internet was developed. Based on the research of the subject area, models and algorithms are built, architectural and technological decisions based on ontologies are developed to create a system of replenishment and storage, analysis and delivery, upon request, of information activity, information on

Web pages in social networks.The ontology is implemented together with the knowledge base - using ontologies and the SPARQL language, where a formal description of the queries to the system, creating guarantees for their computation and additional capabilities, provides effective verification of the system code at all stages of its life cycle. Such a structure is a distinguishing feature of the developed ontology and knowledge base. It allows you to efficiently process user requests.Be course of cybersecurity raises a host of questions about intellectual property protection and new tools and regulations have to be developed in order to solve this problem. A technological approach to protecting privacy might by cryptography although it might be claimed that cryptography presents a serious barrier to ICT criminal investigations. Therefore, it must be studied how people assign credibility to the information they collect to invent and develop new credibility systems to help consumers to manage the information overload.And the ontology for information security and the base of precedents were used to develop a software package designed to manage risks while ensuring network security, monitoring and preventing threats. We propose an ontological model to represent diverse social relationships and manage self-presentation of social web users.This model regulates personal information disclosure since social role and relationship quality between the users. We also present results of our user study, which demonstrates that relationship quality plays vital role to control personal information security disclosure in social networks on the Internet, and quality of relationship between users can be easily inferred from user interaction patterns in online social networks. The proposed method on the Internet user identification allows to achieve about 70 percent probability for systems monitoring the status of social networking resources.

References

[1] Vasenin V.A. To the creation of an international system for monitoring and analyzing the information space for the prevention and cessation of cyber conflicts // Information Technologies, 2012, No 9, 2-10 pp.

[2] Mirzagitov A.A., et al. Methods of developing an ontology for information security, based on the precedent approach, Vestn. Novosib. state. un-ta. Series: Information technology, 2013, vol.11, No 3, 37-46 pp.

[3] Domarev V.V. The security of information technology. Systems approach. K: DiaSoft, 2004, 992 pp.

[4] Aktayeva Al. & etc. - Technique of identification of users of social networks on ontologies // International scientific journal "Modern IT and IT - education", vol.12, No2, 2016, 26-34 pp.

[5] Aktayeva Al. & etc. - Cognitive Ontology of information security priorities in social networks // International Scientific and Practical Conference Open Semantic Technologies for Intelligent Systems - OSTIS-2017, 365-376 pp., http://proc.ostis.net/eng/ main.html

[6] Schumacher M. Security Engeneering with Patterns, LNSC 2754. Springer-Verlag Berlin Heidelberg, 2003, 87-96 pp.

[7] Jutla D.N., Bodorik P., Gao D. Management of Private Data: Web Services Addressing User Privacy and Economic, Social, and Ethical Concerns, in Secure Data Management. Toronto, Canada, 2004, 100-117 pp.

[8] Undercoffer J. Modeling Computer Attacks: An Ontology for Intrusion Detection. University of Maryland, Baltimore, 2004

[9] Stepanov P.A. Automation of the processing of natural language texts / / Vestnik NSU, Serie: Information Technology, 2013, vol. 11,No 2, 109-115 pp.

[10] Palchunov D.E., Stepanov PA Application of model-theoretic methods for extracting ontological knowledge in the domain of information security // Program Engineering, 2013, No 11

[11] Palchunov D. Ye., Yakhyayeva G. E., Hamutskaya A. A. The program system of information risk management RiskPanel // Program Engineering, 2011, No 7, 29-36 pp.

[12] Golomazov D.D. Selection of terms from the collection of texts with a specified thematic division / D.D. Golomazov // Information Technologies, 2010, No, 8-13 pp.

[13] Kolin K.K. Philosophy of information: the structure of reality and the phenomenon of information // Metaphysics, 2013, No, 61-84 pp.

[14] Ursul A.D. Nature of information: a philosophical essay. - 2 nd ed. - Chelyabinsk, 2010, 231 pp.

[15] ISO/IEC 27032: 2012. Security techniques. Guidelines for cybersecurity

[16] http://www.ontology-of-designing.ru/article /2014 4%2814%29 /7 Vorozhtsova.pdf

[17] https://core.ac.uk/download/pdf/53068632.pdf/data of the application 01.09.2017/

[18] https://online.zakon.kz/m/Document/?doc-id=39754354

[19] http://www.lib.tpu.ru/fulltext/v/Bulletin-TPU/2014/v324/i5/08.pdf/data of the application 01.09.2017/

[20] https://link.springer.com/chapter/10.1007/978-3-319-45880-9-5

[21] https://www.slideshare.net/guidelines-introduction-to-iso-27032

КОГНИТИВНЫЕ ОНТОЛОГИЧЕСКИЕ МОДЕЛИ КИБЕРБЕЗОПАСНОСТИ СОЦИАЛЬНЫХ СЕТЕЙ

А.Актаева
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан
Р. Ниязова
Евразийский национальный университет им. Л.Н.Гумилева
Астана, Казахстан
Е.Макатов
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан
А.Даутов
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан

Рассматривается структура и основные принципы технологии повышения вероятности идентификации субъектов информационных процессов открытых ресурсов сети Интернет на основе методов онтологии. На основе этой онтологии была реализована база знаний, предназначенная для создания программных систем, поддерживающих обеспечение информационной безопасности. Разработанная онтологическая база знаний была использована при разработке программного комплекса, предназначенного для идентификации пользователя социальных сетей при обеспечении информационной безопасности, отслеживания и предотвращения угроз.

Principles of ostis-system of automatic diagnosis design

Rostovtsev V. N.
Republican Scientific and Practical Center for Medical Technologies, Informatization, Administration and Management of Health
Minsk, Belarus
vnrost@rambler.ru

Rodionova O. S.
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
oks.rodionova@mail.ru

***Abstract*—The work is devoted to the development of intelligent ostis-system of automatic diagnosis. This article presents analysis of well-known solutions to automation of physician's activities on handing down a medical decision problem. Article also deals with principles of knowledge base and knowledge processing machine for ostis-system of automatic diagnosis design based on semantic networks with set-theoretical interpretation.**

***Keywords*—medical diagnosis, automatisation, intelligent system**

I. Introduction

Possibility of setting and solving problems of automatic medical diagnosis without physician's participation has appeared in response to creating a new technological generation of wave diagnosis in the early part of the current century. This generation is the third after electrography technologies (gas discharge rendering) and frequency-resonant diagnosis. The new wave diagnosis was dubbed functional spectral-dynamic diagnosis (FSD-diagnosis) and honor of its creating belongs to S.M. Zakirov [9].

FSD-diagnosis can be undertaken by the following areas of activity:

1) nosological diagnosis (diagnosis of diseases);
2) system diagnosis (assessment of body systems functional status);
3) etiological diagnosis (detection of active viruses, bacteria, etc);
4) patient-specific complementary medicated products adjustment.

FSD-diagnosis is accomplished through Medical spectral-dynamic complex (MSDC). It's a computer appliance related to physical medicine solutions and designed for performing quick, delicate and quite universal (in reference to common diseases) diagnosis [4].

It's possible to mark the following FSD-diagnosis advantages:

1) simplicity of assessment procedure;
2) low assessment time (signal recording takes 35 seconds);
3) there's no necessity to process acupoints;
4) passive diagnosis mode, i.e. there's no effect on the human body;
5) instrumentation transportability;
6) accessibility for any physician;
7) diagnostic accuracy (more than 90 %);
8) MSDC marker base includes large amount of diagnostic markers (more than 8,000).

At the level of implemented by MSDC analysis, i.e. during the process of concrete processes presence recognition, three major modules are working with the patient's code including pattern recognition module, dynamics analysis module and abnormal aptitudes analysis module.

Pattern recognition module compares patient's code (pattern) with the codes (patterns) of prescribed diagnostic markers set and calculates the recognition probability metric (in %) for each marker.

Dynamics analysis module puts diagnostic markers in order of corresponding processes activity.

Abnormal aptitudes analysis module evaluates the level of abnormal aptitudes between spectral-dynamic codes of markers and patient on ordinal scale from 1 to 6.

The methodological basis for creating the automatic FSD-diagnosis systems, including telemedicine, is the combination of the passive mode of this diagnosis and its high informational content (diagnostic markers multitude). [3], [9].

II. Problem domain definition

At the physician level, the doctor uses dozens of markers (and sometimes hundreds) to perform the FSD-diagnosis of a particular disease and analyzes three quotients for each marker, including the similarity quotient (marker recognition probability in %), the process activity index (position in the list of the specific body system markers) and an indicator of the level of abnormal aptitudes (on the ordinal scale).

Consequently, physician has to manually process quotients values for markers multitude available from assessment to hand down a decision. This lead to substantial increase in duration of single patient diagnosis and to slowdown in physician's work efficiency due to fatigue cumulation through samely repeating actions performing and hence lead to risk of handing down a wrong decision increase.

At present, there is no such diagnosis automation tool that could integratedly assess (on the strength of all the informative parameters) the acute patient conditions in the express-analysis mode and would help the specialist to establish the definite diagnosis [7].

Based on the above, it follows that it's necessary to automate the pack of physician's measures in furtherance of handing down a diagnostic decision as soon as such opportunity exists in principle.

The existing diagnosis process automation tools require considerable costs to commissioning as well as to training of qualified professionals capable of maintaining such system and manipulating it at first hand [7].

III. Common approaches analysis

Medical diagnosis tasks are nothing more than recognition tasks with unfixed set of recognition objects.

Such knowledge representation languages as rule-based language (by using CLIPS and OPS) and first-order predicate logic (by using Prolog) are available for representation knowledge from medical diagnosis domain. Nevertheless, there are series of restrictions that are imposed upon the models listed above. These restrictions are hindering the designing of flexible and multifunctional system.

The production model has the disadvantage that when a fairly large number of productions (of the order of several hundred) is accumulated, they begin to contradict each other, in which case adding a new rule to the system causes difficulties for the developer and the expert. In addition, there is principal difficulties emerge during the system work validation due to inherent to the system nondeterminacy (ambiguous selection of executing production from the activated productions scope) [11].

One of the logical model disadvantages is that it is impossible to determine the truth or falsity of a sentence using rules specifying the language syntax. The sentence may be syntactically valid, but it could occur completely pointless. Another disadvantage is that most intelligent tasks are exemplified by completeness lack, deficiency and ill-posedness, that hinders their formalising based on sententional calculus [1], [8].

IV. Proposed approach

It's proposed to use Open Semantic Technology for Intelligent Systems (OSTIS) for automatic diagnosis system design [10].

Graph-dynamic models of special form – semantic models of representation and processing of knowledge based on semantic networks with set-theoretical interpretation – are used within the mentioned Technology as a formal basis for the designing intelligent systems.

Technology inherits all the fundamental advantages of semantic networks, and knowledge representation in the form of semantic networks enables to essentially simplify the knowledge integration procedure and to reduce this procedure to determining and merging synonymic elements of integrable semantic networks.

Approach based on semantic networks with set-theoretical interpretation has the following advantages in medical diagnosis domain:

1) knowledge representation by applying semantic networks is easier to understand and easier to see;
2) semantic networks enable to work not only with numeric parameters and characteristics values but with any nominal and order objects;
3) it's easier to follow and refine decisional process in semantic model of knowledge processing, i. e. it's enough to modify few infence rules or add some new ones for modification the decisional process without affecting the entire infence machine in case of making wrong decision or appearing a necessity of expanding the method;
4) system built upon semantic networks can explain the way it made one or another decision at any time, by user request.

The way of internal knowledge representation within OSTIS Technology is called SC-code (Semantic Code). Signs included into SC-code texts are called sc-elements [10]. Nodes of semantic network represented by SC-code are called sc-nodes and connections between them are called sc-connectors. Oriented connections are called sc-arcs, non-oriented ones are called sc-edges.

Semantic knowledge processing (problem solving) model within Technology represents an abstract multi-agent system consists of abstract semantic memory storing semantic networks and set of agents oriented on processing the semantic networks stored in mentioned semantic memory.

Practicability of automatic diagnosis intelligent system design applying OSTIS Technology relies on lack of sufficient tools of physician's activities on handing down a medical decision automation and predominance of approach based on semantic networks with set-theoretical interpretation over other knowledge representation models within medical diagnosis domain.

Systems designed by OSTIS Technology are called ostis-systems. Every ostis-system includes the following architectural components:

1) knowledge base;
2) knowledge processing machine;
3) user interface.

Consider next principles of ostis-system of automatic diagnosis design and composition of components it includes.

V. Principles of ostis-system of automatic diagnosis design

Proposed automatic diagnosis system has to solve the following key diagnosis problems:

- functional diagnostics (diagnosis of functional organism systems states);
- prenosological diagnostics (diagnosis of exposures to diseases);
- prednosological diagnostics (diagnosis of latent diseases);
- nosological diagnostics (diagnosis of symptomatic diseases);
- differential diagnostics diseases and other states.

The set of concrete states or diseases estimated probabilities is an input data for designed system.

The estimation of each state probability is making on the following order scale: very low, low, small, high, very high. Prior and post-hoc significances for each state are determined at some context. Interim decisions hierarchy with ranking will be generated in terms of those significances. Also it's determined which estimate scale values will be taken as crucial or near crucial values for each state.

Interim decisions in considered system are handing down in terms of patterns predefined within the scope of estimates subset designated for analysis per each state.

An end-point decision based on analysis and union of all interim decisions made for each assessed state is an output data for designed system.

Operation principle of proposed ostis-system of automatic diagnosis can be submitted in a form of scheme displayed on Figure 1.

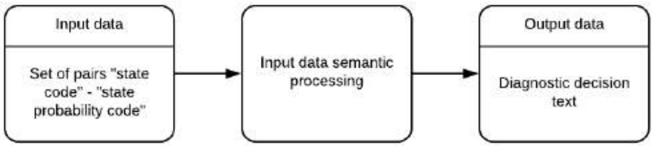


Figure 1. Operation scheme of ostis-system of automatic diagnosis

Consider some indications on described above scheme:

- state code – unique numeric identifier of state stored in ostis-system of automatic diagnosis knowledge base;
- state probability code – one of the following numeric values coding mentioned above state estimation scale: 1 - very low, 2 - low, 3 - small, 4 - high, 5 - very high;
- input data semantic processing – forming the diagnosis decision text based on semantic connections between states existing in knowledge base.

The algorithm of handing down a diagnosis decision appears as follows:

1) to sort the input estimates set in order of predefined priorities;
2) if it's enough data for handing down a decision, then make an interim decision for each estimate by using corresponding patterns;
3) if it's enough data for handing down a decision, then:
   a) to request additional data;
   b) if additional data is received, then return to step 2);
   c) if it's impossible to receive additional data, then hand down a decision of additional assessment requirement for current estimate;
4) to hand down an end-point decision (functional, prenosological, prednosological, nosological, permitting, limitative or prohibitive) or a decision of additional assessment requirement based on obtained interim decisions.

Consider the operation of ostis-system of automatic diagnosis on the following example. The estimates of two states are given as input data: Hepatitis C (state code - 1) and hepatoprotectors requirement state (state code - 2).

Some possible estimates of each state and based on them decisions are presented further on SCg language [2].

Memory state for small probability of both macroorganism states case is presented on Figure 2.

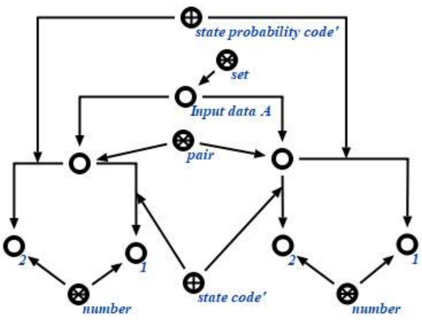


Figure 2. Input data for small probability of both states case

Decision of Hepatitis C clearance is handing down as the first state estimate is lower than high. The forming decision result is shown on Figure 3.

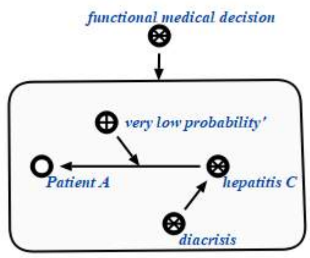


Figure 3. Output data for small probability of both states case

Input data for high probability of the first macroorganism state and small probability of the second macroorganism state case is presented on Figure 4.

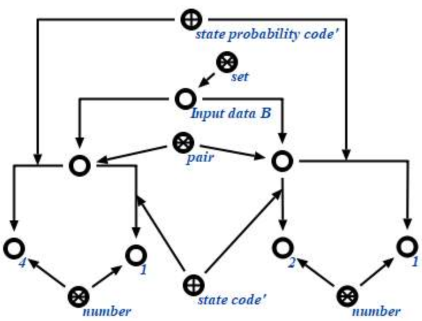


Figure 4. Input data for high probability of the first state and small probability of the second state case

Hepatoprotectors requirement estimate verifying is performing as Hepatitis C estimate is high. The second state estimate is lower than high in this way decision of Hepatitis C risk state is forming. Memory state as result of forming decision is shown on Figure 5.

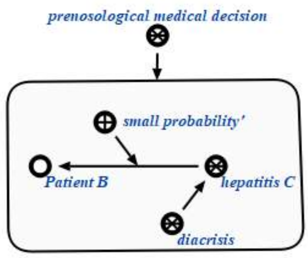


Figure 5. Output data for high probability of the first state and small probability of the second state case

High probability of the Hepatitis C and very high probability of the Hepatoprotectors requirement state case is presented on Figure 6.

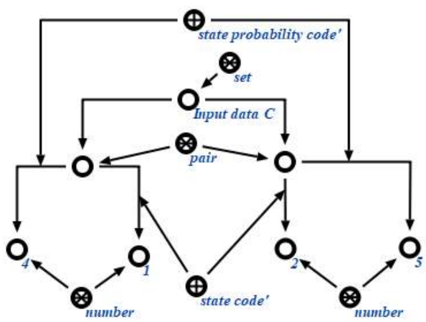


Figure 6. Input data for high probability of the first state and very high probability of the second state case

Hepatoprotectors requirement estimate verifying is performing as Hepatitis C estimate is high. The second state estimate is very high, consequently decision of Hepatitis C latency stage is forming. Result of forming described decision is shown on Figure 7.

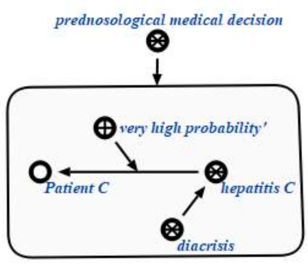


Figure 7. Output data for high probability of the first state and very high probability of the second state case

VI. Structure of ostis-system of automatic diagnosis knowledge base

Each ostis-system knowledge base is characterised by some highest level subject domain decomposition on more partial subject domains. As such, it's necessary to divide all knowledge presented in particular domain to thematic subject domain for designing a knowledge base of concrete ostis-system. Besides, it's necessary to determine the maximum class of researched objects, not maximum class of researched objects and researching relations set for each subject domain (if there are any). Maximum class of researched objects is a class of such entities described within subject domain that can't be outlined by means of more general entities on the scope of this subject domain. Subsequently, classes of entities that can't be subsumed as maximum classes of researched objects should be labeled as not maximum classes of researched objects [6].

A. Subject domains of ostis-system of automatic diagnosis knowledge base hierarchy

Subject domains determined within the medical diagnosis domain and their connections with subject domains from IMS knowledge base [2] are presented further on SCn language [2]:

1) Subject domain of medical diagnosis (medical diagnosis is considered as activity field in this case).

Subject domain of medical diagnosis
*<= particular subject domain**:
Subject domain of actions and tasks
$\ni$ *maximum class of researched objects'*:
medical diagnosis
$\ni$ *not maximum class of researched objects'*:
diacrisis
$\in$ *subject domain*
$\in$ *structure*

2) Subject domain of macroorganism states.

Subject domain of macroorganism states
*<= particular subject domain**:
Subject domain of temporal entities
$\ni$ *maximum class of researched objects'*:
macroorganism state
$\ni$ *not maximum class of researched objects'*:
- *urogenital system state*
- *organism system state*
- *parasitic system state*
- *cardiovascular system state*
- *excitatory system state*
- *bronchopulmonary system state*
- *osteoarticular system state*
- *fungal system state*
- *immune system state*
- *detoxification system state*
- *endoecological system state*
- *virous system state*
- *digestive system state*

• *bacteritic system state*
• *ENT-system state*

∈ *subject domain*
∈ *structure*

It should be noted that each organism system can has a set of states (among them are diseases) and a set of diagnostic decisions corresponding to them. It's possible to hand down a set of diagnostic decisions relating to each endoecological system, among them are decisions relevant to causation of concrete disease.

3) Subject domain of medical decisions.

Subject domain of medical decisions
*<= particular subject domain**:
    Subject domain of temporal entities
∋ *maximum class of researched objects'*:
  medical decision
∋ *not maximum class of researched objects'*:
  • *prohibitive medical decision*
  • *functional medical decision*
  • *limitative medical decision*
  • *nosological medical decision*
  • *permitting medical decision*
  • *prenosological medical decision*
  • *endoecological medical decision*
  • *differential medical decision*
  • *etiological medical decision*
  • *prednosological medical decision*
∈ *subject domain*
∈ *structure*

B. Family of ostis-system of automatic diagnosis knowledge base ontologies

The system of ontologies, i. e. system of subject domain concepts properties and interconnections descriptions, was designed for Subject domain of macroorganism states on the authority of approach to knowledge bases designing described in OSTIS Technology [6]. Mentioned system of ontologies includes the following ontologies kinds:

1) structural specification;
2) set-theoretical ontology;
3) terminological ontology;
4) ontology of tasks classes and solution methods;
5) logical ontology;
6) integrated ontology [6].

Consider some of determined ontologies kinds for ostis-system of automatic diagnosis knowledge base.

Structural specification is an ontology that describes the roles of concepts included into subject domain and connections between described subject domain and other subject domains.

Set-theoretical ontology is an ontology that describes the set-theoretical connections between concepts of described subject domain.

Fragment of Subject domain of macroorganism states structural specification is presented on Figure 8.

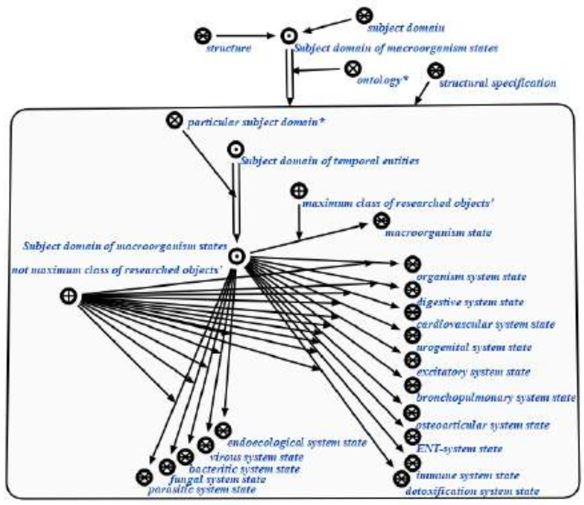


Figure 8. Fragment of Subject domain of macroorganism states structural specification

Fragment of Subject domain of macroorganism states set-theoretical ontology is shown on Figure 9.

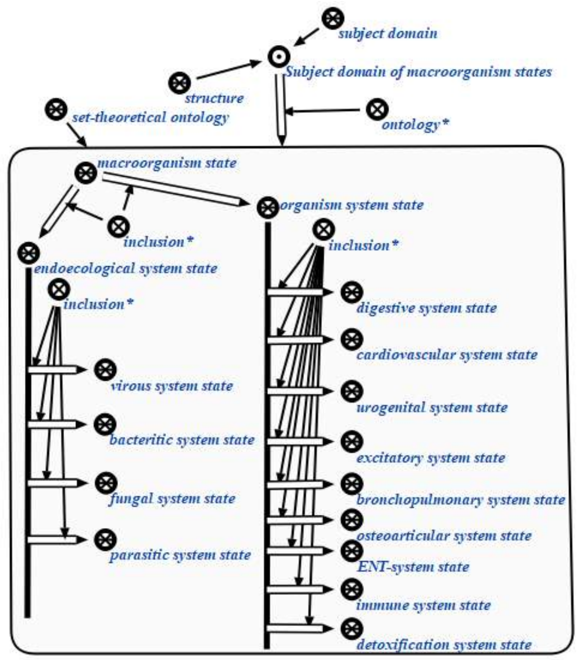


Figure 9. Fragment of Subject domain of macroorganism states set-theoretical ontology

VII. Structure of ostis-system of automatic diagnosis knowledge processing machine

Problem solving process in each ostis-system reduces to decomposition some common action to subactions which performing will lead to problem solving. Knowledge processing machine of each ostis-system represents collective of sc-agents aimed to performing the actions, which signs appear in unified semantic memory during the system work [5].

Hierarchy of sc-agents, which collective activity is aimed on end-point medical decision forming based on input set of states estimates analysis, was designed in terms of described above algorithm of handing down a diagnosis decision. Determined sc-agents hierarchy is presented further on SCn language.

Knowledge processing machine of ostis-system of automatic diagnosis
<= *abstract sc-agent decomposition**:
{
- *Abstract sc-agent of task for handing down a medical decision forming*
- *Astract sc-agent of input estimates set sorting*
- *Abstract sc-agent of input data sufficiency for handing down a decision checking*
- *Abstract sc-agent of interim decision forming*
- *Abstract sc-agent of task for additional data request forming*
- *Abstract sc-agent of additional data receiving*
- *Abstract sc-agent of end-point decision forming*

}

VIII. Conclusion

Principles of ostis-system of automatic diagnosis design by using OSTIS Technology are described in the work. Handing down medical decisions by solving different types of medical diagnosis problems is the purpose of proposed system. In future, it's possible to expand the capabilities of this system by not only considering a macroorganism as diagnosis object but any system that can be amenable to diagnosis whether it be a biological or technological system. Operating principles of the system, knowledge base structure top-levels and algorithms of forming decisions will stay the same on that assumption yet the application field of the system will significantly extend.

References

[1] T. A. Gavrilova, V. F. Khoroshevskii, "Bazy znanii intellektual'nykh sistem. Uchebnik [Knowledge bases of intelligent systems. Textbook]". Saint-Petersburg: Piter, 2001, 383p. (in Russian)

[2] IMS.OSTIS website. Available at: http://www.ims.ostis.net/ (accessed 2017, Dec)

[3] V. N. Rostovtsev, "Kontseptsiya kompleksnogo proekta "Telediagnostika" [The concept of complex project "Telediagnosis"]" in Vesti instituta sovremennyh znaniy [News of present knowledge institute], Minsk: Institut sovremennyh znaniy, 2014, №1, P. 64-67. (in Russian)

[4] Medical Spectral-Dynamic Complex website. Available at: http://www.kmsd.by/, http://www.kmsd.su/ (accessed 2017, Nov)

[5] D. V. Shunkevich, "Ontology-based design of knowledge processing machines" in Open semantic technologies for intelligent systems (OSTIS-2017). Minsk: BSUIR, 2017, P. 73–94.

[6] I. T. Davydenko, "Ontology-based knowledge base design" in Open semantic technologies for intelligent systems (OSTIS-2017). Minsk: BSUIR, 2017, P. 57–72.

[7] N. A. Sergantova, P. K. Strokov, "Osobennosti primeneniya sredstv avtomaticheskoy diagnostiki dlya vyyavleniya "Ostrykh" sostoyaniy patsiyentov [Features of application means automatic diagnostics for identification "Sharp" conditions of patients]" in Izvestiya YUFU. Tekhnicheskiye nauki [Izvestiya SFedU. Engineering sciences]. Taganrog: SFedU, 2012, №9, P. 134-139. (in Russian)

[8] V. E. Kuznetsov, "Predstavleniye v EVM neformal'nykh protsedur: produktsionnyye sistemy [Informal procedures representation in ECM: production systems]". Moscow: Nauka, 1989, 160p. (in Russian)

[9] V. N. Rostovtsev, "Resheniye problemy ranney diagnostiki [Solving the Problem of Early Diagnostics]" in Spravochnik vracha obschey praktiki [Guide of general practitioner]. Moscow: Panorama, 2016, №4, P. 10-15. (in Russian)

[10] V. V. Golenkov, N. A. Guliakina, "Semanticheskaya tekhnologiya komponentnogo proyektirovaniya sistem, upravlyayemykh znaniyami [Semantic technology of component design of systems, managed by knowledges]" in Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem [Open semantic technologies for intelligent systems] (OSTIS-2015). Minsk: BSUIR, 2015, P. 57–78. (in Russian)

[11] D. V. Smolin, "Vvedeniye v iskusstvennyy intellekt. Konspekt lektsiy [Introduction to artificial intelligence. Compendium of lectures]". Moscow: Physmatlit, 2004, 208p. (in Russian)

ПРИНЦИПЫ ПОСТРОЕНИЯ OSTIS-СИСТЕМЫ АВТОМАТИЧЕСКОЙ ДИАГНОСТИКИ

Ростовцев В. Н. (РНПЦ МТ),
Родионова О. С. (БГУИР)

Работа посвящена разработке интеллектуальной ostis-системы автоматической диагностики. В статье проанализированы известные подходы к решению задачи автоматизации деятельности врача по вынесению медицинского заключения, а также рассмотрены принципы построения базы знаний и машины обработки знаний для ostis-системы автоматической диагностики на основе семантических сетей с теоретико-множественной интерпретацией.

The structure of historical intelligent systems knowledge base

Hubarevich Nastassia
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
stasia@tut.by

Vityaz Siarhej
The Center for the Belarusian Culture, Language and Literature researches of the National Academy of Science of Belarus
Minsk, Belarus
vit@history.minsk.by

Grigyanets Romuald
The State Scientific Institution "The United Institute of Informatics Problems of the National Academy of Science of Belarus
Minsk, Belarus
griganec@bas-net.by

***Abstract*—The paper considers the approach to the formalization of historical knowledge on the basis of historical sources, which possess variability, incompleteness and unreliability of knowledge, OSTIS technology is used as a technological basis, and knowledge is represented in the form of semantic networks with a basic set-theoretic interpretation.**

***Keywords*—digital history, knowledge representation, methods of knowledge processing**

I. INTRODUCTION

In modern science there are most popular researches at the junction of two or more sciences. One of these areas includes the Digital history, which is gaining increasing development in foreign, Russian and Belarusian science.

Computer technologies came to the historical researches like the separate field in the 80s, together with the creation of the association "History and computing" in Britain, when the history reseachers began to realize and identify themselves separately from the general area "Computer and the humanities". The main accent in both areas was made on information support of the humanitarian researches and, in particular, historical researches. Gradually the foreshortening was changed to instrumental support of research, and using the mathematical and statistical methods in humanitarian research. And was fixed the new term "History computation".

Later, it became evident the insufficiency and limitations of statistical and mathematical research methods for humanitarian, and especially historical knowledge. First of all, this was due to the fact that these methods had an application mainly in economic and demographic research in the field of history. "Computing" field showed itself as excessively rigid," instrumental ", remote from interests and researches of the majority of historians [1]. Especially when reseachers began to use the computation methods and modern information technologies for analyzing weakly structured text sources, the need to develop another, source-oriented approach, which would take into account the specifics of historical sources with their "irregularities" and "blurring" of data was realized [2]. In this regard, there was a turn-out from methods to sources, which was marked by an increase in interest in the problem of creating information resources in the field of history. In science this direction was called Digital history.

Digital history is defined as a broad field of applied development for improving the work of historians - researchers and teachers - based on computer technologies, to facilitate users' access to digitized historical resources, to increase their visibility and visual representation, to expand the possibilities of working with data [3].

Currently, distinguish four most priority area s in the digital history evolution [4]:

- The area, which is connected with virtual reconstruction, spatial visualization and use of GIS technologies in historical research [5], [6], [7].
- The second one is associated with the creation of big online resources containing materials on history with open databases, with the possibility of third-party replenishment and editing [8]
- The third area is connected with usage interactive hypermedia technologies, nonlinear information presentation (interactive narrative). [9], [10].
- The fourth is associated with the resources of collective authorship (Web 2.0) elaboration.

The further evolution of each area poses the question of the historical knowledge representation and the methods of knowledge processing, considering the particular qualities of historical sources. In questions related to the historical knowledge representation it is necessary to pay attention to the following problems:

- The problem of incomplete, fuzzy and contradictory knowledge representation. For example, the different naming of the organization in different sources and even in the single one.
- The problem of the temporal characteristics of historical objects representation, such as the changes in tracing of streets, the street names, house numbering etc.
- Problem with the exact dates of the historical situations and events.
- Problem of the automation of the process of correlation the different pieces of information from various sources

with each other. For example the data of street names and its building numbering is contained in various sources, and we need to reconstruct all addresses of each building on the street during the whole time of its existence.
- The problem of the multilingualism of sources. So the sources on the history of the city of Minsk are presented in Russian, Belarusian, Polish, rarely German, French and Latin, and there is a need to information correlation.
- The problem of automatic knowledge base filling from the partially structured sources on natural language.

II. PROPOSED APPROACHT

The paper presents an approach to the formalization of information from historical sources. As a technological basis for the implementation,***OSTIS technology*** is used [11]. Systems based on this technology are called ***ostis-systems***, and they consist of a knowledge base that stores all information about the entities described, and knowledge processing machines, a team of agents that perform operations on knowledge.

The knowledge base of the system is a semantic network, where the vertices of the network represent concepts, and the edges indicate the relationship between concepts. And there are some different between the sign of entity(sc-element) and the identifier (name) its entity in various external languages. This separation allows, without changing the description of the entity itself and its connections with other entities in the system, to change the identifier language used only for interaction with the user of the system.

In the historical knowledge base the main identifier for the of entities assigned by the name of the first mention in the source, and to present the changes in entity names, the temporal decomposition is used. The variability of these names is formalized as synonymy. For example the organization BSSR Council of People's Commissars have as the the temporal parts the names "BSSR Supreme Soviet", "Supreme Soviet of the Republic of Belarus", " National Assembly of the Republic of Belarus". And the synonyms of the firs name will be Sovnarcom, SNK etc.

At present time in the historical knowledge base with the following degree of completeness, the following ***subject domains (SD)*** are identified and described:

- ***SD of artifacts*** describes all historically valuable and artificially created material entities as a result of a purposeful activity.
- ***SD of urban planning*** describes immovable monuments of history and culture.
- ***SD of persons and social communities*** considers a person and all arising from his activity communities of people.
- ***SD of ideas*** describes compiled on the basis of purposeful activity results.
- ***SD of historical actions and events***
- ***SD of historical sources***, which describes sources, the information from which formed the basis of the knowledge base.

The most of concepts in the system are the temporal entities, which have the temporal characteristics, such as duration in time, the time of the beginning and the time of ending.

All changes of entities and their relations with other entities are represented as situation – structure that contains at least one temporal entity. The combination of several temporal entities and the connections between them in the structure means that at some point in time all these entities in their interrelationship existed simultaneously.

Situations are divided into the following classes:

- ***One-time situation*** – the situations which mostly represented the moment of the emergence of some nonmaterial entity, which means that from this moment this entity in the form of some idea will always exist. For example, the situation of authorship between a person and his work;
- ***Completed situation*** – situation which had in the past the time of the beginning and ending also ;
- ***Non-completed situation*** is a situation that began in the past but has not been completed yet, such as the existence of some person or building.

The time of the situation existence can be set by:

- ***The time point***, which has discrete and arbitrarily exact values. For example, September 1, 1939 4.40 AM and September, 2 1945 4.02 PM – the begin and the end date of the World War II.
- ***The period of time*** – duration between two time points, such as in the sentences "Minsk City Theater was built in the XVII century and destroyed in the 60s of the XX century".

This approach to the formalization of knowledge in historical systems makes it possible:

- Eliminate duplication in the knowledge base of elements that denote the same entity;
- Present fuzzy time boundaries of the duration of events preserving the possibility of machine processing and analysis of information;
- Specify temporary data with any necessary degree of detail, in any chronological system, and be able to automatically translate and compare dates from one system of chronology to another;
- Describe the temporal properties of not only specific entities, but also whole situations with any necessary number of elements;
- Simply and flexibly make changes into information about entities. By introducing or deleting some information about entities, there is no need to rewrite the situation with the participation of this entity;
- Convenient search by association. That is, we can query the system for any time slices relative to any objects in the knowledge base, we can request specific types of relationships of a given entity with a certain class. For example, we can inquire about the state of some building

at a particular time, find out all the tenants in this time period, and then find out the other addresses of these tenants in the city.

Every situation in knowledge base by the relation source* connected with the specific document, where the information about the represented situation was given.

As sources of information are used: Memorial books of minsk province (1845-1917)[12], Minsk hadbookes (1926, 1967)[13], book "The Minsk main street" (1880-1940) [14] .

III. EXAMPLES OF IMPLEMENTATION

Consider concrete examples of the implementation of these solutions in the historical knowledge base of the city of Minsk.

The figures 1 and 2 demonstrates an example of the street temporal decomposition, where there are different names in different time periods are represented. The second figure demonstrate the formalization of the period of existence of the concrete street name.

The building has the following characteristics: geolocation (coordinates on the map), building number at different time periods, owners, residents of this building, the organizations in it, architectural style of the building at different time periods, type of building, construction material, architectural elements. The description can be supplemented with any other necessary information.

As can be seen in figure 3, the building called "Przheliaskovski House" from 1896 to 1912 had number 18, during the period 1913-1938. the building had number 14, and in the period from 1939 to 1951 it was number 12.

There is no need to restore and enter manually the building address. The system have the knowledge about the date of changing the street names and the date of changing the numbers of each building, and will be able not only restore the addresses of concrete building automatically for the entire period, was described in the system, but also to search for the building at the address in exact period of time.

The organization characteristics in the system are the structure, composition, belonging to the classes of organizations etc. Example on the figure4

Persons are characterized in the system by the following signs: date of birth and death, place of work, study, residence, positions held, participation in organizations, class affiliation, major achievements, authorship, etc.

The description of any concept of the real world in the system makes it possible to present other information about the object in the form of drawings, video fragments, etc. For a person, these can be his portraits, voice recordings, for buildings - plans, measurements, sketches and so on. Such representation of knowledge about the history of the city gives wide opportunities of knowledge processing.

The knowledge about the concrete time of the persons employee in the particular organization, we can automatically find out the name of the organization in any given period, the buildings it was located, the co-workers of the person in this organization. That gives wide opportunities not only for nonlinear presentation of information, but also broad opportunities for its analysis for the presence of implicit relationships between entities. It gives the possibility to find out all places in the city connected with any person in knowledge base. Agents will correlate the time of work of the person in the organization with the time of location of concrete organization in different buildings, all places of persons living and will give the answer, as on the figure 5.

IV. CONCLUSION

The proposed approach to the presentation of historical knowledge makes it possible to fully reflect information from sources, taking into account their incompleteness, inaccuracy and fragmentation, and also to present historical knowledge in a formal form accessible to machine processing.

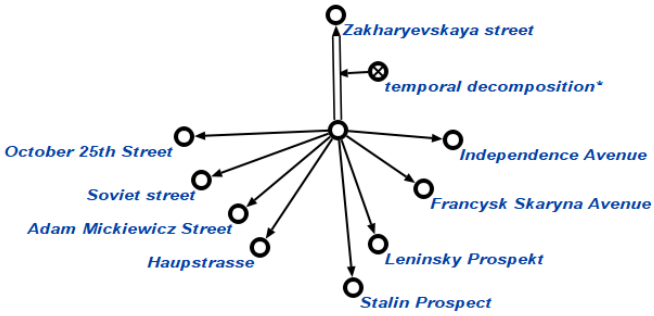


Figure 1. An example of decomposition of Zakharevskaya street.

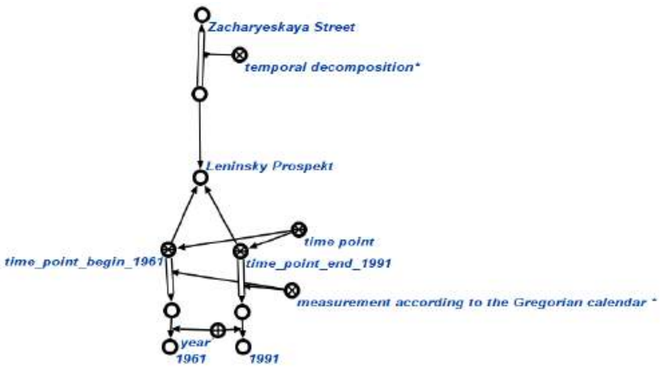


Figure 2. An example of dating a temporal part of the street.

REFERENCES

[1] Garskova I.M. Analysis of the historiography of historical informatics as a scientific.*Kharkiv historiographical collection*, 2010, vol.10, pp. 138-171

[2] Garskova I.M. Historical informatics and quantitative history: continuity and interaction. *Analysis and modeling of socio-historical processes*, Moscow, 2006, pp. 63-64.

[3] Borodkin L.I. Historical informatics: overload? *Bulletin of the Perm university*, 2011, Iss. 2(16), pp. 5-11.

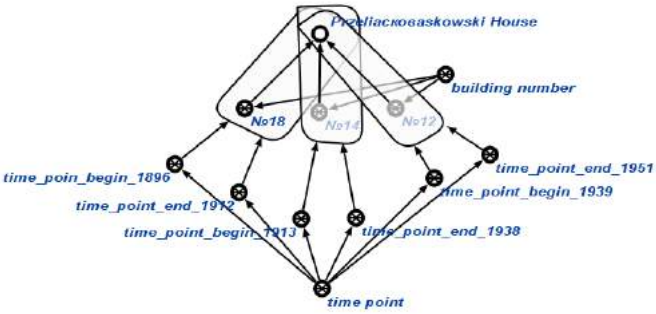


Figure 3. An example of dating a temporal part of the street.

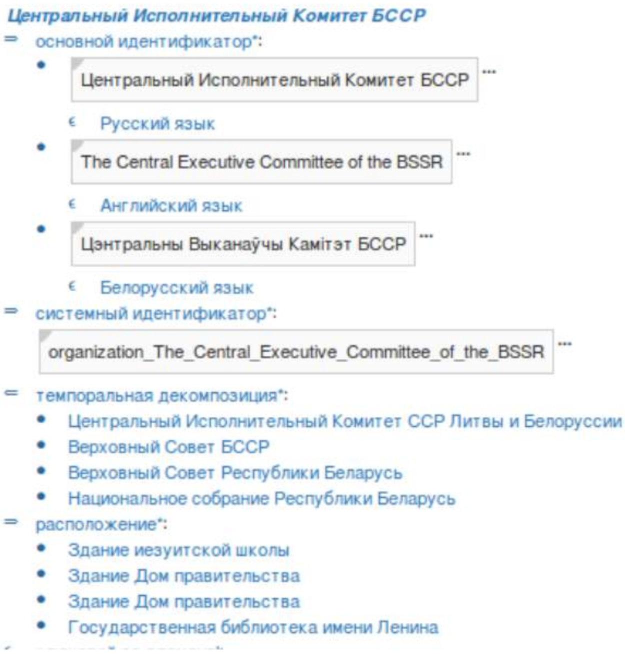


Figure 4. An example of dating a temporal part of the street.

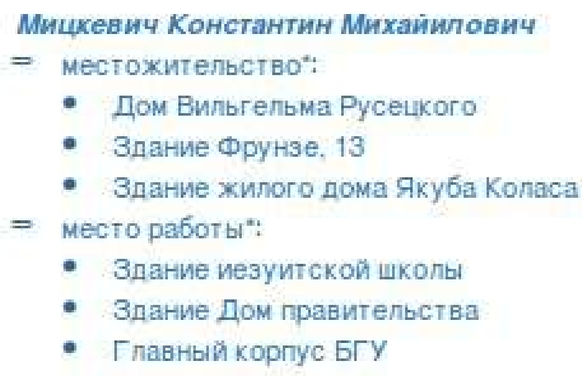


Figure 5. An example of dating a temporal part of the street.

[4] Arthur P. Exhibiting history: the digital future. *Journal of the National Museum of Australia*, 2008. Vol. 3, No 1

[5] Bryukhanova1 E.A., Rygalova M.V. Historical GIS on-line: a Review of Foreign and Russian Projects. *Journal of the altai state university*, 2016, No 2, pp.56-59

[6] Nemtinov V. A., Morozov V. V. Virtual modeling of objects of cultural and historical heritage with using gis technologies. *Bulletin of the tambov state technical university*, 2011, vol.17, No 3, pp. 709-714

[7] Ivakin Yan A., Potapichev Sergey N. Geochronology tracking — specialized gis-tool for history researches. *Historical Information Science*, 2016, No 1-2, pp. 3-11

[8] Makewell Terry The National Museums Online Learning Project Federated Collections Search: Searching Across Museum And Gallery Collections In An Integrated Fashion. Available at: https://www.museumsandtheweb.com/mw2008/papers/makewell/makewell.html

[9] The Labyrinth Project. Available at: https://dornsife.usc.edu/labyrinth/about.html

[10] The Big History Project. Available at: https://www.bighistoryproject.com/

[11] The IMS.OSTIS website. Available at: http://www.ims.ostis.net.

[12] Memorial books of minsk province (1845-1917) Available at: http://infobelarus.nlb.by/7n3119cd/7n3119cd.html

[13] Minsk. Brief handbooke Available at: http://minchanin.esmasoft.com/books/minsk67/

[14] Marakou L. The main street of Minsk (1880-1940). Book 1. Minsk, Publisher ‘Mastackaja litaratura”, 2012. 260P.

СТРУКТУРА БАЗ ЗНАНИЙ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ПО ИСТОРИИ

Губаревич А.В. *, Витязь С.П.**, Григянец Р.Б.***
Белорусский государственный университет информатики и радиоэлектроники
**Центр исследование белорусской культуры, языка и литературы НАН РБ
***ОИПИ НАН РБ
Минск, Беларусь

В работе рассматривается подход к формализации исторического знания на основе исторических источников, обладающих вариативностью, неполнотой и недостоверностью знаний. В качестве технологической основы используется технология OSTIS, и знания представляются в виде семантических сетей с базовой теоретико-множественной интерпретацией.

Assessing Student Learning Outcomes Using Mixed Diagnostic Tests and Cognitive Graphic Tools

Anna E. Yankovskaya*†‡§, Yury N. Dementyev‡, Artem V. Yamshanov§ and Danil Y. Lyapunov‡§
*Tomsk State University of Architecture and Civil Engineering, 2 Solianaya Sq., Tomsk, Russia
†National Research Tomsk State University, 36 Lenin Ave., Tomsk, Russia
‡National Research Tomsk Polytechnic University, 30 Lenin Ave., Tomsk, Russia
§Tomsk State University of Control Systems and Radioelectronics, 40 Lenin Ave., Tomsk, Russia
Email: ayyankov@gmail.com, dementev@tpu.ru, yav@keva.tusur.ru, lyapdy@gmail.com

Abstract—**In this paper, we provide an approach to assessing student learning outcomes by using mixed diagnostic tests. These tests represent an optimal compromise between unconditional and conditional components and facilitate the development of individual learning paths, which, in turn, would provide students with the opportunity for self-guided learning. To construct individual learning paths, we apply an intelligent learning and testing system. Thus, each student becomes able to predict their learning outcomes following the respective learning path designed. In addition, we describe a cognitive graphic tool the 2-simplex prism to cognitively visualize the results of student learning assessment. We assume that our approach can be used when assessing student learning outcomes within any subject and propose applying this approach as a tool to enhance both student and teacher activity.**

Keywords—**intelligent learning and testing system, cognitive graphic tools, blended learning, mixed diagnostic tests, semantic network, 2-simplex prism, e-learning course, student-oriented approach, prediction of learning results**

I. Introduction

Among the existing approaches to student activity assessment, we outline using cognitive graphics as a promising avenue for education research. The importance of cognitive graphic tools in education has been pointed out by an extensive body of literature [1]–[3]. A group of researshers in [1] highlighted the fact that using cognitive graphics for assessment can enhance student learning performance by giving an opportunity to design a learning path for each student. Moreover, effective tools of cognitive assessment contribute to learning content improvement, revealing successful ways for teachers to design learning activities.

Axelrod [4], Pospelov [5], Zenkin [6], Kobrinskiy [7], [8], Albu and Horoshevskiy [9], Kolesnikov et al. [10] and Yankovskaya et al. [1] have contributed significantly to the development of cognitive graphic tools for different problem areas. In 1996, Yankovskaya [11] introduced a versatile means of assessing student knowledge of a particular topic by using mixed diagnostic tests (MDTs). These tests represent an optimal compromise between unconditional and conditional components. We will illustrate the essentials of MDTs design in the next section.

In their research, Kulikovskikh, Prokhorov and Suchkova [12] outlined the following essential learner characteristics to be assessed: 1) learning style, 2) background, 3) motivation, and 4) level of knowledge. The researchers also provided a versatile tool to assess test results and reveal the probability of guessing for each student. Moreover, using fuzzy learning performance assessment proposed by Proknorov and Kulikovskikh [13] facilitates the enhancement of student learning and the development of cognitive graphics.

Nevertheless, as well as positive findings that have accrued, taking into account the specificity of e-learning, an increasingly critical research agenda is emerging. For example, different learning activities require different assessment approaches. Moreover, both students and teachers need to monitor systematically all the aforementioned learner characteristics and visualize them in a single reference frame by cognitive graphic tools.

We wish therefore to focus on issues associated with the cognitive visualization of student learning assessment. Furthermore, we present an Intelligent Learning and Testing System (ILTS) that guides students through the learning module and provides the cognitive visualization of their activity.

II. Construction of Mixed Diagnostic Tests

We constructed MDTs [11] using a syllabus for the discipline of “Power Electronics”. Briefly, in the course of training, students perform learning activities and do these tests. At first, students take an unconditional diagnostic test to reveal their initial level of knowledge. The questions in such a type of test are randomly presented to a learner to estimate the basic knowledge of a subject. Then, if a student has passed the unconditional diagnostic test, they take a conditional diagnostic test. In this type of test, each subsequent question depends on the answer to the previous one. Thus, the result of the test strongly depends on the learner decision.

We represent the construction of MDT on a graph diagram (Fig. 1). In the upper part of Fig. 1, you can see a table with the rows denoting the questions of the unconditional diagnostic test. The number of answers to each question corresponds to the columns of the table. In this case, n is the number of

The research is funded by the Russian Foundation for Basic Research, Project Number 16-07-00859a, and by Competitiveness Enhancement Program Foundation of the Tomsk Polytechnic University, Project Number TPU CEP_IPE_97\2017.

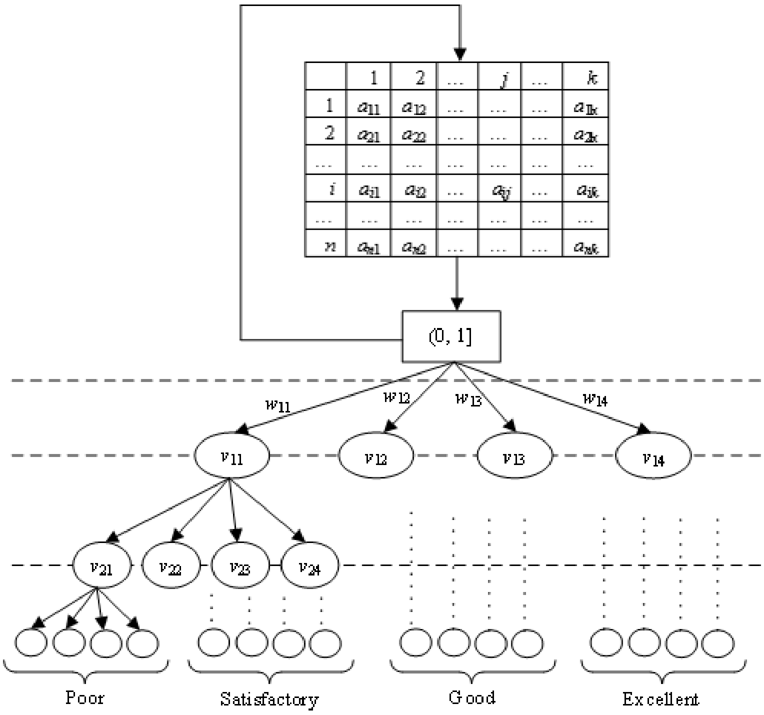


Figure 1. Mixed diagnostic test construction.

questions, k is the maximum number of various answers to each question ($k \geq 2$). Some cells of the table can be empty. The element of the table $a_{i,j}$ is the weight of the j-th answer to the i-th question, $0 \leq a_{i,j} \leq 1$.

Each edge on the diagram (Fig. 1) matches the weight $w_{i,j}$, which corresponds to the complexity of a question in the conditional diagnostic test. Herein, i is the level number, j is the number of the edge. A test question corresponds to each node of the graph $v_{i,j}$, where i is a level number and j is a node number at the current level.

After completing the test, each student gets their result in the form of a conventional grade (poor, satisfactory, good, excellent) and visually observes the overall evaluation of their knowledge using cognitive graphics.

III. Essentials of the Intelligent Learning and Testing System

To introduce the automation of learning and cognitive assessment, we have been developing the ILTS for more than ten years. Here, we use a semantic web to represent the ILTS that supports a student within each learning module in the discipline of "Power Electronics" (Fig. 2).

Using the ILTS, we subdivided the student learning within each module into five subsequent steps.

- Step 1. The students learn all the necessary topics provided by the learning module, which include video lectures with an interactive multimedia content. To create the content, we use the module that transforms the knowledge into the teaching materials using the knowledge database [14].
- Step 2. The students take the MDT and obtain certain results. The tests are constructed by using the knowledge database as well as teacher expert knowledge.
- Step 3. The pattern recognition module converts the MDT results into the assessment using a cognitive graphic tool the 2-simplex prism. We will represent this tool in the subsequent section.
- Step 4. The assessment results are recorded in the outcomes database.
- Step 5. We interpret the assessment using learning outcomes interpretation module. At this stage, the teacher or an expert explains each student how to use the 2-simplex prism for their effective self-evaluation. This helps to reveal student strengths and weaknesses in particular topics of the module.

Student actions may be twofold after the above steps. Thus, if a student needs help with some topics of the module, which is revealed by the supplementary module, they may repeat the corresponding part of the course for better comprehension. Provided a student is satisfied with their learning outcomes, they may go to the next module.

Each learning step is essential for students to reveal and develop their individual approach to comprehending a particular module and spread this approach onto the subsequent learning modules. For this reason, the ILTS records each student step and creates a learner action card (LAC) [14]. The LAC is available to both the student and the teacher so that they

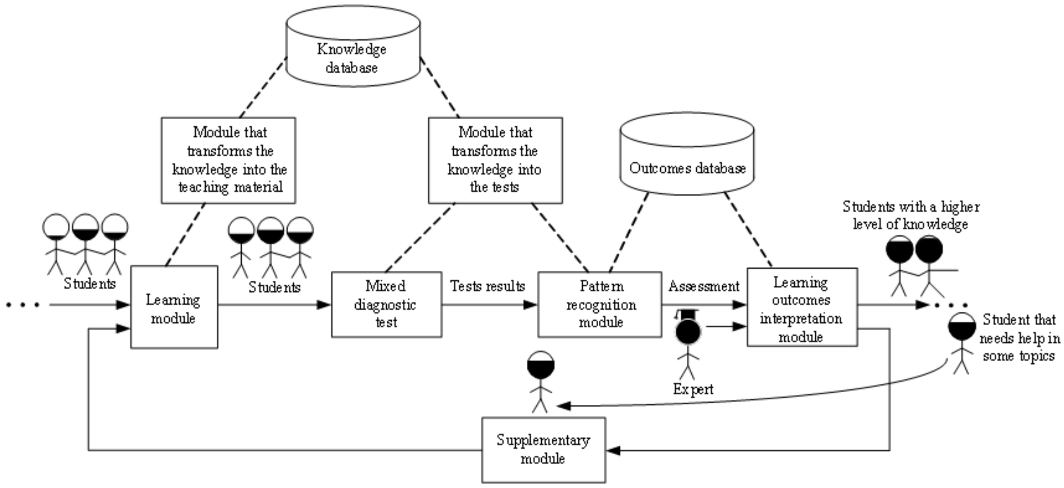


Figure 2. A semantic web illustrating the intelligent learning and testing system within one module of the discipline under study.

could analyze the student knowledge gaps and develop an appropriate learning path for better performance.

IV. Results and Discussion. Interpretation of Learning Outcomes

In this section, we apply the approach to the usage of the 2-simplex prism as a tool to interpret student learning outcomes provided by the ILTS. The prism was first entirely described in [1]. We used the advantages of this tool to assess student performance in the discipline of "Power Electronics". An example of student learning paths designed using the 2-simplex prism is given in Fig. 3.

As shown in Fig. 3, the four points (the small circles of different colors) lie within the 2-simplexes (cross-sections of the 2-simplex prism). These points represent the results of the four tests. The prism faces correspond to grades: 1) "excellent", 2) "good", and 3) "satisfactory". The height from the point to a face within each of the 2-simplexes denotes the proximity degree of the assessment to the particular grade. The distance between two subsequent 2-simplexes indicates the interval between two tests. The dashed black line within the 2-simplex prism shows the evolution of student knowledge level based on the test results at time T_1, T_2, T_3 and T_4.

We observe in Fig. 3 that at time T_1 the student obtained the grade between "satisfactory" and "good", yet close to the level of "satisfactory". Then, having analyzed the test results, the student set a goal of improving the performance and developed a plan on how to achieve this goal (e.g. by taking additional classes or using supplementary tutorials). This led to better assessment results after the 2nd test at time T_2, when the grade was close to "good". Subsequently correcting the individual learning plan, the student finally was able to reach the goal. Thus, at time T_3, the grade was between "good" and "excellent", whereas at time T_4 it approached the level close to "excellent".

We note that we can use another grade system if a student encounters unexpected problems (e.g. health problems, moving, accident, etc.), that could affect significantly their learning performance. In this case, we should use the following grades: 1) "poor", 2) "satisfactory" and 3) "good". Moreover, we are able to switch from one grade system to the other in the course of training depending on the student performance. This means that we use two 2-simplex prisms taking into account the current student grade.

The efficacy of using the 2-simplex prism in the course of learning could be attributed to its versatility and cognitive properties. Although here (Fig. 3) we may observe a linear dependence of the assessment on time, further research is needed to confirm this idea and reveal the peculiarities of the cognitive process over time.

Conclusion

Cognitive visualization of a grade provides not only the assessment of student learning activities, but also sound information on how to achieve that grade and justify this decision. Therefore, each student has the opportunity to take appropriate actions towards the improvement of the grade on a particular topic of the discipline.

The LAC created by the ILTS alleviates revealing student knowledge gaps and helps to eliminate them on time.

The linear dependence of the assessment on time can give us a tool to predict student learning outcomes in the course of their training. This prediction involves using the first order polynomial and is based on the previous as well as the current grade.

In this paper, we have proposed a cognitive way to represent numerical data (grades) by associating a grade with the student knowledge level and visualizing it in a 2-simplex prism. This will presumably be the topic for our further research.

It is noteworthy that we also may use a letter grade assessment to describe student learning performance (A, B, C, D, F, etc.), which is especially relevant when developing learning courses using Bologna declaration [15] to ensure comparability in the standards and quality of higher education

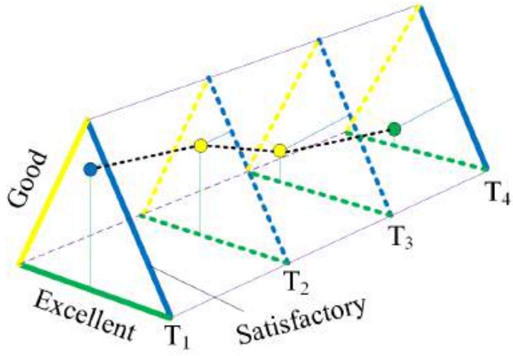


Figure 3. Learning outcomes assessment using cognitive graphic tool the 2-simplex prism.

qualifications. Consequently, the cognitive representation of the letter grade assessment system may also be a promising avenue in our future research.

References

[1] A. Yankovskaya, A. Yamshanov, N. Krivdyuk 2-simpleks prizma kognitivnoe sredstvo prinyatiya i obosnovaniya reshenii v intellektual'nykh dinamicheskikh sistemakh [2-simplex Prism is a Cognitive Tool for Decision Making and Its Justification in Intelligent Dynamical Systems]. Mashinnoe obuchenie i analiz dannykh [Machine Learning and Data Analysis], 2015, vol. 1, no 14. pp. 1930-1938.

[2] J. Dirksen Design for how people learn. New Readers, 2012. 272 p.

[3] L. LeFever The Art of Explanation. Wiley and Sons, 2012. 378 p.

[4] R. Axelrod The structure of decision: Cognitive maps of political elites. Princeton University Press, 1976. 422p.

[5] D. Pospelov Kognitivnaya grafika — okno v novyi mir [Cognitive graphics is a window to a new world]. Programmnye produkty i sistemy [Software Products ans Systems], 1992, vol. 2, pp. 4–6.

[6] A. Zenkin Znanie-porozhdayushchie tekhnologii kognitivnoi real'nosti [Knowledge as a generating technology of the cognitive reality]. Novosti iskusstvennogo intellekta [Artificial Intelligence News], 1996, vol. 2, pp. 72-78.

[7] B. Kobrinskiy K voprosu ucheta obraznogo myshleniya i intuitsii v ekspertnykh meditsinskikh sistemakh [On a question of accounting for creative thinking and intuition in expert medical systems]. Sbornik trudov 5-oi natsional'noi konferentsii s mezhdunarodnym uchastiem "Iskusstvennyi intellekt-96" [Proceedings of the 5th National Conference with International Participation "Artificial Intelligence – 96"], 1996, vol. 2, pp. 110–117.

[8] Vizual'nyye obrazy-simvoly v intellektual'nyh sistemah [Visual patterns-simbols in intelligent systems], 14 Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiyem [14the National Conference on Artificial Intelligence with International Participation], 2014, pp. 189-196.

[9] V. Albu, V. Horoshevskiy COGR - sistema kognitivnoy grafiki. Razrabotka, realizatsiya i primeneniye [COGR-a system of cognitive graphics. Developing, realization and application]. Izvestiya Akademii Nauk SSSR. Tehnicheskaya kibernetika [Bulletin of the Russian Academy of Sciences. Technical cybernetics], 1990, V.5, pp. 105-118.

[10] A. Kolesnikov, S. Listopad, V. Danishevskiy, K.Suvorov Tehnologii vizual'nogo upravleniya i kognitivnoy grafiki v funktsional'nyh gibridnyh intellektual'nyh sistemah podderzhki prin'atiya resheniy [Technologies of visual control and cognitive graphics in functional hybrid intelligent systems for decision-making support] Sbornik trudov 3 Vserossiyskoy Pospelovskoy konferentsii s mezhdunarodnym uchastiyem "Gibridnyye i sinergetichestiye intellectual'nyye sistemy. 2016. pp. 166-173.

[11] A. Yankovskaya Design of Optimal Mixed Diagnostic Test With Reference to the Problems of Evolutionary Computation. Proceedings of the 1st International Conference on Evolutionary Computation and Its Applications, 1996, pp. 292-297.

[12] I. Kulikovskikh, S. Prokhorov, S. Suchkova Promoting collaborative learning through regulation of guessing in clickers. Computers in Human Behavior, 2017, vol. 75, pp. 81–91.

[13] S. Prokhorov, I. Kulikovskikh Fuzzy Learning Performance Assessment Based on Decision Making Under Internal Uncertainty. Proceedings of the 7th Computer Science and Electronic Engineering Conference, University of Essex, 2015, pp. 65-70.

[14] A. Yankovskaya, Y. Shurygyn, A. Yamshanov, N. Krivdyuk Opredelenie urovnya usvoennykh znanii po obuchayushchemu kursu, predstavlennomu semanticheskoi set'yu [Determining of the Knowledge Level Acquired Using the Learning Course Represented by a Semantic Web]. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2015, pp. 331–338.

[15] The Bologna Declaration of 19 June 1999. Available at: https://en.wikipedia.org/wiki/Bologna_declaration(accessed 2017, Dec).

Оценка результатов обучения студентов с применением смешанных диагностических тестов и средств когнитивной графики

Анна Ефимовна Янковская[1,2,3,4], Юрий Николаевич Дементьев[3], Артем Вячеславович Ямшанов[4], Данил Юрьевич Ляпунов[3,4]

[1]ТГАСУ; [2]НИ ТГУ; [3]НИ ТПУ; [4]ТУСУР.

В работе предлагается подход к оценке результатов обучения студентов с использованием смешанных диагностических тестов. Данный вид тестов представляет собой оптимальное сочетание безусловной и условных составляющих, что позволяет разрабатывать индивидуальную траекторию обучения, которая, в свою очередь, обеспечивает возможность самообучения. Для построения индивидуальных траекторий обучения применяется интеллектуальная обучающе-тестирующая система. Таким образом, каждый студент имеет возможность прогнозировать результат обучения в соответствии с построенной индивидуальной траекторией обучения. Кроме того, в работе описано средство когнитивной графики 2-симплекс призма для когнитивной визуализации результатов оценки студентов. Предлагаемый подход может быть использован для оценки результатов обучения студентов в рамках любой дисциплины. Этот подход полезен как для студента, так и для преподавателя с целью повышения качества обучения.

Certain Approaches to Automated Designing of Competence-Oriented Models for Knowledge Engineers Using the Tutoring Integrated Expert Systems

Rybina G.V., Fontalina E.S., Sorokin I.A.
National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)
Moscow, Russian Federation
galina@ailab.mephi.ru

***Abstract*—The development of engineering knowledge led to the emergence of new professions, which are widely popular professional competence and personal qualities. This paper provides a methodological and technological experience in automated construction competence-oriented models specialists in the field of knowledge engineering, in particular, specialists of the profession "system analyst" with the use of tutoring integrated expert systems.**

***Keywords*—tutoring integrated expert systems, competence-oriented model of a specialist, engineer knowledge, student model, AT-TECHNOLOGY workbench.**

I. Introduction

Implementation of advanced concepts of transition to new professional tutoring and training computer-aided technologies for IT professionals implies a flexible common use of professional educational standards and the automated designing technology for the intelligent tutoring systems (ITS), in particular, tutoring integrated expert systems (IES) of various architectural types [1], [2]. It is ITS architectures that procure the advanced tools of the students' intelligent tutoring, monitoring and testing and help shape the competence-oriented models for future professionals [1] as the ultimate result. It is noteworthy that the competence format of the new Federal State Educational Standard for Higher Professional Education [3] envisages that the professional education quality should be assessed based on the graduate's competencies meaning the integral result shown after the education plan completion. The young professional's competencies shall enable him or her to successfully work in the selected occupation area and acquire social, personal and general cultural qualities that promote his/her social mobility and sustainability on the labor market.

In terms of the European Educational System, the notion of "competence" includes, besides the cognitive, operational and technological component, incentive, ethical, social and behavioral components that determine the system of the graduate's values as well [4]. Thus, as concerns the knowledge, skills and capacities widely used in the current Federal State Educational Standard for Higher Professional Education, the competence-oriented models for professionals of different levels have a comprehensive, integral nature that includes the set of knowledge, skills, capacities and social and personal qualities [5]. Accordingly, graduates' competencies improve in case of gradual completion of bachelor and master educational programs.

The key problem is to make sure that the professional standards, in this case IT standards, will not remain a certain methodical basis, but will be pro-actively included in the automated shaping of future professionals' competencies that show what the student will know (i.e. the student's theoretical training) and be able to do (i.e. ability to leverage theoretical knowledge in addressing practical tasks) upon completion of a certain educational plan (direction) and how well he or she is able to apply the knowledge, skills and personal qualities gained at the university for successful professional activity.

Similar experience has already been gained in the Cybernetics Department of MEPhI when training professionals in Applied Mathematics and Informatics, which is currently pro-actively used as part of the Software Engineering direction. This was facilitated by the research and development project of creating the intelligent technology for building a broad TIES category, including tutoring IES and web-IES [1], [2], [6].

The problem-oriented methodology of TIES designing [1], [2] and its supporting unique workbench type tool, AT-TECHNOLOGY workbench, enable the development of state-of-the-art IES and web-IES with advanced intelligent tutoring, monitoring and testing tools for students, through designing and software support to the individual competence-oriented models of students (with the view to the personality's psychological profile), adaptive tutoring models and explanations, models of course/discipline ontologies. etc. The main provisions of the problem-oriented technique and the description of functional options of various versions of AT-TECHNOLOGY workbench are available in several monographs and numerous papers, e.g. [1], [2], [6] etc.

In order to implement intelligent tutoring based on designing and use of tutoring IES and web-IES in the training

[1]The work was performed with the Russian Foundation for Basic Research support (Project No. 18-01-00457)

process, the professional standards for Information Technologies [5] were rather efficiently used as the basic information methodical resources when designing professional competence models, in particular, for such occupations as “programmer”, “system analyst”, “IT system specialist”, “system architect” etc.

Such occupations as “system analyst”, the demand for which is coming closer to that for programmers in the contemporary market for high information technologies, should be specially designated among the listed professions. At present, these professionals in knowledge and technologies of designing intelligent systems are called differently: knowledge engineers, knowledge analysts, cognitologist engineers, task setters, and much attention is paid to their training. For instance, the systemic view of training the knowledge engineers is available in [7], [8] where the authors' experience in training knowledge engineers and business analysts in the last 15 years is summarized. Many interesting findings on this problem are available in [9], [10], [11], [12], [13], [14] and other papers.

In general, as the experience of development and use of tutoring IES and web-IES in training suggested, the following objectives are the critical in shaping the professional and general competences are:

- Selective sampling at each tutoring stage (bachelor, master) of the knowledge, skills and capacities that the students must gain (applied course/discipline ontologies, generalized ontologies of individual training fields are used);
- Improvement of the monitoring and testing techniques, both for the purpose of shaping of current competence-oriented models for students and upon tutoring completion (web-testing of the students with generation of options based on the genetic algorithm is used);
- Efficient accounting for personal features of the students in selecting and shaping of the learning strategies and impacts, including the development of special corrective training impacts intended to develop the student's individual personal features (the students' psychological testing findings together with different types of learnings interactions are used);
- Use of additional (repeated) tutoring based on the identified gaps in knowledge and skills etc. (applied in the aggregate of tutoring relations for different students' clusters).

As concerns the information required for shaping the social and personal competencies (from the group of general competencies) that take into account the students' personal features, here we can partially make use of the information provided in professional standards in the Self-Development job description for each specialization [3]. In addition, in order to identify the personal features, there exists a great number of psychological tests, questionnaires, special websites etc. For instance, in the context of IES and web-IES [1], [2], [6], the option of detecting approx. 20 personal qualities and their correlation with an individual tutoring model was implemented. The main problem here is to find and select the expert information that signals the degree, in which each particular competence is displayed for each of the personal features.

It is noteworthy that there is no general classification of competencies now, but the designation of professional and versatile competencies is a generally acceptable point of view. Further specification depends on the specific features of the profession, the traditions of the university that trains professionals in this domain and other particular features.

This paper is intended to review the methodical and technological experience in automated designing and use of tutoring IES and web-IES of competence-oriented models for future knowledge engineers (“system analyst” profession).

II. The dynamic designing of competence-oriented models for future professionals based on the review of the monitoring of IELS students' functioning processes

Tutoring IES and web-IES developed in the laboratory of “Intelligent Systems and Technologies” of the Cybernetics Department of MEPhI have been pro-actively used since 2008 for automated support to basic courses and disciplines in such fields as Applied Mathematics and Informatics and Software Engineering, in particular: Introduction in Intelligent Systems, Intelligent Dialog Systems, Dynamic Intelligent Systems, Designing the Knowledge-Based Cybernetic Systems, Modern Intelligent System Architectures, Intelligent Information Systems.

For all of these courses and disciplines, the applied ontologies were implemented and are dynamically supported using AT-TECHNOLOGY workbench basic tools. The ontologies jointly form the generalized ontology of Intelligent Systems and Technologies as the tutoring and methodic basis for training knowledge engineers. Significant methodological and technological experience was gained in automated maintenance of a significant number of individual models for the students in the above disciplines (over 2,500 models) and the appropriate tutoring models, the joint analysis of which enables forecasting so-called “perfect” model of a young professional, in particular, a systemic analyst (knowledge engineer).

The possibility of such forecast implementation is largely determined by the particular features of development and use of tutoring IES and web-IES related to automation of virtually all processes arising during tutoring and control of students' knowledge/skills. The entire information about the students, course/discipline topics, academic progress, students' monitoring results, individual recommendations based on the academic progress etc. are in the single environment and at any time available to the student and/or the academic progress supervisor, which is achieved by special monitoring tools for IELS students' functioning process.

Pursuant to [2] and other papers, IES functioning is monitored from two angles. One of them is related to the role and part of tutoring IES in terms of tutoring management at the university, i.e. the use of IES for support of basic tutoring stages: holding of classes (lectures, workshops, laboratory work), holding of regular inspections, both during tutoring and at tutoring control points envisaged in the curricula of a certain course/discipline, as well as control efforts as part of credits and exams.

Another aspect is the review of a set of functional objectives typical of intelligent tutoring. Addressing the basic intelligent tutoring problems has been reviewed in various papers [2], [6] etc. many times; so, let's note that, in the context of summarized ontology of Intelligent Systems and Technologies and shaping the single ontological knowledge and skills space, we have managed to virtually implement the full set of functional objectives typical of the intelligent tutoring technology, namely [2], [6] etc.:

- Individual planning of the method to study the specific training course (specification of the personal tutoring trajectory/strategy based on the course/discipline ontologies, customized monitoring and identification of "gaps" in students' knowledge and skills, enhancement of individual tutoring with reference to the student's psychological profile, etc.);
- Intelligent analysis of solutions to training objectives (simulation of reasonings of the students who address training objectives of different type, in particular, using non-formalized techniques; identification of error types and reasons for their manifestations in knowledge and skills, instead of stating them; feedback via dynamic updates of the students' knowledge and skills; forecast of exam grades etc.);
- Intelligent decision support (use of the conventional expert system (ES) technology and IELS technology for intelligent support at each stage of addressing tutoring tasks, including the expanded explanations of How? and What? type, selection of the solution options, prompt of the next solution stage etc.)

Thus, the monitoring of tutoring IES and web-IES functioning in this case is associated with "tracking" and review of all processes of shaping the student's individual model for each student for a particular discipline by identifying the current knowledge/skills level using web testing and other methods, as well as by shaping the psychological profile of the student's personality as an important component of the student's model.

It is noteworthy that, according to the problem-oriented methodology, the dynamic comparison of web testing findings with the respective fragment of applied course/discipline ontology forms the basis of the approach to designing the student's current competence-oriented model. The result is so-called "problem zones" [1], [2] in students' knowledge in individual sections/subsections and designing of the current competencies that jointly reflect the state of the student's model not only in terms of knowledge level but also by shaping the conceptual and technological link with the elicitation of the ability to solve certain types of educational unformalized objectives recommended in [15] or educational items on knowledge engineering, e.g. [8]. It is also necessary to constantly draft the lists of students (cohorts) with high and/or low knowledge/skills, conduct systemic statistical data processing and also generate current and final reports (lists) for departments and dean's offices.

The final term logs that reflect the students competence-oriented models contain comprehensive information on the students: their grades received during control efforts (exams, credits, tests) intended to elicit the level of knowledge and skills, the current professional competencies, information on the psychological testing passing, information on independent work, the final grade forecast, and include the actual grade received in the exam (the logs are issued for all students trained in a particular course/ discipline).

Analytical and statistical processing of tutoring IES application efficiency is critical for shaping the future professional's model. By introducing special parameters describing an individual student and a certain cohort (cluster) of students These parameters were derived expertly, based on the review of a rather representative data volume (approx. 2,000 students' models), and largely target the basic structure of the student's model [1], [2], the components of which are: student's knowledge model, student's skills model, psychological profile, competencies model and other components)

The experience suggests that the parameters (indicators) shaped as a result of [2], [6] were the most sought of in terms of designing the competence-oriented models for future professionals.

- Review of "problem zones" for each student for particular courses/disciplines and their clusterization;
- Individual tutoring planning (typology and succession of training impacts; influence of training impacts on knowledge upgrading; search for the most efficient training impacts);
- Calculation of the correlation between the current knowledge and skills levels for the appropriate course/ discipline topics
- Account for the student's psychological profile (personal degree of achievement of target competencies for specific courses/disciplines etc.)
- Exam grade forecast based on academic progress (review of the reasoning when addressing particular educational and training tasks)

Besides, use a whole series of parameters for information processing for the entire students' cohort (group, flow, etc.) namely: aggregate analysis of "problem zones" for specific courses/ disciplines and their clusterization; estimate and clusterization of individual tutoring plans for specific courses/ subjects; exam session result forecast (the link between knowledge and skill level and exam grade by course; review and clusterization of students' psychological types, etc.)

Now let's look at practical examples of designing the "perfect" model of a graduate specialist in systemic analysis field (knowledge engineer) according to the base competencies of the current Federal State Educational Standard for Higher Professional Education in a greater detail.

III. Particular features of training knowledge engineers based on use of tutoring IES

The successes of knowledge engineering in development of models and methods for knowledge transfer from different knowledge sources into a software, called the expert system (ES), and in a wider sense, the knowledge-based systems (KBS) [2], [6], [8], [15] resulted in a new occupation that requires not only professional competencies but also individual personal features. Thus, training of knowledge engineers implies not only gaining professional knowledge, skills and capacities in development (KBS) (ES) but also without regard to his or her psychological profile.

Let's provide some examples from the experience of practical application of tutoring IES and web-IES in the training process at MEPhI, with the emphasis on the technology for automated shaping of the competence-oriented model for a knowledge engineer, with reference to the set of training plans and the generalized ontology of Intelligent Systems and Technologies in Software Engineering training domain.

A. Shaping the specialists' professional competencies

According to the Federal State Educational Standard 3+, the two following competencies are used as the base ones for training knowledge engineers:

- PK-1 is the formalization capacity in his or her subject area, with the view to the limitations on study methods in use;
- PK-2 is the readiness to use methods and instrumental means of study of professional business items

Achievement of these target competencies is facilitated by the common ontological space of knowledge and skills formed out of applied course/discipline ontologies of several tutoring IES and web-IES [2], [6]. It is noteworthy that the general competencies model that is the component of the basic ontology model in the shape of the semantic network [2], in the applied course/ discipline ontologies, is represented as the hierarchy of the discipline/ problem-oriented private competencies (with weight ratios) that reflect the methods of teaching some specific courses.

Table I shows the examples of specification of basic target competencies of PK-1 and PK-2 according to the ontology of the Introduction to Intelligent Systems course based on [15].

Table I
Examples of specification of target competencies PK-1 and PK-2

PK1	Ability to formalize his or her subject area, with the view to the limitations on study methods in use
PK11	Know and be able to use the systemic analysis methods to assess whether the intelligent system technology is acceptable or not
PK12	Know and be able to select the knowledge presentation models for designing specific intelligent systems
PK13	Possess the skills in simulating reasoning and building-up modern problem solvers (output mediums) for intelligent systems
PK14	Know the principal types of non-formalized problems and be able to design models and methods for solving non-formalized problems of different types
PK15	Know the methods of gaining knowledge from different knowledge sources (experts, texts, databases) and be able to apply them in practice
PK16	Know and be able to apply the modern methods for intelligent systems (imitation, evolutionary, neuro-network, unclear etc.)
PK2	Readiness to use methods and instrumental means of study of professional business items
PK21	Know the main architectures of static, dynamic, integrated and hybrid intelligent systems and be able to design and develop them
PK22	Know the methods of designing knowledge bases for different problem/ discipline areas
PK23	Know the composition and structure of the main tools and be able to reasonably select and apply them when implementing different intelligent systems
PK24	Master the basic methods in designing, development, testing and support of specific intellectual system classes

Let's briefly look at certain approaches to elicitation of individual professional competencies. Section 1 has already emphasized the particular features of elicitation of students' knowledge in control efforts (exams, tests, credits) by dynamic shaping of the student's current competence-oriented model that is based on the answers to special web tests and subsequent comparisons with the fragment of course/discipline ontology [2], [6].

Now, we emphasize the problems of elicitation of the skills in application of theoretical knowledge in practical sessions. As the experience suggested, training of knowledge engineers implies learning the skills and capacities in solving practical tasks related to the ability to build the simplest situation models in the problem domain based on the "self-expert" principle, proceeding from products, frames and semantic networks [8], [15].

Thus, when designing the tutoring IES, we created and tested during the training process at MEPhI and other universities the special software that implement the "manual" techniques of addressing non-formalized objectives (non-formalized problems), in particular, those shown in [2], [15]. For instance, to elicit the competencies of PK1, PK11, PK12, PK13, PK14, PK15, PK16, etc. types as listed in Table 1, simulation of the reasoning of the students who solve four types of such non-formalized problems as [6] is used: simulation of the strategies of direct/ reverse reasoning in IES (PK13, PK14, PK15), simulation of simplest situations of the problem domain with the use of frames and semantic networks (PK12, PK13) etc.

Figure. 1 shows the fragment of the applied ontology, Introduction to Intelligent Systems, which comprises: Network Models of Knowledge Presentation, Frames, Frame Definition, Theoretical and Multiple Frame Description, Notion of Frame Prototype and Frame Copy, Homogeneous and Heterogeneous Frame Networks, LKP (language of knowledge presentation) based on frames, Release on Frames Such competencies as PK12 and PK13, which make part of the target competence PK1, are referred to this course/ discipline ontology fragment.

In order to elicit the competencies of PK-2 type, the laboratory practicums and practical knowledge related to the

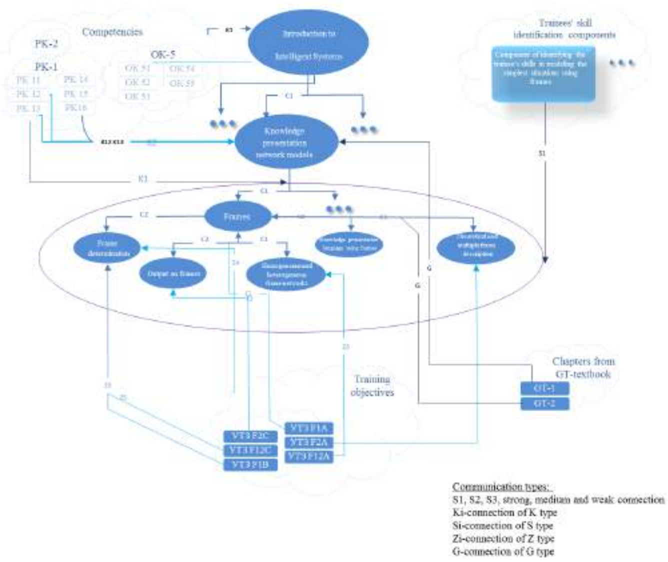


Figure 1. Introduction to Intelligent Systems ontology fragment

instrumental and technological aspects of training knowledge engineers and those oriented on study of the designing technology (KBS) (ES) using the modern arsenal of instrumental tools [2], [6], [15] are used.

Thus, using the Introduction to Intelligent Systems ontology and special means related to building up the student's current models, a certain set of professional requirements (criteria) to competencies of a future knowledge engineer is dynamically created. Table II and Table III contains a set of basic criteria for automated designing of a future specialist's professional competencies (with the recommended expert weights / ratings of each criterion under a 10-score scale) [2]

Now, let's look at the shaping of general cultural competencies and the psychological portrait of a future knowledge engineer

B. Shaping the psychological profile of a future knowledge engineer (in the context of general cultural competencies)

As shown in [15], psychological aspect is critical for knowledge engineering, which is related to knowledge extraction processes, because this particular aspect determines if the knowledge engineer can successfully and efficiently liaise with the knowledge source, the expert. In modern papers, e.g. [2], [7], [8], [15] etc., the suggestion is to take into account a series of personal features or their set as the psychological profile in determination of the so-called "perfect" couple, the knowledge engineer and the expert, to arrange for collective work in creation of the problem area model.

For these purposes, almost two dozens of various author texts are currently used in tutoring IES and web-IES, and the most appropriate test configuration for psychological tests of students depending on the identified competence type is determined using the psychological test generator.

Table II
BASIC REQUIREMENTS TO A FUTURE KNOWLEDGE ENGINEER

s/r #	Requirements	Rating weight
1. Do you have	Broad general scientific training	
	Acquaintance with the reviewing, annotation and other text processing methods	3
	Mastery of quick reading skills	3
	Knowledge of textological methods of deriving knowledge	3
2. Do you know	Acquaintance with the reviewing, annotation and other text processing methods	3
3. Do you master	Mastery of quick reading skills	1
4. Do you know	Knowledge of textological methods of deriving knowledge	9
5. Do you have	Basic training in information sources	
	Qualified knowledge of the methods of knowledge acquisition from different knowledge sources	8-9
	Qualified knowledge of models and methods of knowledge presentation in information sources	8-9
	Qualified knowledge of knowledge processing methods in information sources	8-9
	Qualified knowledge of the systemic analysis basics	8-9
	Mastery of basic information structuring methods	8-9
	Cluster analysis	8-9
	Hierarchical clusterization	8-9
	Designing weighted situations	8-9
	Ranking the selection trees	8-9
	Repertoire grid analysis	8-9
	Possession of KBS (ES) development methods	8-9
	Qualified knowledge of instrumental tools of KBS (ES) designing	8-9
	Holding the practical computer skills, one or several programming languages	8-9
6. Do you master	Qualified knowledge of the methods of knowledge acquisition from different knowledge sources	
	Communications methods	7
	Textological methods	5
	Methods of knowledge acquisition from the database	6
7. Do you master	Qualified knowledge of models and methods of knowledge presentation in information sources	
	Logical models	5
	Inferential	1
	Inductive	1
	Other	1
	Heuristic models	8
	Network	8
	Frames	7
	Semantic networks	5
	Other	3
	Production	8
	Other	3
8. Do you master	Qualified knowledge of knowledge processing methods in information sources	
	Direct or reverse conclusion	8
	Analogy/ precedent based conclusion	5
	Other	3
9. Do you know	Qualified knowledge of the systemic analysis basics	5
10. Are you experienced in	Qualified knowledge of instrumental tools of KBS (ES) designing	9

Table III
BASIC REQUIREMENTS TO A FUTURE KNOWLEDGE ENGINEER

s/r #	Requirements	Rating weight
11. Do you master	Mastery of basic information structuring methods	
	Multiple scaling	5
	Hierarchical clusterization	3
	Designing weighted situations	3
	Conclusion tree ranking	5
	Repertoire grid analysis	5
	Other	3
12. Are you experienced in	Possession of KBS (ES) development methodologies (stages, life cycle, etc.)	7
	Prototyping strategy mastery	8
13. Do you master	Holding the practical computer skills, one or several programming languages	8
14. Do you know	Knowledge of cognitive psychology elements	
	Methods of representing notions and processes in human memory	5
	Main thinking principles (logic, associativity)	5
	Thinking activation methods	5

In terms of the Federal State Education Standard 3+, to shape a knowledge engineer's psychological profile, the following target competencies are the most suitable out of general cultural competencies:

- OK-3, the ability to communicate verbally and in writing in Russian and foreign languages to solve the tasks of inter-personal and inter-cultural communications;
- OK-4, the ability to work as part of a team, by perceiving social, ethnic, confessional and cultural differences with tolerance;
- OK-5, ability to self-manage and self-educate.

However, the use of automated methods for elicidation of the above OK-3, OK-4 and OK-5 competencies is an understudied problem now. There are virtually no papers on establishing the explicit connections between individual personal features of a student's psychological profile and the general cultural competencies. Therefore, as part of taught IELS and web IELS, a cycle of studies intended to specify, in particular, the competencies of OK-4, because the experience related to analysis and clusterization of the students' psychological types has been gained here, and psychological tests are carried out, with the view to the personal degree of achievement of target professional competence of a knowledge engineer.

IV. Conclusion

Thus, the methodical and technological experience gained in training professionals in Applied Mathematics and Informatics and Software Engineering in the domain of automated designing using tutoring IES and web-IES of competence-oriented models of knowledge engineers enable to promptly and efficiently review, adjust (based on the cutting-edge innovations in the professional area) and forecast the level and quality of the graduate professionals' cohort. Such approach does not lay down the new foundations only in relations with employers and potential customers only, but also enables planning of the target training of specialists in various fields among undergraduates.

References

[1] G. V. Rybina, *Theory and technology of construction of integrated expert systems. Monography.* Moscow: Nauchtehlitizdat, 2008.

[2] ——, *Intelligent systems: from A to Z. Monography series in 3 books. Vol. 1. Knowledge-based systems. Integrated expert systems.* Moscow: Nauchtehlitizdat, 2014.

[3] "Portal federal'nyh gosudarstvennyh obrazovatel'nyh standartov vysshego obrazovaniya," http://www.fgosvo.ru.

[4] S. Avdoshin, "Ob opyte ispol'zovaniya professional'nyh standartov pri formirovanii uchebnyh obrazovatel'nyh programm po napravleniyu 230400 «programmnaya inzheneriya»." *Professional'nye standarty v oblasti informacionnyh tekhnologij. M.: APKIT*, pp. 35–44, 2008.

[5] "Razrabotka professional'nyh standartov dlya otrasli informacionnyh tekhnologij. komitet po obrazovaniyu v oblasti it." http://www.apkit.ru/default.asp?artID=5573.

[6] G. Rybina, "Intellektual'naja tehnologija postroenija obuchajushhih integrirovannyh ekspertnyh sistem: novye vozmozhnosti." *Otrytoe obrazovanie.*, vol. 21, no. 4, pp. 43–57, 2017.

[7] L. I. Gavrilova T.A., "Sistemnyj vzglyad na podgotovku inzhenerov po znaniyam i biznes - analitikov," in *Pyatnadcataya nacional'naya konferenciya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2016. Trudy konferencii. V 3h tomah. T.1.*, no. 1, 2016, pp. 16–23.

[8] T. Gavrilova, *Inzheneriya znanij. Modeli i metody*. SPb: Lan, 2016.

[9] O. Fedyaev, "Model' sistemy podgotovki i trudoustrojstva specialistov na osnove programmnyh agentov s nejrosetevoj arhitekturoj," in *Pyatnadcataya nacional'naya konferenciya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2016. Trudy konferencii. V 3h tomah. T.2. Smolensk: Universum*, 2016, pp. 372–381.

[10] M. A. Kirichek A.V., "Aksiomaticheskij bazis kvalimetricheskoj mnogomernoj modeli social'no-professional'noj kompetentnosti molodogo specialista," *Izvestiya Samarskogo nauchnogo centra Rossijskoj akademii nauk*, no. 1, pp. 16–23, 2016.

[11] M. A. P. P. Alekseeva A.V., Mamyrbekov A.K., "Kompetentnostnyj podhod k podgotovke analitikov-tekhnologov," *Izvestiya Samarskogo nauchnogo centra Rossijskoj akademii nauk*, no. 1, pp. 16–23, 2016.

[12] B. A. Koblova N.V., "Realizaciya komptentnostnogo podhoda v processe dopolnitel'noj podgotovki studentov vuza," *Modern Research of Social Problems*, no. 6, pp. 16–23, 2014.

[13] M. Hohlova, "Kompetencii konkurentosposobnogo molodogo specialista: mnenie rabotodatelej i vypusknikov," *Teleskop*, no. 2, pp. 26–31, 2014.

[14] M. I. Bobrov L.K., "Kompetentnostnyj podhod k podgotovke menedzherov biznes-informacii," *FGBOU VO RJeU im. G.V. Plekhanova*, pp. 93–99, 2017.

[15] G. V. Rybina, *Osnovy postroeniya intellektual'nyh sistem.* Moscow: Finansy i statistika, 2014.

НЕКОТОРЫЕ ПОДХОДЫ К АВТОМАТИЗИРОВАННОМУ ПОСТРОЕНИЮ КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫХ МОДЕЛЕЙ ИНЖЕНЕРОВ ПО ЗНАНИЯМ С ИСПОЛЬЗОВАНИЕМ ОБУЧАЮЩИХ ИНТЕГРИРОВАННЫХ ЭКСПЕРТНЫХ СИСТЕМ

Рыбина Г. В., Фонталина Е. С., Сорокин И. А.
Национальный исследовательский ядерный университет МИФИ, Москва, Россия

Развитие инженерии знаний привело к появлению новой профессии, в которой активно востребованы как профессиональные компетенции, так и индивидуальные качества личности. Проанализирован методический и технологический опыт автоматизированного построения компетентностно-ориентированных моделей специалистов в области инженерии знаний, в частности, специалистов по профессии "системный аналитик с использованием обучающих интегрированных экспертных систем.

Ontology of Learning Problems in Boolean Algebra

Nadezhda Gruzdeva
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
Moscow, Russsia
nadezhda.v.gruzdeva@gmail.com

Natalia Bulgakova
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
Moscow, Russia
n.bulgakova.msu@gmail.com

***Abstract*—The article describes an ontology of learning problems in course "Introduction to Boolean Algebra". Nowadays the issue of educational problems classification in computer science is not studied enough. That is why the ontology based on new classification of learning problems. The classification includes Tollinger's taxonomy (cognitive complexity) and structural complexity of the problems (number of input parameters and computational complexity). Depending on the values of these two complexities every learning problem belongs to one of three levels of complexity (easy, medium, difficult). Moreover, all the problems divided into four categories depending on their purpose. The ontology was developed with Protégé 4.3 and OWL DL. It includes over 50 classes and about 60 individual entities. The developed ontology will be useful for teachers to solve the issue of selecting learning problems.**

***Keywords*—ontology, learning problems, complexity, classification, taxonomy, computer science, informatics, Boolean algebra**

I. Introduction

There is an active process of collecting, analyzing and formalizing knowledge for their subsequent representation in the form of ontologies nowadays. Ontology is a hierarchy of concepts of the subject domain, relationships and rules between them. First ontologies were created in 1980s and nowadays a large number either of common ontologies or ontologies of specific areas have already been developed [1, 2, 3, 4, 5]. At the same time, it should be noted that this process is far from completion, because existing ontologies cover only a small part of knowledge areas.

In this article, the area of learning problems for the course "Introduction to the Boolean Algebra" is selected as the specific area for ontology. This choice can be explained by the growing interest in computer science lessons at school. The teacher must carefully select the material for lessons in order to get positive results in education. The ontology is developed for teachers teaching Boolean Algebra to help them to select learning problems more effective. This issue is faced by most teachers because it can be very difficult to find problems for students which would help them to learn the course. A well-chosen set of different tasks can help teachers achieve pedagogical goals such as explaining students how correctly interpret the condition of the problem, break it down into simple components, solve the problem in an effective way, etc.

At the same time, the course "Introduction to the Boolean Algebra" is important part of school informatics. The course "Introduction to the Boolean Algebra" is developed for students of tenth form. It allows students to obtain base knowledge of the Boolean algebra and it causes the development of logical thinking. As practice shows, the uptake of this course by students is difficult due to natural reasons as ambiguous interpretation of some language constructs - "or" in natural language can match either logical "or" or logical "excluding or", the use of three-valued logic in life as opposed to two-valued logic in Boolean algebra. At the same time during problems solving students can work out all difficult moments and learn how they can use the material for practical purposes. Therefore a well-chosen set of problems is an important part of this course.

II. Domain analysis

From pedagogical point of view, all learning problems of this course can be divided into four categories according to their purpose:

- methodically important problems that must be solved on the blackboard by the teacher during the lesson;
- problems for independent work in the classroom, which help to work out solution algorithms during the lesson and don't require checking by the teacher;
- homework problems, which are the most suitable for consolidating the material;
- test problems that allow teacher to check if students get the material well.

Nowadays various classifications of learning problems exist but they are not very suitable for such lessons as informatics. The most suitable for this area Tollinger's taxonomy was taken [6, 7]. It allows to divide the problems into 5 categories according to the increase of their cognitive complexity and operational value.

Moreover, the specificity of the problem area of learning problems for the course "Introduction to Boolean Algebra" makes this classification of problems not suitable for their division, because it does not take into account the computational complexity of the problems. For example, two problems of the same category in Tollinger's taxonomy with different number of input parameters, should be assigned to different complexity categories, since they require different time to solve them.

To solve the issue, a new classification of learning problems that takes into account both the cognitive complexity of the Tollinger's taxonomy and the computational complexity, was developed. New classification of problems was tested in Advanced Educational Scientific Center (faculty) – Kolmogorov's boarding school of Moscow State University and described in the work "Classification of Learning Problems" in course "Introduction to Boolean algebra" Based on Tollinger's taxonomy" [8]. The results of the work were also published in the article in "Informatics in School" [9]. This classification was taken as a basis for creating ontology of learning problems. Development and implementation of the ontology is carried out as N.Bulgakova's master dissertation at the Faculty of Computational Mathematics and Cybernetics of Lomonosov Moscow State University.

During the domain analysis, a set of 60 problems was selected from different sources [10, 11, 12]. This set can be divided into two groups: formalization of the expressions and logical problems. To solve the problems of the first group, students must analyze the text written in natural language and write it in the form of a logical formula, possibly with a further transformation of the resulting formula into a shorter one. To do logical problems students must transform the problem into logical formula and do additional actions with received formula to answer the question posed in the problem.

All problems were classified according to their purpose and complexity according with developed classification. In order to make the ontology easier to extend and supplement with new problems, a formula was chosen. It based on two input parameters - cognitive complexity in Tollinger's taxonomy (K) and computational complexity (C). The cognitive complexity parameter is assigned the category number in Tollinger's taxonomy. The computational complexity depends on number of input parameters that must be taken into account during problem solving. The value of this parameter is expressed by a number, based on the following principles:

- if there are two structural elements in learning problem, then $C = 1$;
- if there are 3 or 4 elements, then $C = 2$;
- if there are 5 or 6 elements, then $C = 3$;
- if there are 7 or 8 elements, then $C = 4$;
- if there are more than 8 elements, then $C = 5$.

The resulting formula helps to automatically calculate the problem class:

- if $2K + C \leq 6$, then easy level of problem complexity;
- if $7 \leq 2K + C \leq 9$, then medium level of problem complexity;
- if $2K + C \geq 10$, then difficult level of problem complexity.

Thus each problem can be associated with an integer lattice point shown in the Fig. 1.

For classifying problems into 4 categories according to their purpose the following principles were used:

- methodically important problems should contain material unknown for students or demonstrate a new algorithm for problems solving;

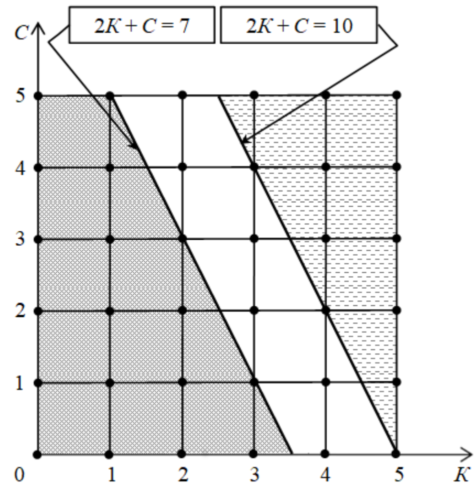


Figure 1. Levels of problem complexity

- problems for independent work in a classroom should have low computational complexity in contrast to problems for homework;
- test problems should show whether the material was learned by the students or not

Example 1. (Formalization problem) You are on duty today, or John, but not both.

The cognitive complexity of this problem is $K = 2$ (Tollinger's taxonomy, category 2.4. Analysis and synthesis), the number of input parameters is $C = 2$, so the resulting complexity is $2*2+2 = 6$ and this problem has easy level of complexity. By purpose this problem is for independent work in the classroom because it has low computational complexity.

Example 2. (Formalization problem) Fog outside or hoarfrost on the trees can be if and only if the thaw is on the street.

The cognitive complexity of this problem is $K = 2$ (Tollinger's taxonomy, category 2.4. Analysis and synthesis), the number of input parameters is $C = 3$, so the resulting complexity is $2 * 2 + 3 = 7$ and this problem has medium level of complexity. By purpose this problem is for homework because it allows students to consolidate new material.

Example 3. (Logical problem) Mrs. Black prepared a wonderful apple strudel for the birthday of little Henry. However, someone stole a treat. Three suspects were soon found, and each made two statements, but it is known that everyone lied at least one time, and exactly one of them is guilty.

A.
1) It was done by either B or I.
2) I agree with the second statement of B.
B.
1) I agree with the first statement of A.
2) I'm not guilty.
C.
1) I agree with the second statement of A.
2) It was done by A, and B helped him.
Who stole the strudel of Mrs. Black?

The cognitive complexity of this problem is $K = 5$ (Tollinger's taxonomy, category 5.2. Problem tasks), the number of input parameters is $C = 3$, so the resulting complexity is $2{*}5{+}3 = 13$ and this problem has difficult level of complexity. By purpose this problem is methodically important because it allows teacher to demonstrate that it is necessary to read the problem carefully and accurately understand its meaning accurately.

Example 4. (Logical problem) Five play cards. In one of the parties in the deck there were five aces: someone was cheating. That's what the players said (everyone once said the truth and once lied):

A.
1) I am innocent.
2) I do not know who was cheating.

B.
1) I am innocent.
2) Neither A nor D was cheating.

C.
1) The first statement of B is true.
2) B is cardsharper.

D.
1) B was cheating.
2) The second time B told the truth.

E.
1) A that day was one of the players.
2) The second statement of D is false.

So who is the cardsharper?

The cognitive complexity of this problem is $K = 5$ (Tollinger's taxonomy, category 5.5. Problems based on rational observations), the number of input parameters is $C = 5$, so the resulting complexity is $2*5+5 = 15$ and this problem has difficult level of complexity. By purpose this problem is for test because the solution of this problem doesn't require analyzing all the statements. In this problem it is necessary to find only one statement that exactly true to reduce computational complexity. Only students who understand solution algorithms of such problems can quickly come to the answer, while for other students it can take a long time.

III. Formalization of domain knowledge

After domain analysis and development of a new classification, the question of a formal description of the developed ontology appeared. The OWL (Web Ontology Language) language was chosen as the formal description language. This language was chosen because there are many ontology editors [13] with the ability to export ontologies written in OWL, for further work with them. OWL has three main modifications: OWL Lite, OWL DL (DL stands for "Description Logic"), OWL Full. Every modification has its own syntax, which defines the semantic power of language. At the same time, the OWL language version, known as OWL Lite, does not include the possibility of specifying a logical relationship between domain concepts. And the language OWL DL makes it possible to describe the logical relationship of the concepts of the domain, which allows to automatically classify learning problems by their purpose and cognitive and computational complexity. So, the OWL DL was chosen for developing the ontology. The main components of the ontology developed with OWL DL are the following:

1) classes - an abstract description of a group of objects. In OWL; there is a Thing subclass, from which other classes are inherited, this class includes all individuals existing in ontology;
2) instances - components of an ontology that are specific objects and are related to the lowest level of ontology;
3) attributes - object properties of classes and data properties.

The Protégé 4.3. ontology editor that allows to visualize the developed ontology and save it as OWL DL code was selected [14] as editor. Moreover, this editor provides the ability to specify instances at the level of classes and subclasses, to track the structure of classes, and has significant reference material.

Currently, the developed ontology includes more than 50 classes, including the Tollinger's taxonomy (33 classes), computational complexity (6 classes) and classification of problems by purpose (4 classes). The scheme of part of ontology is shown in Fig. 2.

In the ontology every learning problem represented as individual entity with description which consists of the problem itself and its solution. Each learning problem is assigned to the cognitive complexity class of Tollinger's taxonomy and to the class of computational complexity. The belonging to the class with level of resulting complexity displayed automatically according to the given logical formula.

In order to add a new problem to the ontology it is necessary to do the following:

1) enter a unique problem identifier;
2) assign the problem with one of the learning topics (tasks for formalization, table method of solving problems, logical reasoning method, etc.);
3) add the problem statement in Russian and in English, add the correct answer;
4) assign one of cognitive complexity classes of Tollinger's taxonomy;
5) assign one of computational complexity classes according to the number of input parameters;
6) assign one of the classes by problem purpose.

Acknowledgment

The described ontology of learning problems in course "Introduction to Boolean Algebra" based on developed classification can really help teachers to find problems for students. Ontology provides users (teachers of informatics) the following opportunities:

- obtain a set of problems with the required characteristics due to the structure of the ontology;
- establish relationship between objects of the ontology, methods of solving learning problems, and complexity of problems;
- add new problems by determining their complexity with developed classification.

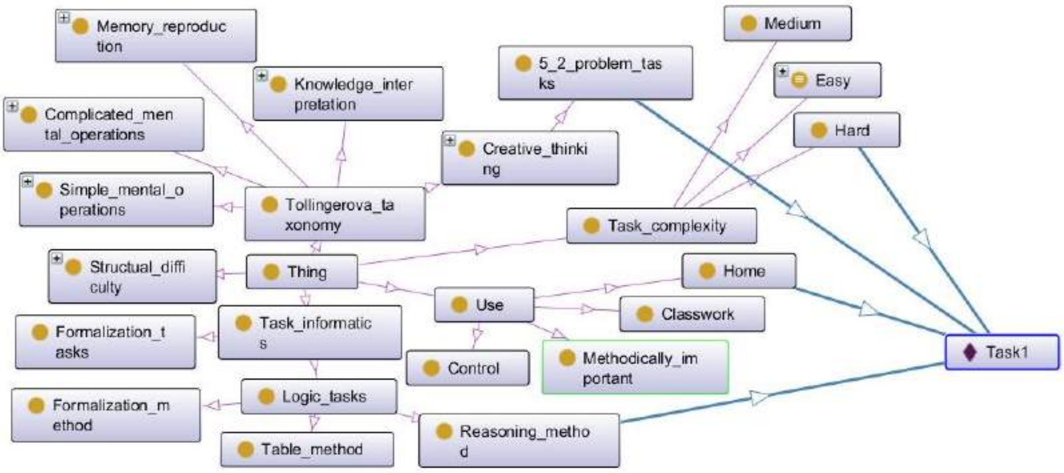


Figure 2. Structure of ontology

Completion of the ontology is planned for March 2018. Testing is preliminary scheduled for April-May 2018 in Advanced Educational Scientific Center (faculty) - Kolmogorov's boarding school of Moscow State University.

References

[1] OWL Reasoning Examples and Hands-On Session. Available at: http://owl.man.ac.uk/2006/07/sssw/.

[2] Ontology examples. Available at: https://github.com/gruninger/colore/tree/master/ontologies

[3] Beta Ontohub, ontology examples. Available at: https://ontohub.org/repositories

[4] The OBO Foundry. Available at: http://www.obofoundry.org/

[5] BioPortal, repository of biomedical ontologies. Available at: http://bioportal.bioontology.org/

[6] D.Tollingerova, D.Goloushova *Psihologiya Proektirovaniya Umstvennogo Razvitiya Detej* [*psychology of designing the mental development of children*], Moscow, Rospedagentstvo, 1994, 48 p.

[7] D.Tollingerova. K Psihologicheskoj Teorii Uchebnyh Zadach [About the psychological theory of learning tasks]. *Socialisticheskaya Shkola* [*Socialist School*], 1976/77, no , pp 156-160.

[8] N.Bulgakova Attestation work *Classification of Learning Problems in course "Introduction to Boolean algebra" Based on Tollinger's taxonomy*, Moscow, Faculty of Educational Studies, Lomonosov Moscow State University, 2017, 55 p.

[9] I.Falina, A.Falin, N.Bulgakova Classification of Learning Tasks in Informatics on Complexity Level. *Informatics in School*, 2017, no 10.

[10] E.Andreeva, L.Bosova, I.Falina *Matematicheskie Osnovy Informatiki. Metodicheskoe Posobie* [*Mathematical Foundations of Computer Science. Methodical Manual*], Moscow, BINOM, Laboratoriya Znanij, 2007, 338 p.

[11] E.Kolman, O.Zih *Zanimatelnaya Logika* [*Entertaining Logic*], Moscow, Nauka, 1966, 128 p.

[12] N.D.Willis. *Mystefing Logic Puzzles*, New York, Sterling Publising Co., Inc., 1998.

[13] D.McGuinness, F.Harmelen *OWL Web Ontology Language Overview*. Available at: http://www.w3.org/TR/owl-features/.

[14] The Protégé Ontology Editor and Knowledge Acquisition System. Available at: http://protege.stanford.edu/.

ОНТОЛОГИЯ УЧЕБНЫХ ЗАДАЧ ПО АЛГЕБРЕ ЛОГИКИ

Груздева Н.В., Булгакова Н.А.
ВМК МГУ

В данной работе описывается онтология учебных задач по курсу "Введение в алгебру логики" предназначенному для школьников старших классов средней школы. Онтология разработана для учителей, ведущих данный курс в средней школе, и предназначена для решения проблемы подбора задач, с которой сталкивается большинство учителей. Для создания онтологии была проведена предварительная классификация задач по сложности. Для классификации использовалась таксономия Толлингеровой, которая позволяет оценить когнитивную сложность задачи, а также структурный анализ задачи, позволяющий определить вычислительную сложность задачи в зависимости от количества входных параметров. Кроме того, для каждой задачи определено её назначение и итоговая сложность, рассчитанная по специальной формуле. Онтология разработана с помощью формального языка OWL и редактора онтологий Protégé 4.3. В онтологии выделено более 50 классов, среди которых 33 отвечают за таксономию Толлингеровой, 6 за вычислительную сложность, 4 за назначение задачи. Задачи представляют собой индивидные сущности, включающие их описание, состоящее из текста условия и текста решения задачи. Для каждой задачи задана принадлежность к классу когнитивной сложности по Толлингеровой и к классу вычислительной сложности. Итоговая сложность задачи выводится автоматически согласно заданным логическим формулам. С учетом итоговой сложности и особенностей решения задачи определяется класс задачи по её назначению.

Knowledge presentation in the content management system eLab-Science

S.N.Sytova, A.P.Dunets, A.N.Kavalenka, A.L.Mazanik, T.P.Sidorovich, S.V.Charapitsa
Institute for Nuclear Problems of Belarusian State University
Minsk, Belarus
sytova@inp.bsu.by

***Abstract*—A short overview of Belarusian content management system (CMS) eLab-Science on the basis of framework eLab is given. eLab is a free software based laboratory information system with elements of electronic document management. This is an electronic system of client-server architecture based on Debian GNU/Linux, the Apache Web server, the Firebird database server, the PHP application server. The system runs under Windows and Linux. The work is carried out through the Web-interface in multi-user mode with the sharing of access rights through widely used browsers.**

CMS eLab-Science was developed to create educational and scientific portals of various profiles such as electronic portal of nuclear knowledge BelNET https://belnet.bsu.by/ and the scientific portal CoExAN https://coexan.bsu.by.The next step is development of Belarusian electronic scientific Archive on the basis of cited CMS. So, knowledge presentation in these portals in the frame of CMS eLab-Science is described.

***Keywords*—content management system, knowledge, taxonomy, electronic scientific archive**

I. INTRODUCTION

Currently, free software occupies a large niche at the world information technology market, providing the user, in contrast to proprietary (licensed, commercial) software, four basic freedoms and rights: unlimited installation, free use, modification and transfer of software [1]. The free software has a number of advantages, facilitating its certification, since source codes and full technical documentation are available for free. The use of such software is the real way to increase the protection of information. This is the reason for the recent trend of IT market towards free software – both in the West and in the post-Soviet space.

This article gives a short overview of Belarusian content management system eLab-Science on the basis of framework eLab. eLab is a free software based laboratory information system with elements of electronic document management. This is an electronic system of client-server architecture based on Debian GNU/Linux, the Apache Web server, the Firebird database server, the PHP application server. The system runs under Windows and Linux. The work is carried out through the Web-interface in multi-user mode with the sharing of access rights through widely used browsers. The eLab main features are the separation of database into system and user databases, he preservation of the current state of user interface, the system operation in real time, with the opening of data pages in less than half a second when working in the internal (corporate) network.

The first modification of the framework eLab was developed for automation of registration and quality control of fuels and lubricants in the Armed Forces of the Republic of Belarus. Since 2012, the "Electronic system for quality control and inventory management of fuel and lubricants eLab-Fuel" stands on alert at 202 Chemistry Center for fuel quality of the Armed Forces of the Republic of Belarus [2]. This center is an accredited testing laboratory in accordance with ISO/MEC 17025. More than 50,000 laboratory tests have been conducted and logged using the system eLab-Fuel over the past time.

Currently, after development of prototype of the software for accounting and control of ionizing radiation sources of eLab-Atom, the work is underway on the Intellectual Information System of the Gosatomnadzor of the Republic of Belarus to control (supervision) on nuclear and radiation safety [3], [4].

Let us point out that the system eLab is built on the basis of the process system approach [5], [6]. It means the identification, understanding and management of interrelated processes in the organization, as well as authorities and responsibilities in business processes managing (sustainable, repetitive activities) that transforms resources at the input to output results. This is necessary to realize the effective functioning of accredited testing laboratories of various types on the basis of key provisions of ISO 9001 and ISO 17025 [7], [8].

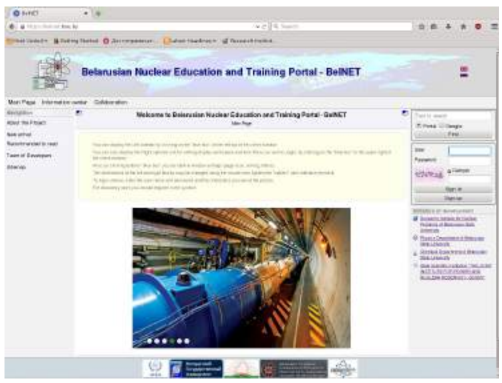


Figure 1. Portal BelNET

On the basis of framework eLab, an original content management system (CMS) eLab-Science was developed to create educational and scientific portals of various profiles.

Figure 2. Portal CoExAN

The electronic portal of nuclear knowledge of educational institutions of the Republic of Belarus BelNET (Belarusian Nuclear Education and Training) was developed and available now https://belnet.bsu.by/ [4]. Also on the basis of CMS eLab-Science the scientific portal of the project of the program Horizon 2020 "Collective Excitations in Advanced Nanostructures" CoExAN was developed https://coexan.bsu.by. The start pages of both portals are given in “Fig. 1” and “Fig. 2”.

Thus, the framework eLab has proved to be flexible and easily customizable for different purposes and projects. The system is in constant development and improvement.

II. Composition of software

Let us describe the composition of software eLab. The system is hosted on a virtual machine (VM) of the VMWare ESX server. It can be hosted on a physical server. The system server is based on the following system components: Apache web server, PHP5 application server, SQL server Firebird 2.5. The users can work both inside the corporate network, and through the Internet, and, if necessary, using VPN (Virtual Private Network). “Fig. 3” shows the organizational structure of the software. Here the system is divided into the Web-server Apache and the Database Store.

The root directory of the system contains libraries (ADOdb, XAJAX) and the following folders:

- **System Core** with the basic files of systems such as libraries of classes and functions;
- **Common files** with ready-made components of the system for use (insertion) in various visual representations (pages) without duplicating the corresponding code;
- **EventLog** with a library for displaying errors and debugging messages;
- **Special** with special modules of the system;
- **Etc** that is a folder in which the connection string to the system database is stored in an encrypted form.

System Core contains modules in PHP, JavaScript and CSS, which provide the overall functionality of the system for all users in a single format: authentication, page design, user controls (buttons, lists, tables), templates, report generators and more. It provides with specialized modules centralized

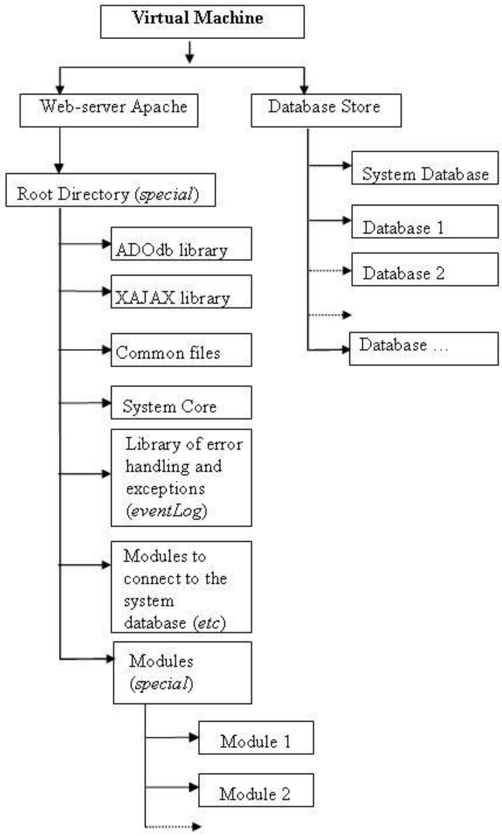


Figure 3. Organizational block diagram of software eLab

validation of input (HTTP requests) and output (HTTP-responses) data, protection from unauthorized code injection, user authentication and access to content in accordance with established permissions for users. The main purpose here is the centralized and fast generation of content pages to display them in different browsers according to stored data, the response to events from different users and corresponding modification of the content and stored data. The content of pages depends on incoming requests and is coordinated with the database of corresponding software product.

Requirements for security in the system eLab, taking into account its implementation in the Armed Forces of the Republic of Belarus and Gosatomnadzor, are increased and implemented as fol-lows. At first, the eLab system provides access to the HTTPS protocol. Secondly, user remote access in the system via VPN is organized. Encryption provides improved eLab security.

III. Content management system eLab-Science

As mentioned above, the original Belarusian content management system of scientific and educational portal eLab-Science developed on the basis of framework eLab was used in portals BelNET https://belnet.bsu.by/ and CoExAN https://coexan.bsu.by. eLab-Science implements all necessary functions of the portal, including the possibility to remote edit the portal structure and document and resource entering, various sorting and filtering, as well as several levels of access to documents depending on user rights, the original mechanism of testing during students laboratory works. eLab-Science provides the following editors for formation and organization of access to resources:

- editor of portal sections,
- resource type editor,
- resource editor,
- systematization of resources,
- access to files,
- portal structure editor,
- content resource editor, including pictures and LaTeX-like formulas input,
- editor of control questions of laboratory work tests,
- editor of answers to the test control questions.

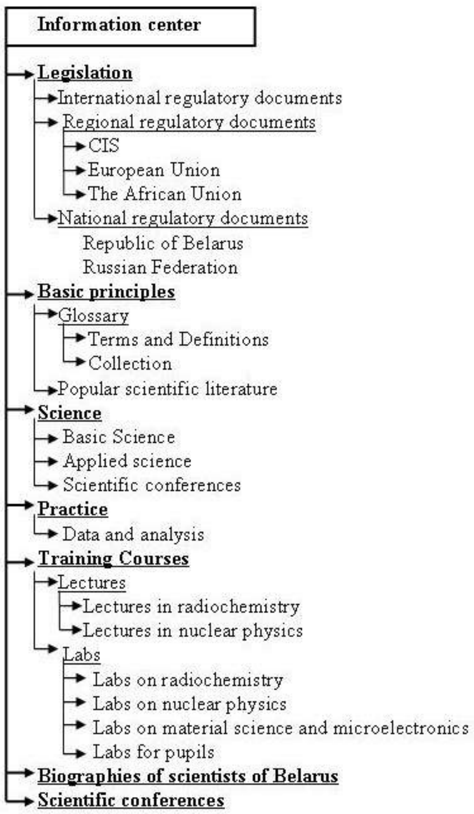


Figure 4. Taxonomy of BelNET

Let us display well-known CMS providing such possibilities. The system Moodle https://moodle.org/ is a learning platform on the basis of free software designed to provide educators, administrators and learners with an integrated system to create online learning sites and personalized courses. The Cyber Learning Platform for Network Education and Training (CLP4NET) http://clp4net-nkm.iaea.org/ was developed on the basis of Moodle as an online platform of educational resources under the aegis of IAEA. MediaWiki https://www.mediawiki.org/wiki/MediaWiki is a free software open source wiki package written in PHP, originally for use on Wikipedia. It is a content management system for different wiki projects and now it is also used by many other projects. Other widely distributed CMS allow to insert formulas as pictures made by some external editors. So, along with the generally recognized CMS like Moodle and MediaWiki, eLab-Science provide the possibility of developing complex scientific texts.

Let us consider one of the fundamental conception in any scientific and educational portal development. This is the taxonomy (from Greek taxis meaning arrangement or division and nomos meaning law) [9]. It is a hierarchical structure with information or knowledge categorized, allowing an understanding of how these various parts relate to each other. Taxonomies are used to organize information in systems, thereby helping users to find it [10]. The main part of the taxonomy of portal BelNET in presented in "Fig. 4" Taxonomy of CoExAN is in the "Fig. 5" Their difference is explained by the different goals and objectives of the portals.

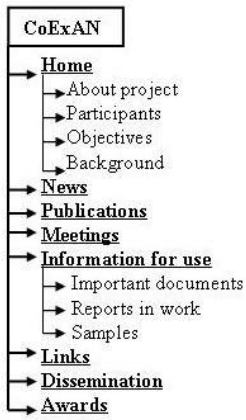


Figure 5. Taxonomy of CoExAN

IV. Online submission, distribution, and archival scientific services in the world

Let us examine the possibilities of online submission, distribution, and archival services in the world in the field of science and technology. arXiv http://xxx.lanl.gov/ is a well-known e-print service in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science funded by Cornell University Library, the Simons Foundation and Los Alamos National Laboratory. ChemRxiv https://chemrxiv.org/ is a free online submission, distribution, and archival service for unpublished preprints in chemistry and related areas developed by the American Chemical Society. bioRxiv https://www.biorxiv.org is a free online archive for unpublished preprints in the life sciences. It is operated by Cold Spring Harbor Laboratory. In these services preprints are in English and available free. The Russian information system "Scientific Archive" https://научныйархив.рф/, as stated there, contains more than 2 million documents (articles, dissertations, author's abstracts) in all areas of scientific knowledge. In this system language is Russian and there is no free access to documents. Search service here raises many questions. There is practically no another archival online scientific services in the world, especially in Russian-speaking countries.

Work on Belarusian electronic scientific Archive on the basis of CMS eLab-Science is at the beginning. eLab-Science has all necessary functions for such task. The Archive features are the following. Archival online service is oriented for publications scientific articles, including preprints, prepublications of the natural and humanitarian profile. Languages of publications are English, Russian, Belarusian with a mandatory summary in English. Service allows to create full-fledged Internet pages on scientific topics with formulas, graphics, drawings, video.

The purpose of the Archive is to promote the dissemination of scientific knowledge, to increase the effectiveness of scientific research, to improve the quality of scientific communication and the transparency of the scientific environment. Draft taxonomy of Archive is given in“Fig. 6”. It is clear that this taxonomy should be expanded.

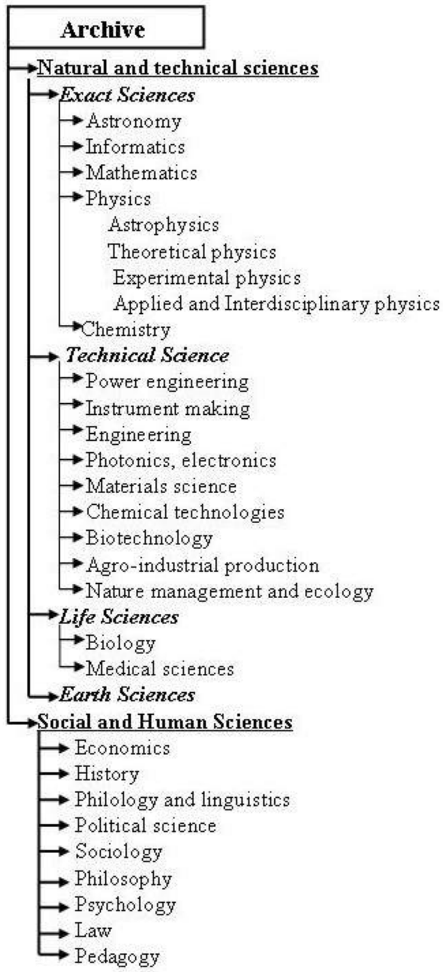


Figure 6. Taxonomy of Belarusian electronic scientific Archive

V. Conclusion

So, Belarusian CMS eLab-Science portal gives possibilities to modern style of knowledge presentation. Development of Belarusian electronic scientific Archive on the basis of eLab-Science will be the important step in Belarusian science.

References

[1] What is free software? GNU Operating System. Available at: https://www.gnu.org/philosophy/free-sw.html (accessed 15 May 2017).

[2] Sytova S.N., Dunets A.P., Kovalenko A.N. et al. Belarusian software for automation of accounting and control systems in the Armed Forces of the Republic of Belarus. 7th International Scientific Conference on Military-Technical Problems, Problems of Defense and Security, Use of Technologies of Dual Use: a collection of scientific articles. Minsk, 2017, pp. 231-222.

[3] Sytova S.N., Dunets A.P., Kovalenko A.N. et al. Information system eLab for accredited testing laboratories. Informatics. 2017. N55.

[4] Sytova S., Dezhurko M., Dubovskaya I. et al. Content of the educational and scientific portal of nuclear knowledge BelNET. The Higher School, 2016, N.5, pp.22-27.

[5] Marca D.A.,McGowan C.L. Sadt: Structured Analysis and Design Techniques. McGraw-Hill, 1987, 392 pp.

[6] Young S. Management: a systems analysis. Scott, Foresman and Company, 1966, 436 pp.

[7] ISO 9001:2015 Quality management systems – Requirements. ISO/CASCO Committee on conformity assessment, 2015, 29 pp.

[8] ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. ISO/CASCO Committee on conformity assessment, 2017, 30 pp.

[9] Development of Knowledge Portals for Nuclear Power Plants. IAEA Nuclear Energy Series No. NG-T-6.2. Vienna: International Atomic Energy Agency, 2016. 56 p.

[10] Fast reactor knowledge preservation system: taxonomy and basic requirements. Vienna: International Atomic Energy Agency, 2008. 79 p.

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ В СИСТЕМЕ УПРАВЛЕНИЯ КОНТЕНТОМ eLab-Science

С.Н. Сытова, А.П. Дунец, А.Н. Коваленко, А.Л. Мазаник, Т.П. Сидорович, С.В.Черепица
Институт ядерных проблем БГУ

Дается краткий обзор белорусской системы управления контентом eLab-Science на основе фреймворка eLab. eLab – это лабораторная информационная система с элементами электронного документооборота на базе свободного программного обеспечения. Это электронная система клиент-серверной архитектуры, основанная на Debian GNU/Linux, веб-сервер Apache, сервер баз данных Firebird, сервер приложений PHP. Система работает под Windows и Linux. Работа выполняется через веб-интерфейс в многопользовательском режиме с разделением прав доступа через широко используемые браузеры. eLab-Science была разработана для создания учебных и научных порталов различных профилей, таких как электронный портал ядерных знаний BelNET https://belnet.bsu.by/ и научный портал CoExAN https://coexan.bsu.by. Следующим шагом является создание на ее основе белорусского электронного научного архива. В работе описывается представление знаний на этих порталах в рамках системы eLab-Science.

Integration of Intelligent and Blockchain Technologies in Information Management Systems

U.A. Vishniakou
Educational Establishment «Belarusian State University of Informatics and Radioelectronics»
Minsk, Belarus
vish2002@list.ru

***Abstract*—Tendencions the use of intelligent technologies in information management (IM) are done. Main directions of semantic technologies are discused. The use of block chain technologies for various actives management is shown. The use of intelligent technologies in information management with cloud computing (CC) and intelligent agents (IA) are represented. The analysis of management activity particulars in area of CC directions is done. As IM trend development of methods and models creation of jointly activity IAs in cloud area is given. As conception development of information management activity the crtiating instrumental platform on the multi agents base, integration semantic and blockchain technologies in cloud area is proposed. The automation design of such intelligent IM system (IIMS) with proposed concepts can be realize on the base of semantic technology.**

***Keywords*—semantic, block chain technologies, information management, intelligent agents, integration, cloud area, cloud computing**

I. Introduction

The information management (IM) theory and practice development is following: with one hand – the strength attention to investigation for information technologies, increase in requirements to the international standards in information management field, the growing expenditures on their support, with another – the non sufficiently IM effect, subjective human factor in control of resources, what the published data to world economy confirm [1].

The first way to improve this situation are the collection and processing information about state and trend for goods and services in would economic and synthesis knowledge what will give new more exactly forecast situation on would markets. The other way is the use in information management intelligent technologies (decision making) and block chain technologies (control and resources accountant) for creating information corporate systems (ICS) of new generation [2] and the clarity of all resources management [3].

II. Intellectualization in information management

One of main task of innovation economic is intellectualization. It idea is the development of effective mechanism of forming, actualization and mass using innovation knowledge in information management. Among such knowledge can propose the following: intelligent agents, intellectual web-services, cloud computing [4]. These proposes can be taken in base during the intellectualization of management activity:

- the use of semantic technologies in IM, electronic marketing, electronic and mobile commerce giving new quality;
- the create of integration decisions on the base of semantic web-services let to decide new management tasks;
- the use of mobile technologies and cloud computing create new platform for IM.

III. The base of semantic technologies

President of W3C Tim Bernes-Lee has proposed the Semantic Web conception in 2001 year on the 8-th consortium W3C conference and has published the article in Scientific American magazine [5]. The distinguish between Semantic Web (SW) and classical Web environment is that – the each page of semantic net include information in two languages: human (shown by browser) and special (conceal from human but it is clear for special programs – intelligent agents (IA). IA finds for the owner task the need information, asked data, checks it on some search criteria and gives the results in convenient for user forms.

Three principles: aggregation, safety and logic are in the base of semantic net. Aggregation is the common using of data. Any type of data can be used in SW, on its base is creating the semantic information (ontology). The last is the fundament of SW and adduce the project area description on some formal language of intelligibility and relations between them.

The security base of SW ensures the digital signature technology, which can be used by IA for checking, that some information received from authentic source.

Logic includes the rule set for description of information data structure, protocols and page description. The RDF language has been developed by W3C for the meta data description in semantic net. It is used for relation description between resources. Propositions coding by RDF can be interpreted by the ontology creating on RDF-Schema and OWL standards for receiving the logical conclusions [5].

The developing of semantic technologies is continuing in Belarus in project OSTIS, the main idea which to create and improve the component design of any intelligence systems. The developing of semantic technologies is continuing in

Belarus in project OSTIS, the main idea it is to create and improve the component design of any intelligence systems. In article [6] the such new open project is devoted, aimed at creation of technology of component design of intellectual systems.

Among the key provisions underlying proposed technology of intelligent systems design, the following provisions apply [6]:

- component design method is based on permanently expandable libraries of reusable components;
- formal models of designed intelligent systems are based on the unified semantic networks, which creates the necessary conditions for the semantic interoperability of intelligent systems and their components;
- for decreasing the complexity of the design and modification intelligent systems, the maximum possible independence of the knowledge base update process from the knowledge processing methods and the technical implementation is used;
- the technology of component design of intelligent systems is realized as intelligent meta system, which is built on the same proposed technology and stores all accumulated to this time models, means (including the library of standard components) and methods that are the part of this technology;
- the permanent development of component design technology is performed within the scope of open source-project.

In the proposed design technology for intelligent systems special attention is paid to the upgrade during their operation process and to meta technology of updating of component design of intelligent systems.

IV. The block chain technologies base

Bloch chain – it is multifunction and multilevel information technology, which destination is the reliable discount of various actives. This technology covers all area of economic activity and has multitude of using. These are: finance, economy; operations with material and nonmaterial actives, counting in the state, private and hybrid organizations. In common, block chain is new paradigm for coordi-nation of any type of management activity [3].

Bloch chain technology (BCT) can be been the economic shell of Internet for online paymentearnings, non centralize exchanges, expenditure of valuable tokens, receive and send of digital actives, issue and execute of smart contracts. Block chain as decentralization technology can bebeen the next stage after mobile and social nets [3].

As technical the block chain technology is the elseone application level over stack of internet protocols. It gives in the Internet the new support element of economic transactions the moment payment in universal crypto currency, or so more complex financial contracts. The block chain technology will state as high economic level of various computers connection set: mobile, digital unit of self fixation, “smart homes, smart autos and smart region” [3].

The functionality implemented within the framework of the paradigm block chain can look like an integrated physical level of calculations with many devices, on top of which there is a layer for servicing payments. But it’s not just about payments, but about micro payments, a decentralized exchange, earning and spending tokens, getting and transferring digital assets, and drawing up and executing clever contracts - that is a full-fledged economic layer that has not yet been available in the Internet [3].

V. New in information management

Intelligent network (web 3.0) is becoming the stage of development of the Internet. The ontology forms the semantics, creating new opportunities for IA to fulfill user requests. This ensure the liberation of the user from the task of documents examination receiving search engine. To cope with the complexity of business processes linking multiple enterprises or value chain in Web 3.0 companies will demand smart processes [2, 7].

Distributed artificial intelligence – DAI is based on agent-based technologies. IA has three properties: autonomy, the ability to respond and to get in touch. IA can chat with other "entities": people, other IA, objects [6]. Adding to that the ability to plan and set goals, to support view models, to reason about actions and to increase the level of knowledge and quality of work through training, and get advanced IA.

IA can be integrated into the structure of cloud computing (CC) that contains specific functions in solution of tasks, data processing and management. IA supports the connection of information and technology, knowledge base and can support the process of logical reasoning (for example, including business regulations). IA allows to enable learning and improvement both at the level of infrastructure (adaptive routing) and application-level (adaptive user interfaces). IAs are used to gather business intelligence (BI) and complex event processing (CEP). The number of visits of pages is out of date. It is important the number of connections in social networks, the number of sent messages and time spent on a particular site [7].

Information receiving and real-time analysis in cloud area is the next task for corporate intelligence, especially when, in order to find the valuable information and to "manage the reputation", it is necessary to move from "search data" to "search blogs". It is necessary to go beyond the Google search engine, to handle the online noise and to understand what is happening in the industry, the situation of the goods and services of the company, i.e., it is need the analytics in Web 3.0.

Using complex event processing for corporate intelligence, you can create feedback between them and management system business processes, which in turn, affects corporate intelligence.

Service Science Management and Engineering (SSME) – a term used by IBM Research in its development of service systems. HP has created a "Research center systems and services". Oracle Corp. joined IBM to create an industrial consortium called the Service Research and Innovation Initiative. A group of NESSI (Networked European Software and Services Initiative) in the European Union created working group on science services. At the University of California Berkeley has a program SSME. This is due to the fact that the services sector now employs more than 50% of the workforce in Brazil, Russia, Japan and Germany, and 75% of the workforce in the US and the UK [8].

Time of monolithic, vertically integrated companies left. Main and auxiliary business processes (BP) of the company occurs in four interrelated areas: suppliers direct procurement, production resources (indirect purchasing), trading partners, customers. These multiple chains should be included in the new business ecosystem in cloud area, combining "all-with-all". They will be available for the connection, gap and the new connection giving the company opportunity to work in multiple markets or to create a new proposal for a "market of one".

Successful companies become representatives of the interests of their clients. They work with suppliers from all over the world in order to offer customers the best value. The answers to the questions who owns the business processes of the entire value chain lie outside of CRM-systems – for new systems of cloud area (CA): relationship management in the value chain and relationship management with the community of customers. Ties are portals, business ecosystems and formation system of information, processes, for example, "individual request for a product" coming through many channels and from many communities of customers [2, 7].

Management CRM 2 is located in CA. The same can be said about the life-cycle management of goods in the new world of innovations generated by consumers: computer aided design, management of supply chains and contractual relationships will have to go beyond the boundaries of the enterprise to include customers and partners in the design and manufacture. The since is not company "owns", the entire sequence of value creation, the business processes of companies and their management will be unified and moved to CA.

This architecture includes the knowledge base in the form of production rules, logical inference mechanism, receptors and effectors of the agent module for communicating with other agents. Applied to the problem of market analysis, receptors convey facts about the outside influences in the knowledge base. The logical result of the output produced solution which is passed to the effecter about changes in the external environment.

Distributed decision of IM system can be used by different types of agents: agent-subordinator, set of agents, an agent-integrator. Agents can be linked together in the complex architecture, which can be horizontal or vertical. In result of analysis of the information processes in distributed IM system it can be considered agents that delineate access rights of users, agents, detection of needs (that is, state changes of the market environment), agents of discovery innovation, agents that build the behavior for the spread of innovation, the agents are the coordinators of whole multi-agent system.

VI. Requiem's to new IM system

In the article [9] are analyzed the main developments in the IM intelligent system (IMIS) and the main trends of their development. With result of BCT integration the list of criteria and their values which can meet the IMIS:

- multi level monitoring environment, collect data on market condition from various sources at different levels of observation – the level of network, server, CA and social systems;
- adaptability, the ability to detect modified implementation of known and new innovations of the market;
- proactively, a built-in reaction mechanisms on the emergence of innovations;
- openness, possibility of adding new analyze resource for information system, its control;
- management, IS needs to combine both centralized and distributed control and discount;
- security, IS must have a means of protecting their components.

As a result, presented the following solutions for multi agent IS of innovation market:

- the structure of multi-agent detection system of innovation including agents which allows to make conclusion about the state and prospects of development of the market;
- method of adoption by agents of the joint decision, in order to form the agent round table and on the basis of the analysis of information obtained from different sources, to appreciate the state of the market;
- detection technique the news using multi agent technology to teach a multi agent system for the detection of innovation and use it to further planning the new goods (services).

VII. The concept of system IM development

As trends and concepts of development on the use of intellectual and block chain technologies in IM systems are invited to:

- the improving architecture of IM systems in cloud environments, to ensure effective management under conditions of uncertainty state of information environment;
- the development of new models of IM in CA with the IA use on the basis of choice of optimal variant of response to market events;
- the improvement of instrumental and software systems for IM with intelligent decision support and research on the effectiveness of methods, models and algorithms;
- the development of theoretical foundations, models and tools, cloud tool platform design intelligent systems IM based on semantic technologies;
- the development of application software of workstations or sites for managers and marketers using block chain technology.

As an example, the structure of a distributed IM system is described. It includes an intelligent portal, with has the support of intelligent agents and block technology. The portal includes a knowledge base (KB), an intellectual solver, an explanatory subsystem, an interface with agents, and a KB editor. Intelligent agents scan the sites of the industry and the sites of scientific institutions and form the knowledge base of the portal. Solver finds coincidences and issues order variants. As a result of human negotiations, smart contracts are being formed, which are supported by block chain technology.

The automation design of such intelligent IM system (IIMS) with proposed concepts can be realize on the base of semantic technology [10]. The automation design of ontology's for IIMS is based on worked out theory, approaches, models and tools [11, 12].

VIII. Conclusion

First direction of the development of IM intelligent systems is the further development of models, methods, architectures and software tools for the solution of problems of adaptation in the markets. The second direction is – the development of models, methods, architectures and software tools for gathering, structuring information from the Internet, the formation of specialized knowledge bases and decision support. The third direction is – the creation of a cloud-based tool platform design of IM intelligent systems based on semantic and block chain technology.

References

[1] Intelligent technologies in management [Electronic resources]. Mode of accesses : http://www. Eurasiancommission. org/ru/act/promi_agroprom/dep_prom/SiteAssets/2017.pdf. Data of of accesses : 9.12.2017.

[2] Vishniakou, U. A. *Information management and security: methods, models, software and hardware solutions. Monograph.* / U. A. Vishniakou. Minsk. MIU, 2014. 287 pp.

[3] Swan M. *Block chain. Scheme new economy* / M. Swan. "Olymp-Business", 2015. 142 p.

[4] Vishniakou, U. A. *Information management. manual with the stamp of EM* /U. A. Vishniakou. Minsk: Bestprint. 2015. 305 pp.

[5] Berners-Lee T. *The Semantic Web, T. Berners-Lee, J. Hendler, O. Lassila* // Scientific American, May 2001. P. 28-37.

[6] Golenkov, V.V., Guliakina N.A. Open project aimed at creating of technology of component design of intelligent systems. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], Minsk, BSUIR. P. 55-78.

[7] Fingar, P. *Cloud computing – the business platform of the XXI century* / P. Fingar. M.: Aquamarine Book, 2011. 256 pp.

[8] Ritz, John. *Cloud computing* / John. Ritz. S. Petersburg: BHV, 2011. 288 p.

[9] Vishniakou, U. A. *Development of intelligent control by using cloud technologies* / U. A. Vishniakou // Informatics, 2016, No. 2. P. 113-120.

[10] Golenkov, V.V., Guliakina N.A. Semantic technologies of component design systems manage by knowledge. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]. Minsk, BSUIR, 2015. P. 57-78.

[11] Golenkov, V.V. Ontology-based design of intelligent systems. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]. Minsk, BSUIR, 2017. P. 37-56.

[12] Davydenko, I.T. Ontology-based knowledge base design. Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]. Minsk, BSUIR, 2017. P. 57-72.

ИНТЕГРАЦИЯ ИНТЕЛЛЕКТУАЛЬНЫХ И БЛОКЧЕЙН ТЕХНОЛОГИЙ В ИНФОРМАЦИОННОМ УПРАВЛЕНИИ

Вишняков В.А.

Приведены тенденции использования интеллектуальных технологий в информационном управлении (ИУ). Обсуждены основные идеи семантических технологий. Рассмотрено использование технологии блокчейн для управления различными активами. Представлено использование интеллектуальных технологий в информационном управлении с использованием облачных вычислений (ОВ) и интеллектуальных агентов (ИА). Как тенденция развития ИУ рассмотрено совершенствование методов и моделей совместной деятельности ИА в облачной среде. В качестве развития информационного управления предложено создание инструментальной платформы на базе многоагентного подхода, интеграции семантических и блокчейн технологий в облачной среде. Автоматизация проектирования этой платформы будет использовать семантическую технология проекта ОСТИС.

Ключевые слова: семантические, блокчейн технологии, информационное управление, интеллектуальные агенты, интеграция, облачая среда

АВТОРСКИЙ УКАЗАТЕЛЬ

AUTHOR INDEX

Научное издание

Открытые семантические технологии проектирования интеллектуальных систем

Open Semantic Technologies for Intelligent Systems

МАТЕРИАЛЫ
МЕЖДУНАРОДНОЙ
НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ

(Республика Беларусь, Минск, 15–17 февраля 2018)

Основан в 2017 году

Выпуск 2

В авторской редакции
Ответственный за выпуск *В. В. Голенков*
Компьютерная вёрстка *Н. В. Гракова, А. В. Губаревич*

Подписано в печать 30.01.2018. Формат 60×84 1/8. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 46,5. Уч.-изд. л. 66,4. Тираж 140 экз. Заказ 19.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий №1/238 от 24.03.2014,
№2/113 от 07.04.2014, №3/615 от 07.04.2014.
ЛП №02330/264 от 14.04.2014.
220013, Минск, П. Бровки, 6

OSTIS-
2019

IX международная научно-техническая конференция «Открытые семантические технологии проектирования интеллектуальных систем»

Open Semantic Technologies for Intelligent Systems

21 – 23 февраля 2019 г. Минск. Республика Беларусь

И Н Ф О Р М А Ц И О Н Н О Е   П И С Ь М О

Приглашаем принять участие в IX Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2019), которая будет посвящена вопросам области применения семантических технологий.

Конференция пройдет в период с ***21 по 23 февраля 2019*** года в Белорусском государственном университете информатики и радиоэлектроники, г. Минск, Республика Беларусь.

Язык статей сборника конференции: английский

Рабочие языки конференции: русский, белорусский, английский.

ОСНОВНЫЕ ОРГАНИЗАТОРЫ КОНФЕРЕНЦИИ

- Российская ассоциация искусственного интеллекта (РАИИ)
- Белорусский государственный университет информатики и радиоэлектроники (БГУИР)
- Государственное учреждение «Администрация Парка высоких технологий» (Республика Беларусь)
- Министерство образования Республики Беларусь
- Министерство связи и информатизации Республики Беларусь

НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ:

- *Принципы, лежащие в основе семантического представления знаний, и их унификация.*
  Типология знаний и особенности семантического представления различного вида знаний и метазнаний.
  Связи между знаниями и отношения, заданные на множестве знаний.
  Семантическая структура глобальной базы знаний, интегрирующей различные накапливаемые знания
- *Языки программирования, ориентированные на параллельную обработку семантического представления баз знаний*
- *Модели решения задач, в основе которых лежит обработка знаний, осуществляемая непосредственно на уровне семантического представления обрабатываемых знаний. Семантические модели информационного поиска, интеграции знаний, анализа корректности и качества баз знаний, сборки информационного мусора, оптимизации баз знаний, дедуктивного и индуктивного вывода в базах знаний, правдоподобных рассуждений, распознавания образов, интеллектуального управления. Интеграция различных моделей решения задач*
- *Семантические модели восприятия информации о внешней среде и отображения этой информации в базу знаний*
- *Семантические модели мультимодальных пользовательских интерфейсов интеллектуальных систем, в основе которых лежит семантическое представление используемых ими знаний, и унификация этих моделей*
- *Семантические модели естественно-языковых пользовательских интерфейсов интеллектуальных систем. Структура семантического представления лингвистических баз знаний, описывающих естественные языки и обеспечивающих решение задач понимания естественно-языковых текстов и речевых сообщений, а также задач синтеза естественно-языковых текстов и речевых сообщений, семантически эквивалентных заданным фрагментам баз знаний*
- *Интегрированные комплексные логико-семантические модели интеллектуальных систем, основанные на семантическом представлении знаний, и их унификация*
- *Различные технические платформы и варианты реализации интерпретаторов унифицированных логико-семантических моделей интеллектуальных систем, основанных на семантическом представлении знаний*
- *Средства и методы, основанные на семантическом представлении знаний и ориентированные на проектирование различных типовых компонентов интеллектуальных систем (баз знаний, программ,*

решателей задач, интерфейсов)

- *Средства и методы, основанные на семантическом представлении знаний и ориентированные на комплексное проектирование различных классов интеллектуальных систем (интеллектуальных справочных систем, интеллектуальных обучающих систем, интеллектуальных систем управления, интеллектуальных робототехнических систем, интеллектуальных систем поддержки проектирования и др.)*
- *Прикладные интеллектуальные системы, основанные на семантическом представлении используемых ими знаний*

ЦЕЛЬ И ФОРМАТ ПРОВЕДЕНИЯ КОНФЕРЕНЦИИ

Целью конференции является обсуждение проблем создания **открытой комплексной семантической технологии компонентного проектирования интеллектуальных систем**. Этим определяется и формат её проведения, предполагающий (1) пленарные доклады, (2) секционные заседания; (3) круглые столы, посвященные обсуждению различных вопросов создания указанной технологии; (4) выставочные презентации докладов.

Выставочная презентация докладов даёт возможность каждому докладчику продемонстрировать результаты своей разработки на выставке. Формат проведения конференции предполагает точное время начала каждого доклада и точное время его выставочной презентации.

Важнейшей задачей конференции является привлечение к её работе не только учёных и аспирантов, но и студенческой молодежи, интересующейся проблемами искусственного интеллекта, а также коммерческих организаций, готовых сотрудничать с научными коллективами, работающими над интеллектуальными системами и созданием современных технологий и их проектированием.

УСЛОВИЯ УЧАСТИЯ В КОНФЕРЕНЦИИ

В конференции имеют право участвовать все те, кто интересуется проблемами искусственного интеллекта, а также коммерческие организации, готовые сотрудничать с научными коллективами, работающими над созданием современных технологий проектирования интеллектуальных систем.

Для участия в конференции OSTIS-2019 необходимо до 15 декабря 2018 года зарегистрироваться в системе СМТ, найти страницу конференции и на ней:

- подать **заявку** на конференцию OSTIS-2019. Каждое поле заявки обязательно для заполнения, в том числе указание того автора, кто будет представлять доклад. Заполняя регистрационную форму, Вы подтверждаете согласие на обработку Оргкомитетом конференции персональных данных, публикацию статей и информации об авторах в печатном и электронном виде. В заявке должна содержаться информация по каждому автору. К заявке доклада должны быть прикреплены **цветные фотографии** всех авторов статьи (это необходимо для публикации Программы конференции);
- загрузить **статью** для публикации в Сборнике материалов конференции OSTIS-2019. Статья на конференцию должна быть оформлена в соответствии с правилами оформления публикуемых материалов и занимать не менее 4 полностью заполненных страниц;
- загрузить **сканированный вариант письма о согласии** на публикацию и размещения передаваемых материалов в сети Интернет.

Если доклад представляется на конкурс докладов молодых учёных или на конкурс программных продуктов молодых учёных, это должно быть явно указано в заявке доклада.

Отбор статей для публикации в Сборнике и участия в работе конференции осуществляется рецензентами и редакционной коллегией сборника.

Заявки и статьи, оформленные без соблюдения предъявляемых требований, не рассматриваются.

До 30 января 2019 года, авторам статей, включённых в Программу конференции, направляются приглашения для участия в конференции.

Участие в конференции не предполагает организационного взноса.

ПОРЯДОК ПРЕДСТАВЛЕНИЯ НАУЧНЫХ СТАТЕЙ

Статьи (только по перечисленным выше направлениям) представляются в готовом для публикации виде (http://proc.ostis.net -> Авторам). Текст статьи должен быть логически законченным и содержать новые научные и практические результаты. От одного автора допускается не более двух статей.

После получения статьи, она отправляется на рецензирование и в срок до 25 января на сайте CMT вы сможете ознакомиться с результатами рецензирования

Оргкомитет оставляет за собой право отказать в приеме статьи в случае, если статья не будет соответствовать требованиям оформления и тематике конференции, а также, если будет отсутствовать заявка доклада, соответствующая этой статье.

КОНКУРС ДОКЛАДОВ МОЛОДЫХ УЧЁНЫХ

Среди авторов доклада, представляемого на конкурс докладов молодых учёных, могут входить учёные со степенями и званиями, но непосредственно представлять доклад должны авторы, не имеющие степеней и званий в возрасте до 35 лет.

Для того, чтобы принять участие в конкурсе научных докладов молодых учёных необходимо:

1) заполнить заявку на участие в конференции, в которой чётко указать своё желание принять участие в данном конкурсе;
2) написать статью на конференцию и загрузить на сайте CMT;
3) заполнить, подписать, отсканировать и отправить по почте письмо о согласии;
4) лично представить доклад на конференции.

КОНКУРС ПРОЕКТОВ МОЛОДЫХ УЧЁНЫХ

Принимать участие в конкурсе проектов молодых учёных могут проекты прикладных интеллектуальных систем и систем ориентированных на поддержку проектирования интеллектуальных систем, при этом представлять проект на конкурсе должен молодой учёный в возрасте до 30 лет, не имеющие учёных степеней.

Для того, чтобы принять участие в конкурсе программных продуктов молодых учёных необходимо:

1) заполнить заявку на участие в конференции), в которой чётко указать своё желание принять участие в данном конкурсе;
2) написать статью на конференцию и загрузить на сайте CMT;
3) лично представить доклад на конференции;
4) провести выставочную презентацию, разработанного программного продукта.

КОНКУРС СТУДЕНЧЕСКИХ ПРОЕКТОВ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

В конкурсе студенческих проектов могут принимать участие проекты, разработчиками которых являются студенты и магистранты высших учебных заведений, консультантами и руководителями проекта могут быть лица, имеющие научную степень и звание. Для того, чтобы принять участие в данном конкурсе необходимо:

1) ознакомиться с положением о конкурсе студенческих проектов (http://www.conf.ostis.net);
2) заполнить заявку на участие в конкурсе студенческих проектов (http://www.conf.ostis.net);

3) подготовить описание проекта (http://www.conf.ostis.net).
4) выслать заявку на участие в конкурсе и описание проекта по электронному адресу конкурса студенческих проектов: ostis.stud@gmail.com.

ПУБЛИКАЦИЯ МАТЕРИАЛОВ КОНФЕРЕНЦИИ

Оргкомитет конференции предполагает публикацию статей, отобранных Программным комитетом по результатам их рецензирования, в Сборнике материалов конференции и на официальном сайте конференции http://conf.ostis.net и официальном сайте сборника http://proc.ostis.net.

По результатам рецензирования автор отправляет оргкомитету письмо о согласии, которое предусматривает дальнейшую возможность размещения статей, вошедших в сборник конференции, в открытом электронном доступе на иных ресурсах по усмотрению редакции сборника.

КЛЮЧЕВЫЕ ДАТЫ КОНФЕРЕНЦИИ

1 октября 2018г.	начало подачи материалов для участия в конференции
15 декабря 2018г. ***25 января 2019г.***	срок получения материалов для участия в конференции Оргкомитетом срок предоставления рецензий на статьи
20 января 2019г.	срок принятия решения о публикации присланных материалов и рассылки приглашений для участия в конференции и сообщение о включении статьи в Сборник материалов конференции OSTIS
10 февраля 2019г.	размещение на сайте конференции http://conf.ostis.net проекта программы конференции
15 февраля 2019г.	размещение на сайте конференции http://proc.ostis.net Сборника материалов и Программы конференции OSTIS-2019
21 февраля 2019г.	регистрация участников и открытие конференции OSTIS-2019
21-23 февраля 2019г.	работа конференции OSTIS-2019
26 февраля 2019г.	публикация фоторепортажа и отчёта о проведённой конференции на сайте конференции: http://conf.ostis.net
15 марта 2019г.	загрузка материалов сборника конференции в РИНЦ

ФОРМИРОВАНИЕ ПРОГРАММЫ КОНФЕРЕНЦИИ

Программа конференции формируется Программным комитетом по результатам рецензирования, представленных статей, а также на основании подтверждения автора(-ов) статьи о прибытии на конференцию.

КОНТАКТНЫЕ ДАННЫЕ ОРГАНИЗАТОРОВ КОНФЕРЕНЦИИ OSTIS

Вся необходимая информация по предстоящей и предыдущих конференциях OSTIS находится на сайте конференции http://conf.ostis.net, а также на сайте материалов конференции http://proc.ostis.net.

По вопросам участия в конференции и решения спорных вопросов обращайтесь: ostisconf@gmail.com.

Методическая и консультативная помощь участникам конференции осуществляется только через электронную почту конференции.

Конференция проходит в Республике Беларусь, г. Минск, ул. Платонова, 39 (5-ый учебный корпус Белорусского государственного университета информатики и радиоэлектроники).

OSTIS-2019

9th international scientific and technical conference

Open Semantic Technologies for Intelligent Systems

February 21-23, 2019 Minsk. Republic of Belarus

C A L L   F O R   P A P E R S

We invite you to take part in IX International Scientific and Technical Conference "Open Semantic Technologies for Intelligent Systems" (OSTIS-2019), which will focus on areas of use of the semantic technologies.

Conference will take place from ***February, 21st to February, 23rd, 2019*** at the Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus.

Conference proceedings language: English

Working languages of the conference: Russian, Belarusian, English

MAIN ORGANIZERS OF THE CONFERENCE

- Russian Association of Artificial Intelligence (RAAI)
- Belarusian State University of Informatics and Radioelectronics (BSUIR)
- State Institution "Administration of High Technologies Park" (Republic of Belarus)
- Ministry of Education
- Ministry of Communications and Informatization

CONFERENCE TOPICS:

- *Underlying principles of semantics-based knowledge representation, and their unification.*
  Types of knowledge and peculiarities of the semantics-based representation of various knowledge and metaknowledge types.
  Links between knowledge; relations, that are defined on the knowledge.
  Semantic structure of a global knowledge base, that integrates various accumulated knowledge.
- *Parallel-oriented programming languages for processing of the semantics-based representation of knowledge bases.*
- *Models for problem solving, that are based on knowledge processing, which occurs directly at the semantics-based representation level of knowledge being processed. Semantic models of information retrieval, knowledge integration, correctness and quality analysis of knowledge bases, garbage collection, knowledge base optimization, deductive and inductive inference in knowledge bases, plausible reasoning, pattern recognition, intelligent control. Integration of various models for problem solving*
- *Semantic models of environment information perception and its translation into the knowledge base.*
- *Semantic models of multimodal user interfaces of intelligent systems, based on the semantic representation of knowledge used by them, and unification of such models.*
- *Semantic models of natural language user interfaces of intelligent systems. The structure of semantic representation of linguistic knowledge bases, which describe natural languages and facilitate solution of natural language text and speech interpretation problems, and of natural language texts and speech messages synthesis, that are semantically equal to certain knowledge base fragments.*
- *Integrated logic-semantic models of intelligent systems, based on semantic knowledge representation, and their unification*
- *Various technical platforms and implementation variants of unified logic-semantic models of intelligent systems, based on semantic knowledge representation*
- *Models and means, that are based on the semantic representation of knowledge and that are oriented to the design of various typical components of intelligent systems (knowledge bases, programs, problem solvers, user interfaces).*
- *Models and means, that are based on semantic representation of knowledge and that are oriented to the complex design of various classes of intelligent systems (intelligent reference systems, intelligent learning systems, intelligent control systems, intelligent robotics systems, intelligent systems for design support etc.)*
- *Applied intelligent systems, that are based on the semantic representation of knowledge used by them*

CONFERENCE GOALS AND FORMAT

The goal of the conference is to discuss problems of creation of the **Open Complex Semantic Technology for Intelligent Systems Design**. This determines the Conference format, which involves (1) plenary reports; (2) workshops; (3) round tables, dedicated to discussion of various questions of creating of such technology; (4) poster sessions.

During the **poster sessions** every participant of the conference will have an opportunity to demonstrate his results. Conference format assumes exact start time of each report, and exact time of its exhibition presentation.

One of the major objectives of the conference is to attract not only scientists and postgraduate students, but also students who are interested in artificial intelligence, as well as commercial organizations willing to collaborate with research groups working on the development of modern technologies for intelligent systems design.

PARTICIPATION TERMS AND CONDITIONS

All those interested in artificial intelligence problems, as well as commercial organizations willing to collaborate with research groups working on the development of modern technologies for intelligent systems design are invited to take part in the Conference.

To participate in the OSTIS-2019 conference, it is necessary to register in the CMT system before December 15, 2018, find conference page, and from there:

- submit a **participation form** for the OSTIS-2019 conference. Each participation form field is required, including indication of the reporter. By filling in the registration form, you agree that your personal data will be processed by the Organizing Committee of the Conference, and that the paper and information about the authors will be published in printed and electronic format. Participation form should contain information on all of the authors. If author(s) are participating with a report, participation form should have their **color photo(s)** attached (they are needed for the Conference Program);
- upload an **article** for publication in the OSTIS-2019 Conference Proceedings. Papers should be formatted according to the provided template (see http://proc.ostis.net/eng/autors.html). Four full pages is a minimum size of a paper.
- send the signed **scan of the letter of consent**

If a report is submitted to participate in one of the contests, this intention should be clearly indicated in the participation form.

The selection of papers for publication in the Conference Proceedings and participation in the Conference is performed by a number of reviewers from among the members of the Conference Program Committee.

Incompliant applications and papers will be rejected.

Authors, whose articles were included in the Conference Program, will receive the invitations for participating in the Conference before January 30th, 2017.

Conference participation does not require any fees.

PAPERS SUBMISSION PROCEDURE

Papers (only on topics mentioned above) should be submitted ready for publication (http://proc.ostis.net/eng -> For authors). The text should be logically complete and contain new scientific and practical results. Each author is allowed to submit two reports maximum.

After the article was submitted, it is sent for review. Review results will become available to the paper author(s) on the CMT website before January 25th.

The Organizing Committee reserves the right to reject any paper, if it does not meet the formatting requirements and the Conference topics, as well as if there was no participation form submitted for the paper.

YOUNG SCIENTIST REPORTS CONTEST

Authors of the report submitted to the contest may include scientists with scientific degrees, but the report should be made by those without a degree and under 35 years old.

To take part in the young scientists report contest, it is necessary to:

1) fill in the participation form, where your participation in the contest is clearly indicated;

2) write an article and upload it to the CMT website;

3) fill in, sign, scan and send letter of consent via the email.

4) make a report at the conference (in person);

YOUNG SCIENTIST PROJECTS CONTEST

Projects of applied intelligent systems and systems aimed at supporting the design of intelligent systems are allowed to take part in the contest; they have to be presented by a scientist without a degree and under 35 years old.

To take part in the young scientist projects contest, it is necessary to:

1) fill in the participation form, where your participation in the contest is clearly indicated;

2) write an article and upload it to the CMT website;

3) make a report at the conference (in person);

4) make an exhibition presentation of the software

STUDENT INTELLIGENT SYSTEM PROJECTS CONTEST

To participate in the contest, a project must meet the following criteria: (a) it was developed by students and/or undergraduates of the higher education institutions, and (b) project consultants and advisors must hold a scientific degree and title. To participate in this contest, it is necessary to:

1) familiarize yourself with contest's terms and conditions (http://www.conf.ostis.net);

2) fill in the participation form for the contest (http://www.conf.ostis.net);

3) prepare a summary of the project (http://www.conf.ostis.net).

4) submit the participation form and project summary to the student projects' email address: ostis.stud@gmail.com.

CONFERENCE PROCEEDINGS PUBLICATION

The Conference Organizing Committee plans to publish the papers selected by the Program Committee based on the results of their review, in the Conference Proceedings, on the official Conference website http://conf.ostis.net and on the Conference Proceedings website http://proc.ostis.net.

Upon successful review author sends a letter of consent to the Organizational Commitee. Author therefore agrees that his paper can be made freely available in electronic form at other resources at the Editorial Board's discretion.

KEY DATES OF THE CONFERENCE

October 1st, 2018	paper submission opens
December 15th, 2018 ***January 25th, 2019***	paper submission deadline paper review deadline
January 20th, 2019	final decision on paper publication; sending out invitations and notifications on inclusion of a paper in the OSTIS-2017 Conference Proceedings
February 10th, 2019	Draft Conference Program publication on the conference website http://conf.ostis.net
February 15th, 2019	Conference Proceedings and Conference program publication on the conference website http://proc.ostis.net
February 21st, 2019	Participant registration and OSTIS-2019 conference opening
February 21st to 23rd, 2019	OSTIS-2019 conference
February 26, 2019	photoreport and conference report publication on the conference website: http://conf.ostis.net
March 15th, 2019	conference proceedings will be uploaded to the Russian Science Citation Index database

CONFERENCE PROGRAM FORMATION

Conference program is formed by the Program Commitee according to the paper review results; author(s)' confirmation of participation is required as well.

CONTACTS

All the necessary information about the forthcoming and previous OSTIS Conferences can be found on the conference website http://conf.ostis.net and http://proc.ostis.net.

For questions regarding conference participation and dispute resolution please contact: ostisconf@gmail.com.

Methodological and advisory support to the conference participants shall be provided through the conference e-mail only.

The conference venue is the 5th academic building of the Belarusian State University of Informatics and Radioelectronics (39, Platonov str., Minsk, Republic of Belarus)

ИТОГИ

Международной научно-технической конференции OSTIS-2017

(Open Semantic Technology for Intelligent Systems − Открытые семантические технологии проектирования интеллектуальных систем)

16-18 февраля 2017 года в Белорусском государственном университете информатики и радиоэлектроники прошла **VII-я Международная научно-техническая конференция** «Открытые семантические технологии проектирования интеллектуальных систем» (**OSTIS-2017**), которая была посвящена памяти выдающегося ученого, профессора Рижского Технического университета, доктора технических наук, хабилитированного доктора компьютерных наук, создатель Латвийской научной школы по теории принятия решений, методам обработки нечёткой информации и интеллектуальным технологиям – Аркадий Николаевич Борисов.

Основной темой VII-ой конференции OSTIS является онтологическое проектирование интеллектуальных систем и их компонентов.

Основной целью ежегодных конференций OSTIS (Open Semantic Technology for Intelligent Systems) является создание условий для расширения сотрудничества различных научных школ, вузов и коммерческих организаций, направленного на разработку и применения комплексной массовой и постоянно совершенствуемой технологии компонентного проектирования интеллектуальных систем.

По результатам рецензирования, представленных статей Программным комитетом были сформированы Сборник материалов конференции и программа конференции. В соответствии с программой был определён **формат проведения конференции OSTIS-2017**.

Мы стремимся к тому, чтобы все желающие могли принять активное участие в обсуждении представленных докладов и подготовились к этому обсуждению. Для этого участники конференции могли познакомиться с текстами докладов, опубликованными на сайте конференции до начала конференции. Благодаря этому докладчикам не было необходимости озвучивать этот текст, а акцентировать внимание на ключевые его положения. Кроме того, это позволило задать докладчикам большее число вопросов и больше времени посвятить обсуждению вопросов, затронутых в докладах.

Каждому докладу по желанию докладчиков было предоставлено место и время для выставочной презентации, где докладчики могли во время проведения выставки обсудить свои научные результаты и продемонстрировать разработанные ими системы.

Всего было **опубликовано 73 статьи**, прошедшие рецензирование Программным комитетом, из которых было **заслушано 40 доклада**. Среди них: 14 докладов докторов и 8 докладов кандидатов наук, а также 15 докладов молодых учёных.

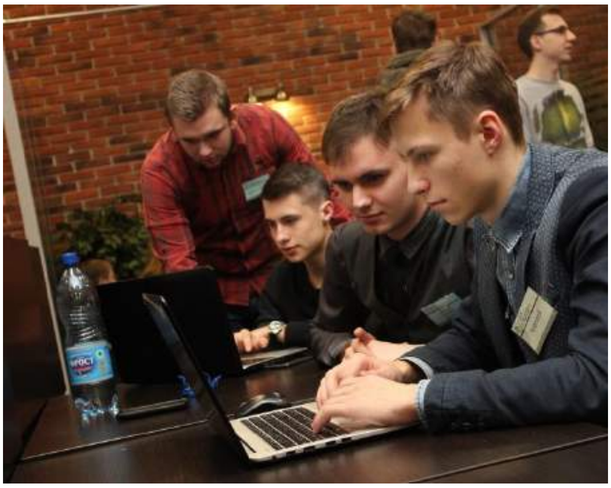
Всего же в работе конференции приняло 25 докторов наук, 31 кандидат наук и более 150 студентов, магистрантов и аспирантов различных вузов и различных городов.

География участников конференции OSTIS-2017 весьма обширна и охватывает 24 города России, Беларуси, Украины, Казахстана, Узбекистана и Латвии: Москва, Санкт-Петербург, Владивосток, Новосибирск, Иркутск, Волгоград, Ростов-на-Дону, Казань, Самара, Ульяновск, Тверь, Пермь, Апатиты, Томск, Тюмень, Йошкар-Ола, Минск, Барановичи, Брест, Гродно, Гомель, Киев, Одесса, Сумы, Кременчуг, Астана, Алмата, Ташкент, Рига.

Доклады, представленные на конференции, уже в первый день вызвали оживленные дискуссии в кулуарах.

Уже второй год подряд второй день конференции проходит в Бизнес-инкубаторе ПВТ – место встречи науки, образования, инженирии и бизнеса.

Традиционно в рамках конференции OSTIS-2017 было проведено 3 конкурса: конкурс докладов молодых ученых, конкурс программных продуктов и конкурс студенческих проектов.





Наталья Бурец и Сергей Леонидович Чехменок отметили понравившиеся проекты ценными призами.

Интеллектуальная система по дискретной математике

Интеллектуальная система по химии

Мобильное приложение для управления автономным роботом-экскурсоводом

Моделирование воздействия импульсных магнитных полей на мозг человека

Туристический веб-портал

Интеллектуальная справочная система «Генеалогия»

В рамках конференции был проведен **круглый стол** на специализированной площадке Бизнес-инкубатора ПВТ, посвященный ***взаимодействию науки, образования, инженерии и бизнеса в области разработки интеллектуальных систем***. В нём приняли участие студенты, магистранты и аспиранты БГУИР, представители науки, образования и бизнеса. Основные тезисы круглого стола:

- В настоящее время важнейшим направлением развития компьютерных систем является использование методов и средств искусственного интеллекта.
- Переход от традиционных компьютерных систем к промышленной разработке интеллектуальных систем требует создания таких технологий проектирования интеллектуальных систем, которые могли бы быстро эволюционировать, но при этом гарантировать совместимость проектируемых интеллектуальных систем (совместимость баз знаний и машин обработки знаний).
- Технологии проектирования интеллектуальных систем должны разрабатываться, с одной стороны, путем интеграции научных результатов, полученных в этой области, с другой стороны, путем накопления и обобщения опыта разработки конкретных интеллектуальных систем, и, с третьей стороны, путем накопления опыта подготовки специалистов, способных разрабатывать интеллектуальные системы и совершенствовать технологии их проектирования.
- Без тесного сотрудничества науки, образования, инженерии и бизнеса прогресс в области промышленной разработки интеллектуальных систем невозможен и эпицентром такого прогресса должно быть создание инфраструктуры, обеспечивающей высокие темпы развития технологий проектирования интеллектуальных систем.
- Если мы хотим занять достойное место на рынке интеллектуальных систем, указанное выше сотрудничество необходимо начинать как можно скорее, т.к. технологии разработки интеллектуальных систем настолько отличаются от традиционных, что быстрый переход к промышленной разработке интеллектуальных систем невозможен.
- От обсуждения целесообразности тесного сотрудничества науки, образования, инженерии и бизнеса в области массовой промышленной разработки интеллектуальных систем, необходимо перейти к уточнению форм и направлений такого сотрудинчества. Если мы не перейдем к разработке промышленных интеллектуальных систем сегодня, то завтра это сделают без нас.


В последний день конференции большинство докладов было посвящено прикладным системам в области искусственного интеллекта.



Перед **закрытием конференци** состоялось активное общение представителей научной среды и молодёжи:









РЕШЕНИЯ КОНФЕРЕНЦИИ:

- провести VIII Международную научно-техническую конференцию OSTIS 15-17 февраля 2018 года;
- расширить географию участников конкурса студенческих проектов интеллектуальных систем и конкурса докладов и проектов молодых учёных;
- в рамках конференции OSTIS-2018 провести конкурс докладов и проектов молодых учёных;
- в рамках конференции OSTIS-2018 провести круглый стол, посвященный проблемам развития рынка систем, основанных на знаниях.