

Abstracts of Plenary Reports

Artificial Intelligence in Medicine: Trouble Spots

B.A. Kobrinskii (*kba_05@mail.ru*)

Federal Research Center «Computer Science and Control»,
RAS, Moscow

The article discusses various aspects of building intelligent medical systems. The main attention is paid to "trouble spots", which include: knowledge extraction, visual images, explainable machine learning and artificial neural networks, diagnostics of combined pathology, support at various stages of clinical decision-making, intellectual cyber-physical systems, orientation on user specialty, extraction intelligent systems of personal data of patients from electronic medical records for the advancement and substantiation of diagnostic hypotheses.

Infocommunication Technology and Natural Language: Regional Experience

D.Sh. Suleymanov (*dvd.slt@gmail.com*)

Academy of Sciences of the Republic of Tatarstan,
Institute of Applied Semiotics, Kazan Federal University, Kazan

The preservation and development of a natural language in the era of widespread digitalization is directly related to the use of language as a means of accumulation, transmission of information, as well as a means of communication in ICT and in the WWW. The paper presents the experience of research activities in Institute of Applied Semiotics of the Academy of Sciences of Tatarstan in NLP scientific area.

Abstracts of Lectures

Functional-Algebraic Models of Intelligent Data Analysis and Methods of Their Construction

I.Z. Batyrshin (*batyr1@gmail.com*)

Centro de Investigacion en Computacion,
Instituto Politecnico Nacional, Mexico

We discuss a functional-algebraic approach to data analysis, the main components of which are functions of relationship between the elements of an involutive set. An involutive set is a set on which a unary operation $N(x)$ can be defined that satisfies the involutivity condition $N(N(x)) = x$. A set with such an operation forms an involutive algebra. Involutive operations and sets arise both in Boolean algebras (negation operations, complement operations) and in linear algebras (mappings to numbers with opposite signs, division operations). The functions of relationship are measures of similarity, proximity, correlation and association coefficients, metrics, and others used in data mining, pattern recognition, and machine learning problems. The paper considers methods for constructing functions of similarity, dissimilarity, and correlation on involutive sets of various types.

Formalization of Insight Processes

V.G. Redko (*vgredko@gmail.com*)
Scientific Research Institute for System Analysis
of the Russian Academy of Sciences

A.V. Samsonovich (*avsamsonovich@mephi.ru*)
National Research Nuclear University “MEPhI”
(Moscow Engineering Physics Institute)

The principles of formalization and computer modeling of insight processes have been developed. Examples of cognitive illumination processes at different evolutionary levels and in different living beings are considered. Examples are given for New Caledonian crows, for chimpanzees, for insight processes of in human scientific cognition. In particular, insights by Archimedes, Poincaré, Newton are considered. According to examples, there are similar features of insight processes: a) there is no visible progress in solving a problem for a long time, b) the problem is remembered and leaves the working memory, c) the solution to the problem is found by a sudden guess, by insight. The work introduces the concepts of the value of prediction, the value of an idea. It is shown that in the process of insight, predictions with high value play an essential role, and it is possible to link high-value ideas presented in explicit and implicit memory. When analyzing the processes of insight, conceptual approaches developed by psychologists were used. Computer modeling of simple abstract examples of processes that characterize insights has been performed.

Abstracts

“Russian Advances in Artificial Intelligence 2021. RAAI 2021”.

Topic 1

General Issues of Artificial Intelligence and Knowledge Engineering

Empirion Ontology for Description of Variables in Datasets Obtained in Experimental Studies

A.M. Begler (*alena.begler@gmail.com*)

I.A. Leshcheva (*leshcheva@gsom.spbu.ru*)

Graduate School of Menegement SPbU, St. Petersburg

By now there are tens of thousands of potentially reusable datasets in open data repositories. However, their reuse is often hampered by insufficient data documentation. Existing ontology models describe data mostly at the project and dataset level and there are few descriptions of variables in the dataset. We propose an empirion ontology that allows to associate variables with the information needed to interpret them – the domain of values and units – and thus allows to connect variables from different datasets. The ontology is available at <https://github.com/jimijimiyo/empirion>.

AI: “ ... As You Call the Boat, so It Will Float”

V.K Finn (*v.k.finn@yandex.ru*)

FRC “Computer Science and Control” RAS

M.A. Mikheyenkova (*m.mikheyenkova@yandex.ru*)

FRC “Computer Science and Control” RAS

M.I. Zabezhailo (*m.zabezhailo@yandex.ru*)

FRC “Computer Science and Control” RAS

Modern content for the concept of AI is discussed. Key specific features characterizing AI as a research and development area are presented. Key differences between intelligent data analysis and computer data analysis (as a whole) are demonstrated.

Cognitive Generalization Skills for Ontology and Knowledge Graphs Development

T.A. Gavrilova (*gavrilova@gsom.spbu.ru*)

Graduate School of Management, St. Petersburg State University, St. Petersburg

Structuring knowledge problems are considered from the cognitive psychology view. The paper discusses the influence of cognitive styles at the training of visual-analytical thinking skills in the framework of ontology engineering training.

Building an Ontological Representation of an Enterprise Database Using Open Knowledge Bases Technologies

V.V. Sherkunov (*v.sherkunov@ulstu.ru*)
Ulyanovsk State Technical University, Ulyanovsk

Currently, ontological engineering has become a developing scientific direction. Since the traditional way of constructing ontologies is its manual compilation by an expert in the subject area, in this regard, such a direction as the automated construction of ontological models has gained interest. This article discusses a variant of the automated construction of ontologies by mapping an existing enterprise database into an ontological representation, with subsequent expansion through open knowledge bases such as WikiData.

Using Ontology Design Patterns to Develop Decision Support Ontology

Yu.A. Zagorulko (*zagor@iis.nsk.su*)
G.B. Zagorulko (*gal@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems of Siberian Branch of RAS,
Novosibirsk, Russia

The paper is devoted to development and application of ontology design patterns in the construction of a decision support ontology. The requirements for the ontology and the main groups of patterns used are described. The paper give examples of content patterns that are used to define the main classes of the ontology and their instances, as well as patterns for defining a multidimensional structure of an ontology and organizing interaction with data from external sources.

Topic 2

Data Mining, Machine Learning, Natural Language Processing, Semantic Web

On VKF Method Foundations Based on Category Theory

D.V. Vinogradov (*krrguest@yandex.ru*)

Federal Research Center Informatics and Management RAS, Moscow

The paper describes the category of semilattices as algebras over a monad, in which the free algebra corresponds to the sets of similarity parents, and the structural operation corresponds to the similarity operations between these parents. This approach allows us to uniformly justify the extendable variants of both the JSM- and the VKF-method of machine learning.

Annotated Russian Corpus PaRuS. Version 1.1

N.A. Vlasova (*nathalie.vlassova@gmail.com*)

I.N. Vozdvizhensky (*vozdvin@yandex.ru*)

Yu.P. Serdyuk (*yury@serdyuk.botik.ru*)

I.V. Trofimov (*itrofimov@gmail.com*)

Program Systems Institute of RAS, Pereslavl-Zalessky

In this article we present an annotated Russian language corpus named PaRuS (Parsed Russian Sentences). We consider the methodology of corpus creation and its latest updates. We also discuss the quality of linguistic annotation in PaRuS.

An Approach to Automated Acquisition of a Subject Domain Thesaurus from a Sample of Scientific Abstracts

A.V. Voinov (*avoinov@gmail.com*)

Saint Petersburg Polytechnic University, Saint Petersburg

An approach to acquire a hierarchical thesaurus of a subject domain, defined by abstracts, resulting from a query to Pubmed. A deep NLP is used together with methods of qualitative data analysis. The latter methods are key to establish an ad hoc semantic hierarchy of terms, which is intrinsic to the subject domain. The proposed method allows to build a balanced tree of clusters, which may serve an initial approximation to an ontology of the subject domain. No apriori ontology of a wider subject domain is needed. Semantic associations, which are implicit to the corpus of analyzed texts, are elicited via custom methods of multidimensional scaling and cluster analysis.

Hybrid Method of Active Learning for Not Collectively Exhaustive Multiclass Classification by Unlabeled Noisy Data

D.I. Dranga (*dranga.daniil@gmail.com*)

National Research University Higher School of Economics, Moscow

R.V. Senchenko (*CdeSenchenkoRV@gmail.com*)

PJSC Sberbank, Retail Block, Customer Care Division, Moscow

The paper investigates the mathematical support of dialog agents, an important component of which is the intent classifier - a model that determines the intention contained in the message. For example, defining intentions to open a card or disable SMS notification by a chatbot in a bank. To train the model, the data marked up by assessors is required, selected from the unmarked dialogs. Object selection is complicated by the fact that the set of intents of interest is a subset of all the possible intents contained in client messages, and many messages do not contain intents. The paper proposes a hybrid algorithm that takes into account the features of the problem and combines unsupervised learning, filtering based on the similarity of object embeddings, and an evaluation of the information usefulness of clusters. The proposed algorithm is universal and does not depend on the chosen text classification model.

Applying Pattern Structures to NLP Tasks

E.F. Goncharova (*egoncharova@hse.ru*)

National Research University Higher School of Economics, Moscow

S.O. Kuznetsov (*skuznetsov@hse.ru*)

National Research University Higher School of Economics, Moscow

The paper presents the research on applicability of pattern structures to the various NLP tasks, and information retrieval, specifically. We provide the modification of one of the most popular structural approach for constructing the representation of the text which utilizes word embeddings. We also analyze the results of experimental evaluation of the observed methods.

Topic 3

Cognitive and psychological research, behavior, ethics, emotions, computer vision and images in artificial intelligence

Building a Model of The Dynamics of Cognitive Activity Based on Analysis of Multi-Dimensional Signals

N.I. Bodrina (*vavilovani@mail.ru*)

N.N. Filatova (*nfilatova99@mail.ru*)

K.V. Sidorov (*bmsidorov@mail.ru*)

Tver State Technical University, Tver, Russia

The paper describes the features of constructing a model of the dynamics of mental (cognitive) activity based on determining the characteristics of the trajectories of a two-dimensional projection of a multidimensional phase portrait on a plane formed by two leads of EEG signal. The results of experiments illustrating an increase in cognitive activity (a decrease in the computation time, a decrease in the number of mistakes) in the process of sequential emotigenic stimulation are presented.

Animate Navigation on Visual Landmarks with Scene Recognition

I.P. Karpova (*karpova_ip@mail.ru*)

HSE University, Moscow

The paper proposes a method of using visual landmarks for memorizing the traversed path by a mobile robot (animate), based on the navigation mechanism of ants. The model of route presentation and the rules of its interpretation are described. They allow another robot to repeat the route taken by the scouting robot. The results of experiments on simulation modeling demonstrate the possibility of using the developed method for solving problems of foraging and reconnaissance.

Schematization of Images of Knowledge in Heterogeneous Model Field of Cognitive Functional Hybrid Intelligent Systems

A.V. Kolesnikov (*avkolesnikov@yandex.ru*)

Immanuel Kant Baltic Federal University

S.B. Rumovskaya (*sophiyabr@gmail.com*)

Kaliningrad Branch of the FRC “Computer Science and Control” of the RAS

S.A. Soldatov (*ssacompany@mail.ru*)

LLC MTS Digital

E.V. Jasinski (*ejasinski96@gmail.com*)

Immanuel Kant Baltic Federal University

Cognitive functional hybrid intelligent systems are considered as automatic partners of operational activities, integrators of cognition through mathematics, morphology and semiotics, isographers and the language of role-based visual relations for schematizing of figurative representations in a heterogeneous model field.

Topic 4

Fuzzy models, soft computing, multi-agent systems and artificial communities

Fuzzy Situational Control of Complex Technical Systems Based on Composite Hybrid Models

V.V. Borisov (*vbor67@mail.ru*)

D.Yu. Avramenko (*leyzi.small@yandex.ru*)

The Branch of National Research University “Moscow Power Engineering Institute” in Smolensk

The paper considers the proposed method of fuzzy situational control of complex technical systems (CTS) based on composite hybrid models. The method consists in using a pre-designed fuzzy situational network to identify current situations, search and select sequences of decisions when translating the CTS into target situations. The method increases the efficiency of system management, depending on the control strategy and constraints. It organically combines the processes of compositional modeling and fuzzy situational control.

Co-Evolutionary Algorithm of Placement Based on Models of Adaptive Behavior of Ant Colony

B.K. Lebedev (*lebedev.b.k@gmail.com*)

O.B. Lebedev (*lebedev.ob@mail.ru*)

Southern Federal University, Taganrog

To solve the placement problem, a co-evolutionary algorithm based on the ant colony method has been developed. The essence of the co-evolutionary approach is that the evolving population is divided into subpopulations that evolve in parallel. Periodically, agents move from one subpopulation to another, passing on their experience. The co-evolutionary approach provides a broader view of the solution space and a higher probability of localizing the global extremum of the problem. This approach allows to partially solve the problem of premature convergence, provides a way out of local optima and increases the speed of obtaining the result.

On the Rational Activity on an Intelligent Agent

M.A. Mikheyenkova (*m.mikheyenkova@yandex.ru*)

FRC “Computer Science and Control” RAS

The paper deals with the formalization of an artificial agent activity using the representation of the agent's actions by logical means. The proposed approach characterizes the rationality of the intelligent (cognitive) agents' activity from the logical consistency point of view. The presentation of rationality based on an argued choice of actions is also considered.

The Model of Spots as the Basis for Qualitative Reasoning and Granular Computing

N.A. Simonov (*nsimonov@ftian.ru*)

Institute of Physics and Technology named after K.A. Valiev RAS, Moscow

The basis of the mathematical apparatus of the recently proposed model of spots, which simulates figures with vague shapes, is considered. The main feature of our approach is that both the shape of the spot and properties of its surrounding space are initially unknown. They can be identified only based on information about the elementary relations of the spot with some other spots. We consider crisp figures to be a special and limiting case of spots.

The proposed model of spots is a mathematical object intended for representation qualitative data or semantic aspects in data, as well as for modeling human perception and reasoning. The considered model is applicable in many areas of artificial intelligence. For example, spots can be used for mental representation in the computational theory of mind, as well as for modeling the semantic properties of knowledge in knowledge bases. The mathematical apparatus of spots can be used to represent and process knowledge, model reasoning, build neural networks of a new type, in the field of natural language processing, and can also become the basis for the theory of granular computations.

The developed mathematical apparatus was tested by solving problems of image reconstruction of flat crisp figures, using only qualitative data on their elementary relationships with a variety of known basis figures.

Topic 5

Intelligent Decision support and Control Systems, Robotic Systems, Intelligent Systems Design Tools

Organization of Models Translation for Model-Based Design Tools

A.A. Burenko (*burenko@sfedu.ru*)

M.Yu. Polenov (*mypolenov@sfedu.ru*)

Institute of Computer Technologies and Information Security
of Southern Federal University, Taganrog

Purpose of the paper is to review current solutions in the field of models' translation for model-based design tools, and to analyze their inherent approaches to models reuse. Relevance of the solutions is reviewed in regards to the modeling of modern complex technical systems. Proposed principles for converting model codes make it possible to upgrade existing tools and to increase the efficiency of the implemented model conversion modules.

On the Implementation of a Temporal Data Representation Model Based on a Non-Relational Graph Database for Real-Time Decision Support Systems

A.P. Ereemeev (*eremeev@appmat.ru*)

N.A. Paniavin (*PaniavinNA@mpei.ru*)

National Research University "Moscow Power Engineering Institute", Moscow

The paper considers the possibilities of implementing a model for representing temporal data and knowledge based on a non-relational graph database using the Neo4j DBMS. The proposed model for the presentation of temporal data is intended to be included in modern intelligent decision support systems focused on dynamic subject/problem areas.

Topic 6

Neural Network Methods, Neuroinformatics

Application of Case-Based Reasoning Approach to Training Neural Networks

P.R. Varshavskiy (*varp@appmat.ru*)

National Research University “Moscow Power Engineering Institute”, Moscow

An.V. Kozhevnikov (*antoko@yandex.ru*)

National Research University “Moscow Power Engineering Institute”, Moscow

Al.V. Kozhevnikov (*alexkoz27@gmail.com*)

Moscow State University, Moscow

The work is devoted to an urgent problem in the field of artificial intelligence related to the research and development of original methods and algorithms for training artificial neural networks. The main attention in the work is paid to the study of the promising possibility of using the case-based reasoning approach for training convolutional neural networks for data mining purposes. An algorithm for training artificial neural networks is proposed using a case-based approach, which uses data on iterations in the process of training neural networks as precedent information, accumulates experience (cases) and adapts to adjust the parameters of the training algorithm for artificial neural networks.

Topic 7

Intelligent Technologies and Applied Intelligent Systems in Industry and Social Sphere

Methods of Data Mining in Criminalistics: Features of Application and of Computer Implementation

S.M. Gusakova (*svem45@yandex.ru*)

Federal Research Center «Computer Science and Control», RAS, Moscow

N.V. Ogoreltseva (*nat@gmail.com*)

Master's Degree in the Department of Intelligent Systems
in the Humanities of RSUH, Moscow

The report discusses the features of the use of data mining methods and their computer implementation in the study of the influence of psychological characteristics of a person on the features of performing a signature, taking into account additional parameters.

Intellectual Capital Ontology of Innovation-Oriented Production Company

V.D. Moskaleva (*lera.mosk@gmail.com*)

O.V. Stoianova (*o.stoianova@spbu.ru*)

Saint-Petersburg State University, Saint-Petersburg

This paper presents an ontology of intellectual capital, designed for research and development management in an innovation-oriented manufacturing company. The ontology is the result of combining the developed private ontologies of innovation, human, organizational (structural) and social capital, as well as the mechanism of knowledge extraction. For each private ontology there is a description of classes, properties and relationships that form a model of knowledge to support decision-making on research and development management. The result can be useful for specialists engaged in the creation of formalized methods and tools of intellectual capital management, as the created ontology has wide opportunities for expansion and scaling.

Method for Determining the Bifurcation Point of a Continuous Multi-Stage Technological Process Using a Multi-Sensor Measurement System

B.V. Palukh (*pboris@tstu.tver.ru*)
Tver State Technical University, Tver

A.N. Vetrov (*vetrov_48@mail.ru*)
Tver State Technical University, Tver

The article deals with the problem of detecting anomalies in complex dynamic systems under conditions of partial uncertainty inherent in continuous multi-stage production. Possible approaches to solving the problem of early warning of critical states of the technological process based on the data of continuous monitoring and analysis of the values of diagnostic indicators are considered. A new method and computational procedure for determining the bifurcation point of a continuous multi-stage technological process are proposed. The state of the system is estimated on the basis of short time series formed by the readings of a variety of sensors of technological equipment. Each sensor registers the values of a multidimensional random process, the characteristics of which are a priori unknown. The proposed method for detecting anomalies in a multi-sensor distributed measurement system allow combining the measurement results obtained from individual sensors. This allows us to draw conclusions about the possible transition of the technological process to a critical state, taking into account the correlation of data in time series. The connection of the proposed method for identifying bifurcation points with the ExpSystem expert system developed by the authors is described.

Topic 8

Intelligent technologies and applied intelligent systems in medicine

To the Creation of an Applied Intellectual System for Diagnostics of the Quality of Life of Patients with Neurological Pathology

A.E. Yankovskaya (ayyankov@gmail.com)
Tomsk State University, Tomsk

V. Obukhovskaya (diada1991@gmail.com)
Siberian State Medical University, Tomsk

The article outlines the basics of creating an applied intelligent system for diagnosing the quality of life (IS DIAKAGE) for patients with neurological pathology. Applied IS DIAKAGE is based on test methods of pattern recognition, non-redundant unconditional diagnostic tests, mixed diagnostic tests and is constructed on the basis of the intelligent tool IMSLOG. Applied IS DIAKAGE is designed to identify various regularities in the quality of life of patients based on parameters (signs) that determine the characteristics of the physical and mental components of the health of patients with neurological pathology, as well as to make diagnostic decisions and justify them using graphical, including cognitive, funds. Diagnostics of the quality of life of patients with neurological pathology on the basis of the DIAKAGE applied IS is very important for the subsequent process of adaptation of patients to the existing disease and the prognosis of rehabilitation measures.

Abstracts

“Artificial Intelligence. RCAI 2021. Lecture Notes in Computer Science. Springer, 2021”

Topic 1

Cognitive Research

Heterogeneous Formal Neurons and Modeling of Multi-Transmitter Neural Ensembles

Nikolay Bazenkov (*n.bazenkov@yandex.ru*)

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences

A multitransmitter neural ensemble is a group of neurons interacting not via isolated synaptic connections, but via emission of neurotransmitters directly into the shared extracellular space (ECS). There are multiple experimental evidence that non-synaptic interactions play the important role in biological neural circuits. We propose a model of multitransmitter neural ensembles where each neuron is represented as a finite state machine. An algorithm of neural interactions via the shared ECS is proposed. This framework allows to capture the variety of spiking behavior observed in biological neurons. The model is intended primarily for simulation of simple neural ensembles where each neuron has a unique internal properties and plays the specific role in the ensemble activity. We show how the model can imitate such neural activity classes as tonic spiking, bursting, post-inhibitory rebound etc. To illustrate the key features of the proposed framework we model two examples of pattern-generating neural ensembles: a half-center oscillator and a feeding network of a pond snail.

Methods for Recognition of Frustration-Derived Reactions on Social Media

Dmitry Devyatkin (*devyatkin@isa.ru*),

Natalia Chudova (*nchudova@gmail.com*),

Anfisa Chuganskaya (*anfisa.makh@gmail.com*)

Federal Research Center “Computer Science and Control,” Russian Academy of Sciences, Moscow, Russia

Daria Sharypina (*sharypinadv@gmail.com*)

Moscow Institute of Physics and Technology, Moscow, Russia

In this paper, we attempted to find speech features of different reactions to frustration to detect and classify them in social media texts. Frustration is a highly motivated situation in which it is impossible to achieve a goal when unexpected external or internal obstacles

are encountered to meet the need. We use a well-recognized typology of the reactions and focus on context-aware but straightforward models and classification features, which can be easily interpreted. The experiments show that pure lexis cannot be used as the only feature for the classification. Only the models, which combine different-level linguistic features, implicitly like in BERT or in the models with the linguistic patterns, provide fair results. From a psychological point of view, some misclassifications of the obtained reaction data can be related to their assignment to one class of extrapunitive reactions. Discussions in social networks suggest a high level of human activity, a desire to seek a solution to the problem in a broader social interaction. Thus, the focus on extrapunitive reactions and an increased emotional component in the form of aggression is a feature of that interaction type. On the one hand, we provide a method to classify the social network messages; on the other hand, the training results can be interpreted and analyzed by experts in psychodiagnostics.

Identification of the Network State Based on the ART-2 Neural Network With a Hierarchical Memory Structure in Parallel Mode

Vladimir Polyakov, Dmitriy Bukhanov and Maxim Panchenko
(*panchenko.maks@gmail.com*)

Belgorod State Technological University named after V.G. Shukhov,
Belgorod, Russia

A new memory structure for artificial neural networks of adaptive-resonance theory is proposed, which has a hierarchical form. For each new memory level, the previous classification value is refined by increasing the similarity parameter. This architecture was used to determine the network state in intrusion detection systems. The paper describes an algorithm for learning the proposed structure of the ART-2m network in parallel mode. A comparative analysis of the time characteristics of the network with the proposed structure when operating in series and parallel modes is carried out. Experiments were carried out using the NSL KDD-2009 sample, the results of which show the possibility of using ART-2m in intrusion detection systems. Most states were determined with sufficient accuracy that is greater than 80 percent.

Topic 2

Data Mining, Machine Learning, Classification

Ranking Weibull Survival Model: Boosting Concordance Index of Weibull Time-to-event Prediction Model with Ranking Losses

Cheloshkina Kseniia (*kcheloshkina@hse.ru*)

Laboratory of Bioinformatics, Faculty of Computer Science, National Research University Higher School of Economics, Moscow, Russia

The concordance index is a widely used metric for the evaluation of time-to-event prediction models. It describes the proportion of correctly ranked pairs of observations by time to event and hence is closely related to ROC AUC. In this paper, we propose to enrich such a baseline model as Weibull time-to-event feed-forward network which optimizes classic in survival analysis log-likelihood with additional concordance-aware loss components. Here we demonstrate that a combination of parametric survival analysis methods with a learning-to-rank approach forces the model to achieve higher concordance. The experiments over real-world datasets demonstrate the highly competitive performance of the proposed method called RWSM (Ranking Weibull Survival Model) in terms of concordance index.

Predicting Different Health and Lifestyle Behaviors of Social Media Users

Karim Khalil (*code.karim@gmail.com*),

Maksim Stankevich (*{stankevich,ivs}@isa.ru*)

Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

Ivan Smirnov, Maria Danina (*mdanina@yandex.ru*)

Federal Research Center Computer Science and Control of RAS, Moscow, Russia

Psychological Institute of Russian Academy of Education, Moscow, Russia

The rise of social media platforms and a growing interest in applying machine learning methods to ever increasing amounts of data creates an opportunity to use data from social media to predict lifestyle choices and behaviors. In this study we examine the possibility of using machine learning methods to classify users, of the Russian-speaking social media network: VKontakte, based on different health related activities and habits. Participants of this study took a survey that had questions about different health-related behaviors and activities and the intensity with which users follow them. We describe the process of gathering, processing, and using this data to train a set of machine learning classifiers and we evaluate the performance of these models in our experimental results. The features that were best able to classify most of the behaviors were collected from user subscription data. The best results were achieved on the questions about limiting alcohol use and limiting laptop and smartphone use (0.73 and 0.74 ROC AUC) with features generated from user profile and subscription data.

Methods for Finding Consequences with Specified Properties

Boris Kulik (*ba-kulik@yandex.ru*)

Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, Saint Petersburg, Russia

Alexander Fridman (*fridman@iimm.ru*)

Institute for Informatics and Mathematical Modelling, Kola Science Centre of the Russian Academy of Sciences, Apatity, Russia

In logical inference, two main tasks are relevant: verifying the correctness of an alleged consequence and inferring a consequence with predetermined properties. The entire theoretical potential of mathematical logic is mainly aimed at solving the former problem. As for the task of finding some desired consequences, it has obtained practically no useful developments. In this paper, we propose methods for solving the problem of finding consequences with certain properties on the basis on n-tuple algebra. These methods are shown to be applicable for solving constraint satisfaction problems as well.

Data Mining Methods for Analysis and Forecast of an Emerging Technology Trend: A Systematic Mapping Study From SCOPUS Papers

Nguyen Thanh Viet (*vietqn1987@gmail.com*)

Volgograd State Technical University, Volgograd, Russia

Alla Kravets^{1,3} (*agk@gde.ru*)

Dubna State University, Dubna, Russia

Duong Quoc Hoang Tu (*ctdqhtu@gmail.com*)

Pham Van Dong University, Quang Ngai, Vietnam

To stay competitive in an environment of rapidly changing science, it is important to monitor the development of existing technology and to discover new and promising technologies. Similarly, it is necessary for a firm to establish a technology development strategy through emerging technology forecast to gain a competitive edge while utilizing limited resources. Numerous methods of emerging technology trend analysis and forecast (TTAF) have been proposed; however, no study described data mining methods' review of this research area in a systematic and structured procedure. Hence, this paper intends to give a review of TTAF data mining methods and shortages by surveying and constructing challenging problems, research and resolving approaches. Moreover, the study highlights adopted data mining methods and types of data sources. Specifically, 50 documents from SCOPUS over a ten-year timespan between 2010 and 2019 were systematically reviewed, and each performing step was followed properly in accordance with systematic mapping study.

**Machine Learning for Assessment of Cardiometabolic Risk Factors
Predictive Potential and Prediction of Obstructive Coronary
Arteries Lesions**

Karina Shakhgeldyan (*carinashakh@gmail.com*), Andrey Vishnevskiy, Elena Emtseva
Vladivostok State University of Economics

Boris Geltser, Vladislav Rublev, Mikhail Tsivanyuk
Far Eastern Federal University, School of Medicine

The aim of this study was searching and validation of new obstructive coronary arteries lesions predictors and prognostic models development for its verification in patients with ischemic heart disease prior to invasive coronary angiography. Research included a step-by-step algorithm for predictors selection and validation as well as thresholds measurements with filtering and wrapping techniques. Cross-validation of predictive models based on multivariate logistic regression, support vector machine and random forest were made by averaging of 4 quality metrics. Based on selected predictors in continuous and categorical forms the best developed predictive model was logistic regression models ensemble with the following quality metrics: area under the ROC curve 0.85, accuracy - 0.80, sensitivity - 0.82, and specificity - 0.73, which is higher than the existing CAD Consortium scale.

Topic 3

Knowledge Engineering

Application of FCA for Domain Model Theory Investigation

Dmitry Palchunov (*palch@math.nsc.ru*)
Sobolev Institute of Mathematics, Novosibirsk, Russian Federation

The article explores the construction of semantic models of subject domains to formalize the existing knowledge about the subject domain and generate new knowledge. Methods of representation and processing of partial knowledge about cases of a given domain are investigated. Formal Concept Analysis has been applied to axiomatize classes of partial models that formally represent knowledge about the domain cases. To solve this problem, the theory of classes of fragments of atomic diagrams of algebraic systems and axiomatizable classes of fragments of atomic diagrams are studied. The apparatus of Boolean-valued and fuzzy models is used to obtain new knowledge about the domain cases.

The Metagraph Model for Complex Networks: Definition, Calculus, and Granulation Issues

Valery Tarassov, Yuriy Kaganov, Yuriy Gapanyuk (*gapyu@bmstu.ru*)
Bauman Moscow State Technical University, Moscow, Russia

This article attempts to look at complex graph models (specifically, the metagraph model) through the prism of information granulation. The basic provisions and definitions of information granulation are explored. The formal definitions of the metagraph data model and metagraph agent model are given. The examples of data metagraph and metagraph rule agents are discussed. The concept of an active metagraph, aimed for the data metagraph model and metagraph agent model combination, is proposed. It is shown that the metagraph vertex, the metagraph edge, attributes, the metagraph fragment, metagraph agents, and the active metagraph can be represented as a special kind of a metavertex. The metagraph calculus adopted for the information granulation is proposed. The metagraph calculus from the point of view of atomicity definition is described. The granularity of the metagraph model, including the atomicity property, the holonic organization, refinement and coarsening, partial ordering, similarity relationship, partitions and coverings, measures of granularity, the granularity of metagraph agents, are discussed.

Subjective Expert Evaluations in the Model-Theoretic Representation of Object Domain Knowledge

Gulnara Yakhyaeva, Vera Skokova (*gul_nara@mail.ru*)
Novosibirsk State University, Novosibirsk, Russia

Often, evaluative knowledge about the object area is formulated not only as an objective (statistical) probability but also as a subjective (expert) probability. Expert evaluations may be incomplete or inconsistent with each other. A tool is needed to check the consistency of expertise. The paper proposes a theoretical-modal formalization of subjective and objective interpretations of probability. This allows us to formulate the criteria for the correctness of the evaluative knowledge received from the experts. The article describes an algorithm for checking the correctness of evaluative knowledge, as well as an algorithm for correcting some incorrectness.

Topic 4

Multiagent Systems and Robotics

Q-Mixing Network for Multi-Agent Pathfinding in Partially Observable Grid Environments

Vasilii Davydov, Alexey Skrynnik, Konstantin Yakovlev, Aleksandr Panov
Moscow Institute of Physics and Technology, Moscow, Russia
Moscow Aviation Institute, Moscow, Russia
Artificial Intelligence Research Institute, Federal Research Center
"Computer Science and Control," Russian Academy of Sciences, Moscow, Russia

In this paper, we consider the problem of multi-agent navigation in partially observable grid environments. This problem is challenging for centralized planning approaches as they typically rely on full knowledge of the environment. To this end, we suggest utilizing the reinforcement learning approach when the agents first learn the policies that map observations to actions and then follow these policies to reach their goals. To tackle the challenge associated with learning cooperative behavior, i.e. in many cases agents need to yield to each other to accomplish a mission, we use a mixing Q-network that complements learning individual policies. In the experimental evaluation, we show that such approach leads to plausible results and scales well to a large number of agents.

Subdefinite Computations for Reducing the Search Space in Mobile Robot Localization Task

Anton Moscovsky (*moscovskyad@gmail.com*)
National Research Center "Kurchatov Institute", Moscow, Russia

The paper describes the application of the technology of subdefinite computations for the problem of localizing a mobile robot by landmarks. The application of subdefinite computations together with probabilistic approaches is considered using the example of a histogram filter. It is shown that the use of this technology can significantly increase the computational efficiency by reducing the space in which the search for the position of the robot takes place. The experimental results are obtained by simulation.

Enhancing Exploration Algorithms for Navigation with Visual SLAM

Kirill Muravyev (*kirill.mouraviev@yandex.ru*),
Andrey Bokovoy, Konstantin Yakovlev (*{bokovoy,yakovlev}@isa.ru*)
Artificial Intelligence Research Institute, Federal Research Center for Computer
Science and Control of Russian Academy of Sciences, Moscow, Russia.
Moscow Institute of Physics and Technology, Dolgoprudny, Russia

Exploration is an important step in autonomous navigation of robotic systems. In this paper we introduce a series of enhancements for exploration algorithms in order to use them with vision-based simultaneous localization and mapping (vSLAM) methods. We evaluate developed approaches in photo-realistic simulator in 2 modes: with ground-truth depths and neural network reconstructed depth maps as vSLAM input. We evaluate standard metrics in order to estimate exploration coverage.

Topic 5

Natural Language Processing

Relying on Discourse Trees to Extract Medical Ontologies from Text

Boris Galitsky (boris.galitsky@oracle.com),
Dmitry Ilvovsky, and Elizaveta Goncharova ({egoncharova,dilvovsky}@hse.ru)
Oracle Inc., Redwood Shores, CA, USA
National Research University Higher School of Economics, Moscow, Russia

We explore the role of discourse analysis in ontology construction. While extracting candidate phrases to form ontology entries from text, it is important to pay attention to which discourse units these phrases occur in. It turns out that not all discourse units are equal in terms of their contribution to forming ontology entries. We survey text mining and ontology information extraction techniques in medical do-main and select the ones where advanced linguistic analysis including the discourse level is leveraged the most to produce a robust and efficient ontology. We evaluate the consistency of the resultant ontology and its role in assuring high search relevance using several real-life medical datasets and prove the importance of introducing discourse information into the ontology construction.

TITANIS: A Tool for Intelligent Text Analysis in Social Media

Ivan Smirnov, Maksim Stankevich, Yulia Kuznetsova, Margarita Suvorova,
Daniil Larionov ({ivs, stankevich, suvorova, dslarionov}@isa.ru),
kuzjum@yandex.ru
Elena Nikitina (yelenon@mail.ru), Mikhail Savelov (savelov.me@phystech.edu)
Oleg Grigoriev (oleggpolikvart@yandex.ru)
Federal Research Center Computer Science and Control of RAS, Moscow, Russia,
National Research University Higher School of Economics, Moscow, Russia
Moscow Institute of Physics and Technology, Moscow, Russia

This paper introduces TITANIS, a new social media text analysis tool specifically designed to assess the reaction of social media users to global events from a psycho-emotional point of view. The tool of-fers an expanded set of text parameters and natural language processing methods suitable for working with texts from social media. In addition to the widely used NLP approaches, such as tf-idf and sentiment analysis, TITANIS includes psycholinguistic, semantic, discursive, and other types of analysis that allow detecting more peculiarities in the texts of users with di' erent psycho-emotional states. The paper describes the structure of the tool and provides insight into the methodological background of its functionality. To demonstrate some capabilities of TITANIS, we applied it to Pikabu data to analyze the user reaction to the self-isolation period and the COVID-19 informational background on social media.

Approach to the Automated development of Scientific Subject Domain Ontologies Based on Heterogeneous Ontology Design Patterns

Yury Zagorulko, Elena Sidorova, Irina Akhmadeeva, Alexey Sery
and Galina Zagorulko (*zagor@iis.nsk.su*)

A.P. Ershov Institute of Informatics Systems of Siberian Branch of the Russian
Academy of Sciences, Novosibirsk, Russia

Our practice has shown that for the development of ontologies of scientific subject domains (SSD), the use of ontology design patterns (ODPs) is especially effective. This is due to the fact that the ontology of any SSD contains, as a rule, a large number of typical fragments that are well described by the ODPs. In addition, since these patterns greatly facilitate the development of an SSD ontology, it is possible to involve experts in the modeled SSD not possessing the skills of ontological modeling, which, in turn, speeds up the development of an SSD ontology. In order to obtain an ontology that describes a given SSD fully enough, it is necessary to process a huge number of publications related to this SSD. We can facilitate and accelerate the process of populating an ontology with information from these sources using lexico-syntactic patterns. The paper presents an approach to automating the development of the SSD ontologies based on a set of heterogeneous ontology design patterns. This set includes two kinds of patterns: (a) patterns intended for ontology developers and (b) lexico-syntactic patterns automatically built on the basis of (a), capable of automatically populating the ontology with the information extracted from natural language texts.

Topic 6

Fuzzy Models and Soft Computing

Logistic-Based Design of Fuzzy Interpretable Classifiers

Alexander Dolgiy, Sergey Kovalev, Anna Kolodenkova, Andrey Sukhanov
(a.suhanov@rfniias.ru)

JSC NIIAS, Moscow, Russia

Rostov State Transport University, Rostov-on-Don, Russia

Samara State Technical University, Samara, Russia

The paper develops an idea of fuzzy evidential classifiers based on modification of logistic regression model and Dempster–Shafer methodology. The proposed approach is integrating the additional linguistic variable into the classifier. This variable considers different shades of truth for class membership hypotheses and enriches available information for decision-making. It leads to identification of pre-failure states and detecting anomalies, inconsistency, and incorrectness in the initial data. As a result of the research, linguistic log-regression model is shown, and its components are justified. The inference procedure based on the model is illustrated. In the end, a simple example of implementation is also shown.

The PC-Algorithm of the Algebraic Bayesian Network Secondary Structure Training

Nikita Kharitonov, Maxim Abramov, and Alexander Tulupyev

St. Petersburg State University, St. Petersburg, Russia

St. Petersburg Federal Research Center of the Russian Academy of Sciences,

St. Petersburg, Russia

Algebraic Bayesian networks and Bayesian belief networks are one of the probabilistic graphical models. One of the main tasks which need to be solved during the networks' handling is the model structure training. This paper is dedicated to the automation of this process for algebraic Bayesian networks.

This work relates to the PC-algorithm for algebraic Bayesian network secondary structure training. The algorithm is based on the PC-algorithm for Belief Bayesian networks training. The algorithm pseudo-code and us-age example are described. The provided algorithm helps investigate the full-automated machine learning of algebraic Bayesian networks. Earlier, the structure was provided by experts.

Topic 7

Intelligent Systems

Knowledge-Based Diagnostic System With a Precedent Library

Nikolay Blagosklonov (nblagosklonov@frcsc.ru),
Valeriya Gribova (gribova@iacp.dvo.ru), Boris Kobrinskii (kba_05@mail.ru)
and Elena Shalfeeva (shalfeeva@iacp.dvo.ru)

Federal Research Center “Computer Science and Control” of the Russian
Academy of Sciences, Moscow, Russian Federation
Institute of Automation and Control Processes of the Far Eastern Branch of the
Russian Academy of Sciences, Vladivostok, Russian Federation

The hypothesis of the presumptive diagnosis before laboratory confirmation is especially important in orphan (rare) hereditary diseases. It is possible to solve this problem using computer-based decision support systems based on knowledge. However, in medical practice, there are cases of an atypical clinical picture in patients with fuzzy manifestations of features. In such cases, it is possible to increase the diagnostic accuracy using a precedent approach. The concept of “synthetic precedent” is introduced, which is the result of the transformation of an atypical case into a synthesized description. The paper presents methods for constructing synthetic precedents of two types. The precedents of the first type are created as a result of extension with the fuzzy boundaries for ordinal variables. The precedents of the second type are received by softening the requirement for the number of necessary signs of a patient to match an atypical case from the precedent library. An approach to the creation of a hybrid system, including a traditional knowledge base and a precedent library, is proposed and demonstrated. The use of the hybrid system increases the accuracy of early diagnosis of orphan diseases in childhood.

Semiotic Models in Monitoring and Decision Support Systems

Alexander Kulinich (alexkul@rambler.ru)
V.A. Trapeznikov Institute of Control Sciences, RAS,
Moscow, Russian Federation

The paper considers an approach to the construction of decision support systems based on the use of the subject area semiotic models and methods of natural language processing. The peculiarity of the proposed approach lies in the fact that the unstructured linguistic information of the Internet is used as the knowledge base of the semiotic model. The structuring of this information is carried out based on a subjective qualitative semiotic model of the situation, built by an expert, and the use of natural language processing methods for texts relevant to the subject area, which are obtained from the Internet. The functional structures of decision support systems in the tasks of monitoring the situation state, making decisions, and supporting the modeling results interpretation in conditions

of uncertainty are given. Qualitative assessments of the possibility of implementing decision support systems based on a subjective semiotic model and natural language processing methods are given. Conclusions based on the results of experiments with a software model of the system for the decision support problem are presented, confirming the effectiveness of this approach.

Cognitive Patterns for Semantic Presentation of Natural-Language Descriptions of Well-Formalizable Problems

Sergey Kurbatov (curbatow.serg@yandex.ru),
Igor Fominykh (igborfomin@mail.ru)
and Aleksandr Vorobyev (abvorobyev@bk.ru)
Research Centre of Electronic Computing, Moscow, Russia
National Research University MPEI, Moscow, Russia

The paper suggests and develops a method for creating cognitive patterns for well-formalizable problems described in a natural language. The method takes into account the full set of human cognitive abilities: the understanding of a natural language in the light of a well-formalizable problem, its formalization for an automatic solution, a graphical display of the solution, and its psychological aspects. The method is focused on a holistic approach to applied artificial intelligence systems considered in terms of interdisciplinarity. We have conducted an experimental study of the method capabilities on a set of geometric problems. Our experiment has included the modification of the source text at the level of morphology, syntax, and significant objects of a geometric problem followed by the preparation of a drawing. We have suggested an extension of the experiment to the problems with physical content.

Detecting Anomalous Behavior of Users of Data Centers Based on the Application of Artificial Neural Networks

Igor Saenko, Igor Kotenko ({ibsaen,ivkote} @comsec.spb.ru),
Fadey Skorik (work_bk@bk.ru), and Mazen Al-Barri (mazenb51@gmail.com)
Saint-Petersburg Federal Research Center of the Russian Academy of Sciences
(SPC RAS), Saint-Petersburg, Russia
Saint-Petersburg Signal Academy, Saint-Petersburg, Russia

The paper discusses an approach to detecting anomalies in the behavior of users of data centers, using a specialized analytical unit based on artificial neural networks. It is proposed to use transaction log records of the databases that are part of the data center as data sets for analysis. An experimental evaluation of the proposed approach is made for several types of analytical units, which include several artificial neural networks. Experiments have demonstrated the high efficiency of the proposed approach.

Topic 8

Tools for Designing Intelligent Systems

A Module for Industrial Safety Inspection Planning Based on Self-organization

Alexander Berman, Olga Nikolaychuk, Alexander Pavlov,
Alexander Yurin (idstu.irk.ru)

Matrosov Institute for System Dynamics and Control Theory of the Siberian
Branch of the Russian Academy of Sciences, Irkutsk, Russia

The paper considers the application of the principles of self-organization for the automation of the business process of industrial safety inspection (ISI). Self-organization refers to the process of ordering elements (components) in an information system due to internal factors, without specific external influence. Self-organization is parametric and structural in nature. The architecture of the system as a whole and the main aspects of the main module, namely, scheduler, which provides self-organization of an open multicomponent information system of the ISI, are described. An ontological model of the object of expertise, the task, the method, and the operation that implements it, as the basis of self-organization, is proposed. The model and algorithms for implementing the scheduler are detailed. In particular, the algorithms of the main operations are described: the formation of the task description, the methodology for solving the problem, and the "intelligent" task execution. It also lists the principles of formation and describes the main local rules of the knowledge base of the scheduler, which are responsible for describing the methodology, the decision process, the coordination of expert opinions, and self-learning. A conceptual description of the implementation of the scheduler based on the components of the software platform is given. Some results of application of the considered approach for the problem of technical diagnostics of ISI are given: fragments of the generated software, forms of the user interface demonstrating the description of the object of expertise, local rules, methods and results of self-organization.

A Study of the Feasibility of Creating of a Real-Time Neural Network Infrared Ground Objects Recognition System

Andrey Maltsev, Anatoly Nekleenov, Dmitriy Otkupman,
and Victoria Ostashenkova (odvk@ya.ru)

JSC Research and Production Enterprise "Impulse", Moscow,
Russian Federation

The problem of implementation of a real-time neural network thermal imaging recognition system, built on widely available components, and allowing placement on a small-sized carrier, is considered. The main criteria for choosing hardware and software parts were the data processing speed and high accuracy of the classification of the detected ground objects.

The Implementation of the Ontological Approach to Control of the Processes of Designing Integrated Expert Systems Based on the Problem-Oriented Methodology

Galina Rybina (gvrybina@yandex.ru), Aleksandr Slinkov (sanch971@mail.ru)
National Research Nuclear University MEPhI, Moscow, Russia

This work focuses on the experimental studies of a new technology for the development of integrated expert systems (IES) based on the further development of the problem-oriented methodology and intelligent software environment of the AT-TECHNOLOGY workbench through the integration of an ontological approach to software design for applied IES and methods of intelligent planning and management of IES development processes with different architectural typology.

The description of prototyping methods for applied IES is based on the use of the basic components of the model of an intelligent software environment with an emphasis on expanding one of the components, the technological knowledge base, through the development of an ontology of typical IES architectures and the implementation of interaction with the intelligent planner.