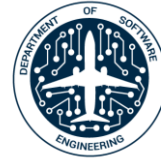


**Ministry of Education and Science of Ukraine
National Aviation University
Software Engineering Department**



**International Conference on
Software Engineering**








April 10 – 12, 2022



Proceedings

Kyiv 2022

Conference organizers and partners

	<p>National Aviation University (Kyiv, Ukraine) www.nau.edu.ua</p>
	<p>Software engineering department www.sed.nau.edu.ua</p>
	<p>Artificial Intelligence and Knowledge Engineering Research Lab (Cairo, Egypt): http://aiasulab.000webhostapp.com/Ambassadors.html</p>
	<p>The Association of Developers and Users of Intelligent Systems http://www.aduis.com.ua/</p>
	<p>Faculty of Computer and Information Sciences, Ain Shams University (Cairo, Egypt): http://cis.asu.edu.eg/</p>
	<p>The Institute for Information Theories and Applications FOI ITHEA® (Bulgariya) http://www.ithea.org/index.html</p>
<p>ITHEA® International Scientific Society</p>	<p>International Scientific Society ITHEA® ISS http://www.ithea.org/iss/iss.html</p>
	<p>V. M. Glushkov Institute of Cybernetics of the NAS of Ukraine (Kyiv, Ukraine) http://incyb.kiev.ua/?lang=en</p>

NATIONAL AVIATION UNIVERSITY

(www.nau.edu.ua)

For almost 80 years of its history more than 200,000 highly skilled professionals have been trained in this higher aviation educational institution. Among them there are well-known scientists, heads of aviation companies, enterprises, organizations and institutions that provide aircraft flights, their maintenance and repair, transportation of passengers and cargo.

Strong scientific and pedagogical schools enable to train not only engineers but also economists, lawyers, environmentalists, translators, psychologists, sociologists and others.

SOFTWARE ENGINEERING DEPARTMENT

www.sed.nau.edu.ua

Software engineering department studies more than 900 students. Students obtain fundamental knowledge and practical skills in software development area as well as theoretical basic in software development life-cycle management.

AIN SHAMS UNIVERSITY AI AND KNOWLEDGE ENGINEERING LAB

<http://aiasulab.000webhostapp.com/>

The Ain Shams University Artificial Intelligence and Knowledge Engineering Lab was founded in 2005 and belongs to the Computer Science department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt.

The Lab is currently composed of 3 full Profs., 3 Assistant Prof., 15 Drs (PhD holders) 5 of them from Arabic countries, 10 PhD students and 5 Msc students. Laboratory is working in the following disciplines:

- Intelligent Information Processing, Data Mining.
- Knowledge Engineering and Expert Systems.
- Image Processing and Pattern Recognition.
- Data Science and Big Data Analytics
- Intelligent Education and Smart Learning.
- Medical Informatics, Bio-informatics and e-Health.
- Computing Intelligence and Machine learning
- Internet of Things (IoT),
- Technologies for HealthCare;
- Space Science and Cosmic Rays Big Data Analytics.

In addition, the lab includes 17 scientific international ambassadors from Germany, Canada, Italy, Georgy, Ghana, Cyprus, Ukraine, Romania, Greece, Japan, Mexico, and Belgium

THE ASSOCIATION OF DEVELOPERS AND USERS OF INTELLIGENT SYSTEMS

<http://www.aduis.com.ua/>

Association of Developers and Users of Intelligent Systems (ADUIS) consists of about one hundred members including ten collective members. The Association was founded in Ukraine in 1992. The main aim of **ADUIS** is to contribute to the development and application of the artificial intelligence methods and techniques.

Association has long-term experience in collaboration with teams, working in different fields of research and development. Methods and programs created in Association were used for revealing regularities, which characterize chemical compounds and materials with desired properties.

**FACULTY OF COMPUTER AND INFORMATION SCIENCES, AIN SHAMS
UNIVERSITY (Cairo, Egypt):**

<http://www.shams.edu.eg/598/news#>

Faculty of computer and information sciences was founded in 1996 and consists of 5 departments, namely: Computer Science, Information Systems, Computers Systems, Scientific Computing and Basic Sciences.

Faculty of computer and information sciences announces it's distinct programs by credit hours system:

- Software Engineering;
- Artificial Intelligence;
- Cyber Security;
- Digital multimedia.

Faculty of computer and information sciences also provides Bioinformatics program which accepts Thanawya Amma science section

**INSTITUTE FOR INFORMATION THEORIES AND APPLICATIONS FOI ITHEA®
(BULGARIA)**

<http://www.ithea.org/>

The Institute for Information Theories and Applications FOI ITHEA® is an international nongovernmental organization functioning since 2002 year. ITHEA® is aimed to support international scientific research through international scientific projects, workshops, conferences, journals, book series, etc.

INTERNATIONAL SCIENTIFIC SOCIETY**ITHEA® ISS (INTERNATIONAL SCIENTIFIC SOCIETY)**

<http://www.ithea.org/iss/iss.html>

ITHEA® International Scientific Society (ITHEA® ISS) is aimed to support growing collaboration between scientists from all over the world.

Till now, the **ITHEA® International Scientific Society** has been joined by more than **4000** members from **53** countries from all over the world:

ITHEA® Publishing House (**ITHEA® PH**) is the official publisher of the works of the ITHEA®ISS. The scope of the issues of the ITHEA®ISS covers the area of Informatics and Computer Science. ITHEA® PH welcomes scientific papers and books connected with any information theory or its application.

ITHEA®ISS has four International Journals, established as independent scientific printed and electronic media, published by ITHEA® Publishing House:

- International Journal “Information Theories and Applications” (IJ ITA), since 1993;
- International Journal “Information Technologies and Knowledge” (IJ ITK), since 2007;
- International Journal “Information Models and Analyses” (IJ IMA), since 2012;
- International Journal “Information Content and Processing” (IJ ICP), since 2014.

Every year, ITHEA®ISS organizes many scientific events called “**ITHEA® ISS Joint International Events of Informatics**” (ITA). ITHEA®ISS has organized 135 International Conferences and workshops and has published 45 scientific books: monographs, thematic collections and proceedings.

All publications of ITHEA®ISS are accessible freely via ITHEA® main site www.ithea.org as well as via ITHEA® Digital Library (IDR) <http://idr.ithea.org>.

The great success of ITHEA® International Journals, Books, and Conferences belongs to the whole of the **ITHEA® International Scientific Society**. Membership in ITHEA® ISS is free.

V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine

<http://incyb.kiev.ua/?lang=en>

M.Glushkov Institute of Cybernetics of the NAS of Ukraine is a Scientific Centre widely known in Ukraine and beyond it engaged in the solution of fundamental and applied issues of informatics and computer science as well as introduction of their methods and means to various spheres of human activities.

Today the main areas of the research of the Institute are:

Development of the general theory and methods of system analysis, mathematical modelling, optimization and artificial intelligence;

Development of the general theory of management, methods and means of arrangement of intelligent control systems of different levels and purposes

Creation of the general theory of computing machines and development of perspective computing equipment, artificial intelligence and informatics;

Creation of perspective systems of mathematical support of general and applied purposes;

Development of new information technologies and intelligent systems; and

Solving of fundamental and applied problems of informatization of the society.

Scientific schools of computer mathematics and discrete optimization, mathematical theory of computing systems and artificial intelligence, system analysis and stochastic programming, mathematical theory of reliability and programming theory were forming for fifty years at the V.M.Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine and they gained world's recognition.

Conference topics

Theoretical Basics of Software Engineering

AGILE and Model-Driven Approaches to Software Development.
Approaches to Software Development Life Cycle Processes Improvement;
Business Process Management and Engineering
Empirical Software Engineering;
Formal Foundations of Software Engineering;
Frameworks and Middleware;
Model-Driven Engineering;
Software Development Technologies;
Software Engineering Standards;
Software Patterns and Refactoring;

Applied Aspects of Software Engineering

Approaches and Models to Estimate Expenses of Future Software Project;
Component-Based Software Engineering;
Requirements Engineering;
Service-Oriented Software Engineering;
Software Designing;
Software Quality;
Software Maintains and Evolution;
Software Testing.

Software Engineering Application Domains

Agent and Multi-Agent Systems;
Biotechnologies and Smart Health Technologies;
Cloud Computing;
Data Bases and Knowledge Bases;
Image Processing and Computer Vision;

Internet;

Security;

Software Development for Mobile Operation Systems.

Real-Time Systems;

Network and Data Communications;

User-Centered Software Engineering.

Steering committee

- ✓ Abdel-Badeeh M. Salem (Ain Shams University and Artificial Intelligence and Knowledge Engineering Research Lab, Egypt);
- ✓ Kateryna Nesterenko (National Aviation University, Ukraine);
- ✓ Krassimir Markov (Institute of Information Theories and Applications, Bulgaria);
- ✓ Olena Chebanyuk (National Aviation University and V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine, Ukraine);
- ✓ Serhii Zybin (National Aviation University, Ukraine);
- ✓ Vitalii Velychko (V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine and Association of Developers and Users of Intellectualized Systems).

Program committee

- ✓ Aykut Karakaya (Bulent Ecevit University, Turkey);
- ✓ Carlos Cares (Software Engineering Studies Center (CEIS). University of La Frontera. Temuco, Chile);
- ✓ Gefei Zhang (Hochschule für Technik und Wirtschaft Berlin, Germany);
- ✓ Ewa Ziemba (University of Economics in Katowice, Poland);
- ✓ Evellin Cardoso (Federal University of Goias, Brazil);
- ✓ Hakan Kutucu (Karabuk University, Turkey);
- ✓ Khaja Mannanuddin (Sumathi Reddy Institute of Technology, India);
- ✓ Krasimira Ivanova (University of Telecommunications and Post, Bulgaria);
- ✓ Michael Gr. Voskoglou (Graduate Technological Educational Institute of Western Greece School of Technological Applications, Greece);
- ✓ Rafael Lahoz Beltrá (Universidad Complutense de Madrid, Spain);
- ✓ Sedat Akleylek (Ondokuz Mayıs University, Turkey);
- ✓ Taras Panchenko (Taras Shevchenko National University of Kyiv, Ukraine);
- ✓ Tea Munjishhvili (I.Javakhishvili Tbilisi State University, Georgia);

-
-
- ✓ Vitalii Prykhodniuk (Central Scientific Research Institute of Armament and Military Equipment of the Armed Forces of Ukraine);
 - ✓ Yuh Wen Chen (Da Yeh University, Taiwan).

Organization committee

- ✓ Mykyta Krainii (National Aviation University, Ukraine);
- ✓ Tetiana Konrad (National Aviation University, Ukraine);
- ✓ Yana Bielozorova (National Aviation University, Ukraine);
- ✓ Yuliia Bezkorovaina (National Aviation University, Ukraine).

Table of contents

<i>SECTION: THEORETICAL BASICS OF SOFTWARE ENGINEERING</i>	14
<i>A MODEL-DRIVEN APPROACH FOR CYBER-PHYSICAL SYSTEMS: A METHOD ENGINEERING PERSPECTIVE</i>	14
Carlos Cares, Olena Chebanyuk, Claudio Navarro	14
<i>USE OF SEMANTIC TECHNOLOGIES IN ADVISORY SOFTWARE DEVELOPMENT</i>	20
Julia Rogushina, Anatoly Gladun, Serhii Pryima, Oksana Strokan	20
<i>A NEW APPROACH TO AUTOMATIC ONTOLOGY CREATION FROM UNTAGGED TEXT ON NATURAL LANGUAGE OF INFLECTIVE TYPE</i>	37
Anna Litvin, Vitalii Velychko, Vladislav Kaverinsky	37
<i>CODE STATIC ANALYSIS BY REVERSE ENGINEERING TOOLS</i>	46
Mykyta Krainiy	46
<i>SECTION: APPLIED ASPECTS OF SOFTWARE ENGINEERING</i>	50
<i>DESIGN PRINCIPLES FOR SOFTWARE USING TECHNIQUES OF BIOLOGIC INFORMATION PROCESSING</i>	50
Karl Javorszky	50
<i>SECTION: SOFTWARE ENGINEERING APPLICATION DOMAINS</i>	62
<i>APPLICATION OF SOFT SETS TO ASSESSMENT OF PROBLEM SOLVING SKILLS</i>	62

Michael Gr. Voskoglou	62
<i>ABOUT THE ISSUE OF JOINT PROCESSING OF IMAGES OF THE EARTH'S SURFACE, OBTAINED FROM DIFFERENT TYPES OF MEDIA.....</i>	<i>69</i>
Valeriy Pisarenko, Julia Pisarenko, Ivan Varava, Olga Gulchak.....	69
<i>PRINCIPLES FOR SELECTING CHARACTERISTIC PEAKS OF SPEECH SIGNAL FOR SPEAKER IDENTIFICATION.....</i>	<i>75</i>
Serhii Zybin, Yana Bielozorova	75
<i>ABOUT THE ORGANIZATION OF REGIONAL SITUATIONAL CENTERS OF THE INTELLECTUAL SYSTEM "CONTROL_TEA" WITH THE USE OF UAVS.....</i>	<i>81</i>
Julia Pisarenko, Katerina Melkumian, Alexander Kryachek, Olga Gavrilyuk, Alexander Koval.....	81
<i>PHOTO SHARPNESS EVALUATION SYSTEM DESIGN.....</i>	<i>86</i>
Oleksandr Dolhanenko, Mariya Shirokopetleva, Natalya Lesna	86
<i>REPRODUCING SOUND WAVES WITH AN ALTERNATIVE DIGITAL-TO-ANALOGUE CONVERTER AND EMITTERS.....</i>	<i>92</i>
Stanislav Danylenko, Oleksander Vechur, Mariya Shirokopetleva	92
<i>TRIP PLANNING SOFTWARE SYSTEM DESIGN</i>	<i>97</i>
Mariia Kryvoruchko, Mariya Shirokopetleva, Oleksander Vechur	97

SECTION: THEORETICAL BASICS OF SOFTWARE ENGINEERING

A MODEL-DRIVEN APPROACH FOR CYBER-PHYSICAL SYSTEMS: A METHOD ENGINEERING PERSPECTIVE

Carlos Cares, Olena Chebanyuk, Claudio Navarro

Abstract: *Two of the most profitable approaches in Software Engineering are reuse and code generation. On this line, Model-Driven Development (MDD) not only includes code generation that is based on transformation rules as well as reuse of parameterized packages or components. In addition, transformation rules should satisfy the next software designing principles: separation of concerns, high cohesion and low coupling. MDD allows managing different views due to multidimensional perspectives and models of the system to be. On the other hand, Cyber-Physical Systems (CPS) are software intensive and its social embedded, productive environment, or critical functionality requires that different functional and non-functional requirements be represented, managed, traceable, verified and validated. Therefore, Model-Driven Development offers a general approach for dealing with these required process features in the context of cyber-physical systems. In software engineering, the task of adaptation and specialization of the software development tasks is known as method engineering. In this paper we offer a set of method engineering representations of a model-driven approach for dealing with the software development for cyber-physical systems. It is an initial formalization work for comparing the requirements of a Cyber-physical engineering process and the current offer of Model-Driven Development, an agenda of future work and challenges to complete the map of results.*

Keywords: *Model-Driven Development, Method Engineering, Cyber-physical Engineering.*

ITHEA Keywords: *D.2.9 Software Engineering Management, D.3.1 Formal definitions and theory, I.2.9 Robotics.*

Conference topic: *Theoretical Basics of Software Engineering*

Introduction

Model-Driven Development approach is a way of software development when software models become main «citizens» of the development process. In order to use models

some widespread graphical notation is used (often UML). Such a notation allows expressing the semantics of the software models in a convenient way for human understanding.

Different software development life cycles require different levels of details in its models while information about software is represented. The aim of these representations is to describe software structure and behavior from different perspectives or views. Additionally, the representations can have engineering requirements for improving maintainability, traceability or even reverse engineering processes.

The context of the current system “as is” and the expected context of the system “to be” are usually required models depending on the particular chosen software process. Thus business processes, business rules, user scenarios, and particular constraints are also targets of a detailed design to represent in some kind of models.

However, an MDD-like software process requires not only the models but also the transformations rules among them. Detailed foundations of software models’ transformation approach, as well as classification of different transformation approaches are summarized in [Chebanyuk and Markov, 2016].

On the other side, cyber-physical systems (CPS) are being proposed for monitoring and controlling complex processes in a real time set. Thus, they require sensors and actuators as well the distributed software for coordinating actions for controlling their parts and components [Lee, 2006]. Of course, the CPS engineering has imposed several challenges and most of them are still without a clear approach for dealing with them. From the software engineering foundations are separation of concerns, integration of concerns, among others. From the information defense are human safety, information security, environment representation [Cares et al., 2019].

Therefore, on one hand we have a promising methodological approach as MDD and, on the other hand, we have the relevant differences, particularities and challenges that implies the engineering for cyber-physical systems. For this scenario, i.e. for the problem of adapting and generating a methodological approach to a particular situation is known

as situational method engineering [Ralyté and Rolland, 2001], when the focus is only on the formalization process we named method engineering [Henderson-Sellers and Ralyté, 2010]. The theoretical assumption is that a method, in particular a software development method, is composed by method's fragments or chunks, and, arranging these chunks in different ways we obtain different methods. From a paradigmatic perspective, method engineering has used class diagrams for the static perspective of method engineering and state transition diagrams for representing the dynamic part of the method engineering, this is to recognize chunks or assembly chunks looking for a proper method. In this paper, we offer a method engineering perspective for generating the basics of an MDD approach for developing cyber-physical systems. A method engineering perspective is the base for generating particular software development methodologies, having a common set of assets, for developing cyber-physical systems. In this context, different types of cyber-physical systems would arrive to different methodologies and, also, similar types of cyber-physical systems but having different goals, would also arrive to different development methodologies.

Method Engineering for a MDD approach for CPS

da Silva [da Silva, 2015] has developed a recognized set of static models for representing the components of a model-driven approach for software development. In Figure 1, we have summarized part of his conceptual model extending it to three cases of transformations in an MDD approach, and, additionally, we have related some of the parent classes of da Silva's representation to basic concepts in method engineering proposed in [Henderson-Sellers and Ralyté, 2010].

Note that as in any other class diagram, the inheritance has conceptual implications, for example, the general conceptualization of chunk implies that there are aggregated and atomic chunks, besides, a chunk has a mandatory description and explicit conditions for reuse. Therefore, the inheritance from *Method Chunk* to *MDDTransformation* means that there are aggregate and atomic MDD transformations. Additionally, each MDD transformation has also a mandatory description.

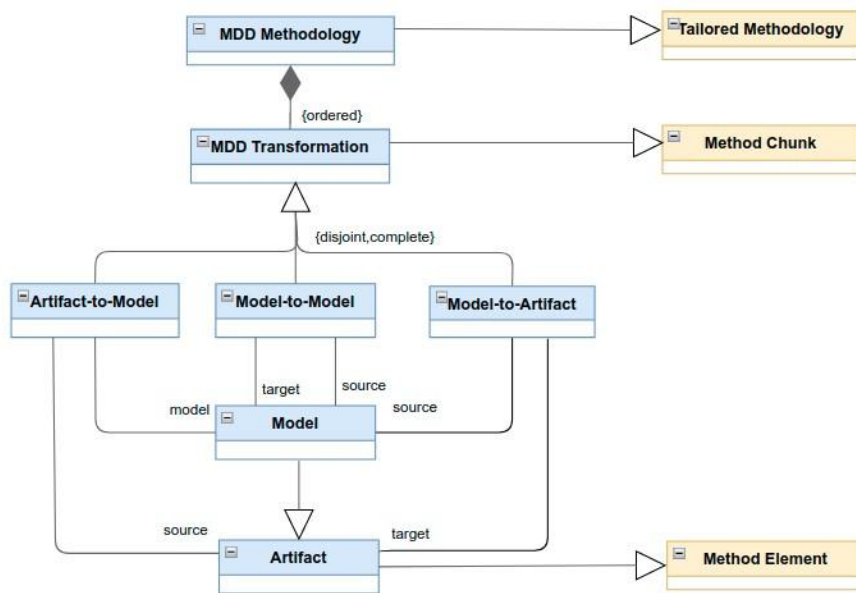


Figure 1: MDD Concepts in the context of Method Engineering

A second recurrent diagram in Method Engineering are state machines for representing the construction and systematic extension of a methodology or of a methodology component. In Figure 2 we have illustrated a CPS example of managing physical components as part of the methodology.

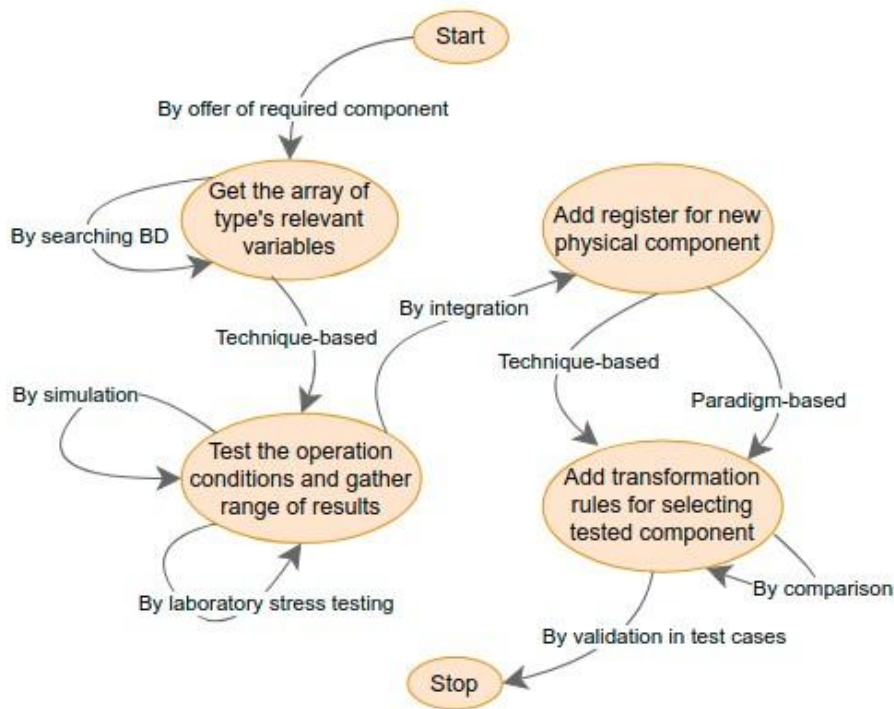


Figure 2: Processing new transformation rules into a CPS-MDD Methodology

In order to put in context this diagram, we can assume the case of Autonomous Underwater Vehicles (AUV) and their propulsion system. A propulsion system may have a single component or several components, but the atomic case is a single component: for example a propeller. Each atomic propulsion component has a set of variables that define it, for example, power consumption, range of propulsion forces, discrete or continuous, energy dissipation, depending on the technology, may be motor-shaft inertia among others. This set of variables, that make a different decision about selecting the component or not, as part of a solution, is "the array of type's relevant variables" referred to in the states after the start in Figure 2. After that the component is tested, its results are registered and transformation rules are modified so that different conditions can select it as component if it corresponds.

Conclusion

The lack of a systematic and disciplined approach for addressing the development of cyber-physical systems has been described and the role of the Method engineering and situational method engineering for generating proper engineering methodologies and techniques have been remarked. By means of two method engineering models we have shown how general MDD concepts match with a method engineering approach and also how to use the classic state machines of situational method engineering for improving MDD transformation rules at the level of mechatronic devices of a cyber-physical system. Of course, the work of generating a particular MDD-like approach requires additional and detailed theoretical-practical work.

Acknowledgements

We want to acknowledge the National Agency of Research and Innovation, ANID, in Chile by funding the project TRABOTS: traceability in the design of cyber-physical systems, code FOVI20010006.

Bibliography

[Cares et al., 2019] Cares C., Sepúlveda S., Navarro C. Agent-Oriented Engineering for Cyber-Physical Systems. In: Rocha Á., Ferrás C., Paredes M. (eds) Information

- Technology and Systems. ICITS 2019. Advances in Intelligent Systems and Computing, vol-918. pp.93-102, Springer, 2019.
- [Chebanyuk and Markov, 2016] Elena Chebanyuk, Krassimir Markov. Model of problem domain “Model-driven architecture formal methods and approaches”. International Journal “Information Content and Processing”, vol-2, n3, ITHEA 2016, p.202-222.
- [da Silva, 2015] Da Silva, A.R., Model-driven engineering: A survey supported by the unified conceptual model. Computer Languages, Systems & Structures, 43, pp.139-155, 2015.
- [Lee, 2006] E.A.Lee. Cyber-physical systems - are computing foundations adequate? position paper in the NSF Workshop on cyber-physical systems: research motivation, techniques and roadmap, vol-2, pp.1-9, Citeseer, 2006.
- [Ralyté and Rolland, 2001] Ralyté, J. and Rolland, C., An assembly process model for method engineering. In Int. Conf. on Advanced Information Systems Engineering (pp. 267-283). Springer, Berlin, Heidelberg, 2001.
- [Henderson-Sellers and Ralyté, 2010] Henderson-Sellers, B. and Ralyté, J., Situational method engineering: state-of-the-art review. Journal of Universal Computer Science, vol-16, n3, pp.424-478, 2010.

Authors` Information



Carlos Cares – *Software Engineering Studies Center (CEIS), University of La Frontera (UFRO). Av. Francisco Salazar 01145, Temuco, Chile; e-mail: carlos.cares@ceisufro.cl*

Major Fields of Scientific Research: *Agent-oriented Software Engineering, Software Engineering for AI-based Systems*



Olena Chebanyuk – *Professor of Software Engineering Department, National Aviation University, Kyiv, Ukraine. e-mail: chebanyuk.elena@ithea.org (chebanyuk.olena@gmail.com)*

Major Fields of Scientific Research: *Model-Driven Architecture, Model-Driven Development, Software architecture, Mobile development, Software development.*



Claudio Navarro – *University of La Frontera (UFRO). Casilla 54-D, Av.Francisco Salazar 01145, Temuco, Chile; e-mail: claudio.navarro@ufrontera.cl*

Major Fields of Scientific Research: *Software Engineering, Semantic Web, Cyber-physical systems.*

USE OF SEMANTIC TECHNOLOGIES IN ADVISORY SOFTWARE DEVELOPMENT

Julia Rogushina, Anatoly Gladun, Serhii Pryima, Oksana Strokan

Abstract: *We analyze specifics of modern semantic technologies used for development of the Web-oriented advisory software. Ontological analysis provides representation of distributed background knowledge about users, their professions, competencies, lifelong learning outcomes, etc., and standards of the Semantic Web project provide the technological foundation for creation of intelligent advisory applications. Validation of learning outcomes of informal and non-formal learning is an important component processed by advisory software that needs in algorithms for semantic matching of user profiles with classifications of professions and qualifications. On base of this research we develop formal models for representation, acquisition and processing of advisory knowledge that can be transformed into intelligent Web-oriented software. Methods for comparison of different advisory information objects based on atomic competences processing are developed. Elements of artificial intelligence and machine learning are used for construction of information objects that are processed by advisory systems. Advantages of proposed approach are demonstrated on example of applied information system AdvisOnt developed to combine the market of educational services with the labor market that can use validation of the informal and non-formal learning outcomes.*

ITHEA Keywords: *1.2.4 Knowledge Representation Formalisms and Methods, 1.2.1 Applications and Expert Systems, H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)*

Conference and topic: *SoftEngine 2022 International Conference on Software Engineering, National Aviation University, Kyiv, Ukraine 10.04-12.04.2022. Conference topics: Software Engineering Application Domains - Data Bases and Knowledge Bases.*

Introduction

Now the Semantic Web and associated semantic technologies become standards of the mainstream software development industry dealt with a huge number of the Web-oriented intelligent information systems (IISs). Such IISs migrate from the processing and storage of large amounts of data to processing and storage of more compact knowledge with a much more complex structure. They apply semantization for different steps of information processing and knowledge management [Warren, 2006] and try to take into

account meaning of data and relations with domain knowledge. Ontology-based semantic technologies (STs) allow to build powerful applied intelligent applications aimed on analysis and modeling of complex objects and processes of different nature. Semantic applications (SAs) are considered in this work as a subset of IISs with some specific features dealt with knowledge processing. The most important of these features is the requirement for use of various external KBs without software changes.

One of the most well-known ST projects is the Semantic Web (SW) [Decker et al, 2000] aimed to transform the Web into global knowledge base (KB). SW provides a large number of standards and tools for knowledge representation and processing by intelligent applications [Sabou, 2008]. The main components of the Semantic Web are ontologies for knowledge representation [Obrst et al, 2007], Web services [Studer, Grimm and Abecker, 2007] for knowledge processing and software agents for representation of individual needs of users [Hendler, 2001].

SW proposes open standards that allow formalizing the semantics of information resources (IRs) and software tools for their search and processing:

- metadata description language RDF [Lassila and Swick, 1998];
- ontology representation language OWL [Bechhofer et al, 2004];
- query language SPARQL [Pérez, Arenas and Gutierrez, 2009] for RDF and OWL.

Use of STs depends on specifics of solved tasks and used IRs. In these work we analyze the specific features of advisory SAs (they are defined by combination of advisory, competence and agricultural knowledge from national and international standards, KBs and IRs) and demonstrate advantages of STs on example of the Web-oriented advisory system AdvisOnt [Pryima et al, 2020].

For this purpose, we consider tasks dealt with combining of the market of educational services with the labor market that can solve problems caused by informal and non-formal learning. Therefore, we take into account specifics of educational domain, its subjects and objects and their relations into the labor market. We consider open sources of information about them and methods of their acquisitions and matching.

Problem definition

AdvisOnt is an intelligent software that provides wide number of advisory functions. This system is aimed for automated semantic matching [Giunchiglia, 2009] of qualifications and competences of various information objects (IOs) – humans, organizations, learning courses, requirements of employer etc. Such matching needs in use of external knowledge bases represented by ontological approach and their integration with internal ontological representation of advisory process. Problem can be considered as a special case of ontology matching [Shvaiko and Euzenat, 2008]. Therefore, design of AdvisOnt provides processing of ontology formats (OWL and RDF) and their transformation into more usable representations.

Ontology of competencies and methods for processing of atomic competences become the ground for semantic integration of such objects described by different terms from various qualification systems. AdvisOnt needs in external knowledge sources and in non-trivial methods of their processing. Therefore we have to use the Semantic Web standards that supports these possibilities that can't be achieved in other ways.

Related work

Literature review proposed by [Mommsen-Ghosh, 2004] defines investment advisory services and their role in solving of problem of information transmission into customer-specific process. In general, the content of advisory services is any information relevant to the decision-making of the advisor. Education and the contact frequency with the advisor affect the level of his individual knowledge that influences on advisory results. Advisory tools vary from simple presentations to complex software for mathematical simulations and statistical analyses. In the concept of service modules, they can be represented by information resources, models and methods. Advisory tools can be used either as a part of a service module where the advisor interprets the results of the tool, thereby enhancing the service brought to the customer; or they can be sold as an isolated module where the tool is put at the customers' disposal via the Web. Service modules serve as the conceptualization to understand the involved domain knowledge.

Knowledge used by advisory services includes organizational routines, processes, practices and norms, framed experiences, values, contextual information and expert insight that provides a framework for evaluating and incorporating new experiences and information. Modern Web-oriented systems use ontological approach for interoperable and personified representation of advisory knowledge [Erriquez and Grasso, 2008]. Important part of this background knowledge [Sabou, 2008] represents learning outcomes of various forms of learning.

Official recognition of non-formal and informal learning outcomes obtained outside formal learning systems of partial qualifications has to take into account all outcomes obtained by persons in process of lifelong learning. Lifelong learning is a key factor in personal and professional development of human defined by Global Standard for Lifelong Learning and Worker Engagement to Support Advanced Manufacturing (<http://www3.weforum.org/docs>). Validation of the results achieved in the process of non-formal and informal learning (knowledge, skills, competencies, etc.) with use of open educational IRs is necessary for access to the labor market and lifelong learning [Colardyn and Bjornavold, 2004]. Information about these outcomes can be proposed by person or be acquired from various external IRs.

Learning outcomes of some person consist of his/her knowledge, skills, abilities and competencies. These outcomes are statements that describe personified knowledge or skills should acquire by the end of a particular assignment, class, course or program, and define for students why that knowledge and those skills can be useful to them.

Recognition of such outcomes that is achieved through non-formal and informal (spontaneous) learning with the help of various open educational resources is necessary for semantic matching of resumes with vacancies of labor market and propositions of learning organizations. Validation allows the recognition of learning outcomes obtained outside the institutions of formal learning (in non-formal and informal education) and is necessary for access to the labor market and lifelong learning [Pryima, Rogushina and Strokan, 2018].

The tools used in this process have to take into account changes in the open world and be dynamic, and they need in semantic retrieval components based on ontological models of user and domain [Pryima et al, 2020].

Semantic technologies

STs aimed at knowledge-level information processing can solve validation problem because they are able to formalize, analyze and process the content of IRs that describe results of learning. Semantics allows to define explicitly meanings and relations between domain concepts represented by data (words, phrases, symbols, etc.) that depend on context. For the same piece of information, semantics can be defined differently depending on ontology used to formalize the user's view of the world. Ontologies can be used as a formal, explicit specification of conceptualization of terms at a certain level of details [Guarino, 1998].

Semantic processing of information includes:

- methods and means of integration and unified representation of heterogeneous distributed knowledge and data;
- retrieval and processing the Web resources as a universal knowledge source of about meaning of information objects.

Semantic technologies in general can be described through a combination of three main components [Gorodetsky and Tushkanova, 2018]: ontologies; semantic resources; models of semantics of natural language (NL) entities. This approach is concentrated on NL processing. In our study we take into account more wide classes of IOs with various structure defined by appropriate ontologies. For example, advisory systems analyze people, organizations, vacancies, learning courses, etc. that can contain NL definitions, multimedia elements and structured data.

Therefore from the point of view of creating semantic applications, these three basic components form a hierarchy where: 1. ontologies and other knowledge bases (KBs) are the upper level of abstraction of the knowledge structure; 2. IO models represent the intermediary level that allows to distinguish typical IOs and their properties and

characteristics; 3. semantic IRs are the lower level that provides information about individuals of classes.

Semantic IRs can have links between content elements (IOs of various types and structure) and with elements of IO models (for example, links with other Wiki pages or with data). Meaning of links is provided by means specific for IR representation and markup. For example, Semantic MediaWiki uses semantic properties.

Structure of IO models can contain relations with other IOs (for example, some IO of category "Person" has semantic link with IO of category "Organizations" by relation "Place of work", and such link can be used into page content only if is present into the IO model). IO models can be formalized by various representations such as templates and forms. Domain ontologies contain classes and individuals of concepts and formalize their properties and characteristics.

Semantic IRs explicitly define relations between content elements and formalized knowledge representations. Usually they use some common standards (such as MPEG21 for multimedia) or problem-specific domain ontologies.

Ontologies in semantic advisory system

Ontological analysis is widely used now for formal modeling of various domains [Gruber, 1991]. Ontology provides a formal explicit description of domain concepts (classes and individuals), their properties, attributes and relations. Moreover, ontology can contain some domain-specific restrictions on use of all these elements and their combinations. Modern intelligent applications use ontologies as interoperable KBs [Staab and Studer, 2013].

A lot of oriented on the Semantic Web software systems use ontologies as a base of domain knowledge for semantic markup of various documents (NL texts, Wiki resources, other semi-structures and structured texts, multimedia context etc.). Popular representation of information in Wiki resources can be semantized in this way. Advantages of the use of ontologies for semantic representation for learning domain and

competencies [Miranda et al, 2016], [Lundqvist, Baker and Williams, 2017] are analyzed by many researchers.

Semantic Wiki resources are Wiki IRs with underlying model of the knowledge described in its pages. Widely used semantic markup of the Web resources is realized by various intelligent extensions of the Wiki resources (such as Semantic MediaWiki, OntoWiki, IkeWiki, SemanticXWiki, and KawaWiki). There is a number of Wiki software that provides semantic functionality. Some of them are standalone Wiki applications, and others are realized as extensions or plugging to standard Wiki software. Semantic Wiki-based IRs differ in their degree of formalization. Some of them support integration with external ontologies (RDF and OWL) and can generate local ontologies for group of the Wiki pages. For example, Semantic Media Wiki provides to users such tools of semantic structuring as categories and semantic properties.

Categories help to link Wiki pages with more general terms and group them, and semantic properties allow to define various semantic features and their values of concept linked with some page. Categories and semantic properties of the Wiki pages can be used as classes and object properties of domain ontology, and names of Wiki-pages – as individuals of ontology. Such domain ontology can be built automatically by special functions of Semantic Media Wiki or by special algorithms according to personal needs of users. Unfortunately, there is no logical or semantic restrictions on ontology building in Semantic Media Wiki. Therefore, ontologies provided such possibilities remain the important content of semantic technologies as a source of domain knowledge

In order to support the process of validation of learning outcomes in both formal and non-formal and informal learning, the European Commission has developed a free internet portal for multilingual classifier ESCO (European Skills, Competencies, Qualifications and Occupations – <https://ec.europa.eu/esco/portal/home>) that joins the labor markets of the EU member states and allows jobseekers and employers to communicate more effectively with definitions of skills, training and work in all European languages. The

main elements of ESCO are professions, skills and qualifications related to the labor, education and training market in the EU (see Figure 1).



Figure1: ESCO taxonomy of professions and skills

ESCO allows users to determine what knowledge and skills are usually required to work in a particular profession. Each ESCO concept is associated with at least one term in all ESCO languages. Thus, ESCO is a source of information on competencies relevant to the labor market in the international dimension, both for the development of higher education standards and for the review of educational programs in higher education, given that professional standards are currently lacking in many professions. ESCO is published as Linked Open Data, and developers can use RDF format. In this work we consider ESCO as source ontology for semantic application that needs in information about structure of skills and competencies.

AdvisOnt system

Agricultural advisory systems are widely used now for fast dissemination of agricultural knowledge and information, introduction of modern scientific research and technologies

in production, mobility and constant advanced training of agricultural specialists. Their implementation becomes an important factor in competitiveness of rural economy.

Development of the agricultural sector causes the dissemination of modern knowledge among agricultural manufacturers, relevant and efficient training and information support of their employees. AdvisOnt is an agro-advisory system that ensures consulting services for the agricultural sector of economy. It implements an ontological representation of advisory knowledge. AdvisOnt provides formalization and harmonization of semantic models of advisory objects with use of semantic identification and documentation of non-formal and informal learning outcomes and competence-based representation of advisory IOs [Rogushina and Priyma, 2017].

We consider this system because in analysis of semantic application we need in information about realization of knowledge base and methods of its processing. Therefore, we can propose objective information about compliance to the Semantic Web requirements only for those software where we participate in its design and development.

General architecture of AdvisOnt.

The general architecture of AdvisOnt defines relations between main subjects of advisory activities (see Figure 2):

- applicant – person needed in some work in agricultural domain and has a set of relevant competencies and skills;
- employer – person or organization needed in employees for execution of some task or work on some position;
- providers of learning services – organizations that propose various (formal, non-formal and informal) learning means for expansion of personal competencies;
- advisors – experts specialized in agricultural domain of fixer region that can use domain knowledge for refinement of mutual interests of employers and applicants and provides advising services if applicant qualification needs in additional learning according to employer demands.

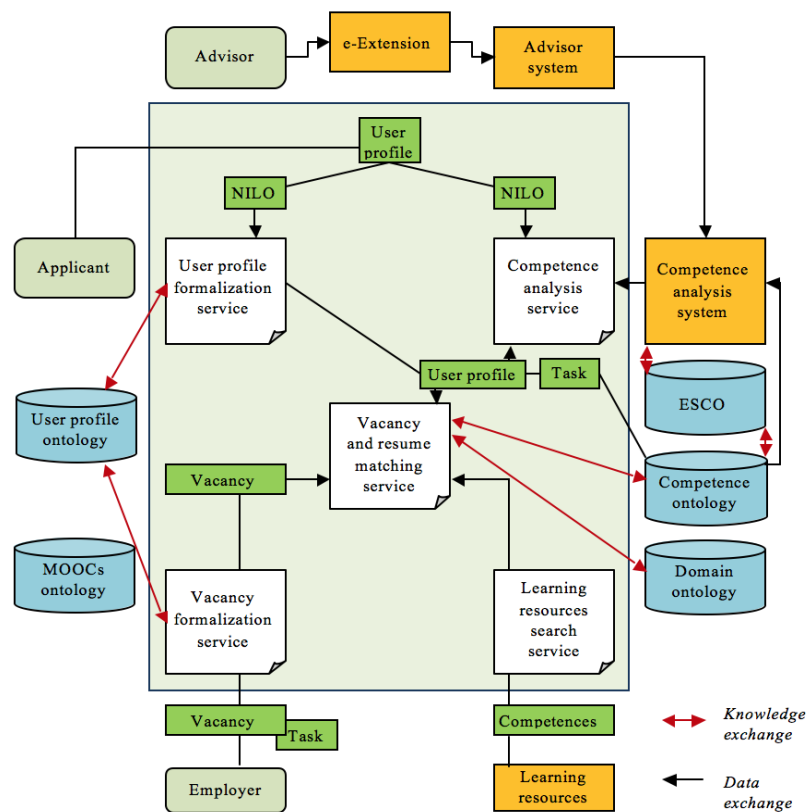


Figure 2: General architecture of AdvisOnt advisory system

AdvisOnt helps in interaction between expert-advisor and other subjects by e-Extension interface and uses external semantic IRs and knowledge bases: ESCO as a source of structured representation of domains competencies and qualifications; user profile ontology to determine the structure of the applicant's model; domain ontologies containing facts and rules of specific agricultural tasks; expert knowledge and soft skills used for semantic formalization and matching of vacancies and resumes; ontology of open online learning services (such as Massive Open Online Courses (www.mooc.org)). All classes of ESCO ontology used by AdvisOnt are stored into Turtle file. SPARQL queries and connectors are used for selection of skills and occupations from this RDF repository. The answers of SPARQL queries can be represented as result sets or RDF graphs. In the same way, the results of requests are returned to the RDF repository. Analysis of this ontology is used to define semantic similarity estimates for competence concepts [Rogushina, 2019]. Domain ontologies are integrated into the RDF repository with use of database of semantic graphs GraphDB. This database complies with W3C

standards and links data from various IRs, indexes them for semantic search and uses elements of NL analysis. GraphDB connectors provide fast search for keywords and aggregations usually realized by external services with use of synchronization on level of entities defined by URI, properties and property values.

We can consider AdvisOnt as SA because it provides [Pryima, Strokan et al, 2020]:

- personified interaction for potential employers and job seekers based on use of personal intelligent agents;
- registration of vacancies and resumes with their semantic analysis for formalized representation of used terms (on base of NL texts markup by ESCO ontology concepts);
- comparison of resumes and vacancies at the semantic level with use of semantic similarity of domain terms semantic relations between professions, knowledge, skills, competencies and qualifications defined by ESCO ontology;
- personified search of educational services and training courses based on validation of learning outcomes of formal, informal and non-formal level;
- comparison of training courses and programs with professions on base of their atomic competencies [Gladun, Khala et al, 2015].

Semantic components of AdvisOnt.

From the point of view of relations between components of semantic technology we can AdvisOnt contains:

- external KBs represented by ESCO ontology, MOOC ontology for learning courses, various agriculture domain ontologies and user profile ontologies from other IISs and internal ontology of competencies and qualifications;
- IO models that formalize structure and features of typical advisory IOs: competencies, skills and professions, applicants, vacancies and resumes, etc.;
- IRs that contain semantic markup based on structure of typical IOs and provide additional actual information about individuals of classes: semantic Wiki resources

that can contain NL text and multimedia (see Figure 3) such as wiki.isofo.kiev.ua and vue.gov.ua:

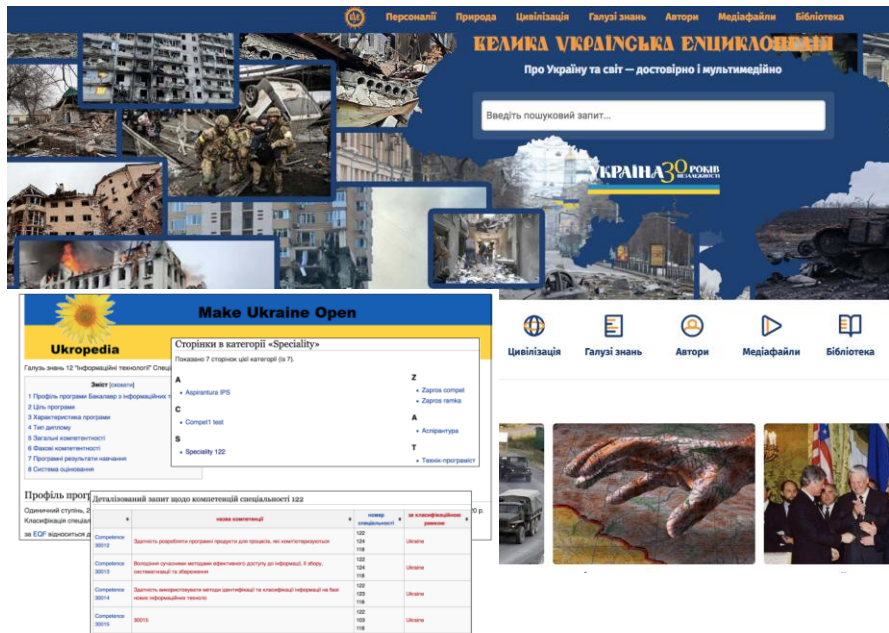


Figure 3: Example of Semantic Wiki resource pages

All these elements are integrated with the help of AdvisOnt ontology that defines relations between main components of advisory process and information sources used in this process (see Figure 4).

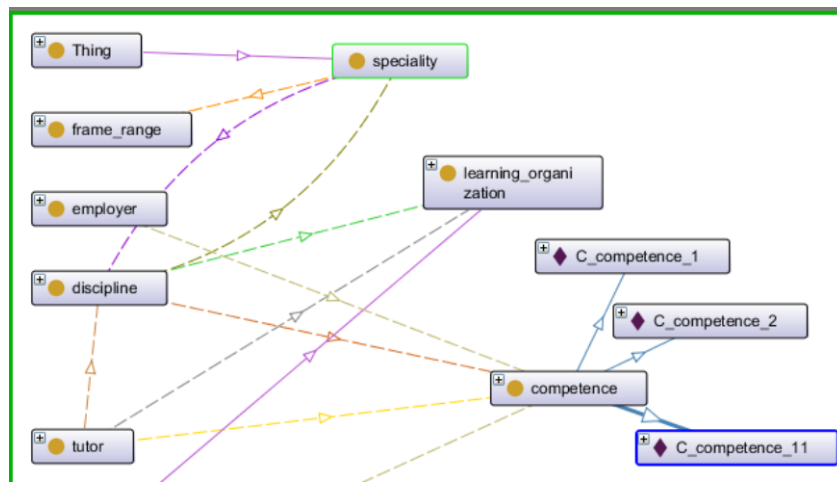


Figure 4: AdvisOnt ontology

This ontology shows the meaning of relation between subjects (employers, tutors, organizations and persons) and objects (competences, skills, professions, resumes, vacancies).

We can rate AdvisOnt as semantic application because this IS conforms to the requirements of the Semantic Web Challenge: Minimum requirements of the Semantic Web application for AdvisOnt can be interpreted like this:

1. data meaning plays a key role in its functioning: AdvisOnt process meaning of vacancies and resumes with use of ontologies to link various terms with concepts and realize original non-trivial approach based on atomic competencies for matching of IOs that cannot be obtained without analysis of their semantics.
2. AdvisOnt uses ontologies and IRs from different owners (ESCO, MOOCs, etc.) that can be hanged by other ones (for example, by ontology of national qualification system or other e-learning platform) without changes of software, these sources are heterogeneous syntactically (ontologies, Wiki IRs, thesauri, etc.) and semantically (use different NL languages and describe various domains), and contain real-world data used by other commercial applications.
3. Search for information is carried out in the real information space of the Web: results of AdvisOnt depend on user requests and actual information retrieved from the Web about vacancies, resumes and learning courses.

AdvisOnt works into the open information space, i.e. recommendations are not absolutely optimal but are based on available data and knowledge of system. This IS is based on processing of IRs that are represented on languages developed by the Semantic Web – RDF and OWL.

Conclusion and future work

We consider semantic approach to development of advisory IISs that provides methods and means of integration and unified representation of heterogeneous distributed knowledge and data. Analysis of the semantic IRs as an universal knowledge source about meaning of information objects helps in selection of external information sources of

advisory system. In future we plan to develop more detailed model of advisory IOs and use AI methods for acquisition of knowledge about these IOs from the Web IRs with various types of semantic markup and their metadata.

Bibliography

- [Bechhofer et al, 2004] Bechhofer, S., Van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., & Stein, L. A. (2004) OWL web ontology language reference. W3C recommendation, 10(2), 1-53. <https://www.w3.org/TR/owl-ref/>. Last accessed 6 Feb 2021.
- [Colardyn and Bjornavold, 2004] Colardyn, D., Bjornavold, J.: Validation of formal, non-formal and informal learning: Policy and practices in EU member states. In: European journal of education, 39(1), 69–89, 2004. DOI:10.1111/j.0141-8211.2004.00167.x
- [Decker et al, 2000] Decker, S., Melnik, S., Van Harmelen, F., Fensel, D., Klein, M., Broekstra, J., ... & Horrocks, I. (2000) The Semantic Web: The roles of XML and RDF. IEEE Internet computing, 4(5), 63-73.
- [Erriquez and Grasso, 2008] Erriquez, E., Grasso, F. (2008). Generation of personalised advisory messages: an ontology-based approach. In 2008 21st IEEE International Symposium on Computer-Based Medical Systems (pp. 437-442). DOI:10.1109/CBMS.2008.103.
- [Giunchiglia, 2009] Giunchiglia F., Shvaiko P., Yatskevich M. Semantic matching. Encyclopedia of Database Systems, 2009.
- [Gladun et al, 2015] Gladun, A., Khala, K., Sammour, G., Al-Zoubi, A., Schreurs, J. (2015) Semantic web and ontologies for personalisation of learning in MOOCs. In: Proceedings of IEEE 7th International Conference on Intelligent Computing and Information Systems, ICICIS-2015, 185–190. DOI: 10.1109/IntelICIS.2015.7397219.
- [Gorodetsky and Tushkanova, 2018] Gorodetsky, V.I., Tushkanova, O.N.: Semantic technologies for semantic applications. Part 1. Basic components of semantic technologies, In: Artificial Intelligence and Decision Making, 4, 61-71, (2018). DOI:10.3103/S0147688219050022
- [Gruber, 1991] Gruber, T.R. (1991) The role of common ontology in achieving shareable, reusable knowledge bases, Kr, 91, 601-602, <https://pdfs.semanticscholar.org/0256/614d1dc55962708db9b791f70a29370f1d76.pdf>. Last accessed 6 Feb 2021.
- [Guarino, 1998] Guarino N. (1998) Formal Ontology in Information Systems. Formal Ontology in Information Systems. Proc. of FOIS'98, 3-15. <https://www.ukma.edu.ua/~gor/courses/AIUS2009/LecturesAndMaterials/Formal%20ontology%20and%20Information%20Systems.pdf>. Last accessed 6 Feb 2021.

-
-
- [Hendler, 2001] Hendler, J.: Agents and the Semantic Web, *IEEE Intelligent systems*, 16(2), 2001.30-37. DOI: 10.1109/5254.920597.
- [Lassila and Swick, 1998] Lassila O., Swick R. (1998) Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation. <http://www.w3.org/TR/REC-rdf-syntax>. Last accessed 6 Feb 2021.
- [Lundqvist, Baker and Williams, 2017] Lundqvist K.O., Baker K.D., Williams S.A. An ontological approach to competency management, 2017, <http://www.eifel.org/publications/proceedings/ilf07/Contribution110.doc.pdf>. Last accessed 5 Feb 2021.
- [Miranda et al, 2016], Miranda, S., Orciuoli, F., Loia, V., Sampson, D.: An Ontology-based Model for Competence Management, *Data and Knowledge Engineering*, 107, 51-66, 2016. <https://doi.org/10.1016/j.datak.2016.12.001>.
- [Mommsen-Ghosh, 2004] Mommsen-Ghosh, D. (2004). Ontology Based Repository for Specifying Investment Advisory Services as a Knowledge Product. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.144.4017&rep=rep1&type=pdf>
- [Obrst et al, 2007] Obrst, L., Ceusters, W., Mani, I., Ray, S., Smith, B.: The evaluation of ontologies Semantic Web. In: Springer US. 2007, pp. 139–158.- <http://philpapers.org/archive/OBRTEO-6.pdf>. Last accessed 12 Feb 2021.
- [Pérez, Arenas and Gutierrez, 2009] Pérez, J., Arenas, M., & Gutierrez, C. (2009). Semantics and complexity of SPARQL. *ACM Transactions on Database Systems (TODS)*, 34(3), 1-45. <https://doi.org/10.1145/1567274.1567278>.
- [Pryima et al, 2020] Pryima, S., Strokan, O., Rogushina, J., Gladun, A., Mozgovenko, A. Ontology-based Methods and Tools for Validation of Non-formal Learning Outcomes. 12th International Scientific and Practical Conference of Programming (UkrPROG 2020), CEUR Workshop Proceedings, 2866 50–60, (2020). http://ceur-ws.org/Vol-2866/ceur_50-60Priyma5.pdf. Last accessed 12 Feb 2021.
- [Pryima, Strokan et al, 2020] Pryima, S., Strokan, O., Rogushina, J., Gladun, A., Lubko, D., Malkina, V.: Ontological Analysis of Outcomes of Non-formal and Informal Learning for Agro-Advisory System AdvisOnt, *International Conference on Technologies and Innovation*, pp. 3-17, Springer, Cham, (2020). https://link.springer.com/chapter/10.1007/978-3-030-62015-8_1. Last accessed 12 Feb 2021.
- [Pryima, Rogushina and Strokan, 2018] Pryima, S., Rogushina, J., Strokan, O.: Use of semantic technologies in the process of recognizing the outcomes of non-formal and informal learning, 11th International Conference of Programming UkrPROG, CEUR

- Workshop Proceedings, 2139, 226–235, (2018). <http://ceur-ws.org/Vol-2139/226-235.pdf>.
- [Rogushina, 2019] Rogushina, J.: Use of Semantic Similarity Estimates for Unstructured Data Analysis, XIX International Scientific and Practical Conference "Information Technologies and Security" (ITS 2019), CEUR 2577, 246-258, (2019). <http://ceur-ws.org/Vol-2577/paper20.pdf>. Last accessed 8 Dec 2020.
- [Sabou, 2008] Sabou M. Exploring the semantic web as background knowledge for ontology matching / Sabou M., d'Aquin M., Motta E. // Journal on Data Semantics, (2008). DOI: 10.1007/978-3-540-92148-6_6.
- [Shvaiko and Euzenat, 2008] Shvaiko P., Euzenat J. Ten challenges for ontology matching. Proc. of the 7th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE), 2008. <https://hal.inria.fr/hal-00825956/document>
- [Staab and Studer, 2013]. Staab, S., Studer, R.: Handbook on ontologies, Springer Science and Business Media, (2013).
- [Studer, Grimm and Abecker, 2007] Studer, R., Grimm, S., Abecker, A.: Semantic web services. Springer, (2007).
- [Warren, 2006] Warren P. Knowledge Management and the Semantic Web: From Scenario to Technology. IEEE Intelligent Systems. 2006. V. 21, 1.53-59.

Authors' Information



Julia Rogushina – Institute of Software Systems of the National Academy of Sciences of Ukraine, PhD (Physics and Mathematics); Senior Researcher, 40, Acad. Glushkov av., 03187, Kyiv, Ukraine; e-mail: ladamandraka2010@gmail.com

Major Fields of Scientific Research: Ontological Analysis, Semantic Web Applications, Semantic Wiki, Big Data Analytics



Anatoly Gladun – International Research and Training Center for Information Technologies and Systems under NAS and MES of Ukraine; PhD (Engineering), Senior Researcher of the Department of Complex Research of Information Technologies and Systems. 40, Acad. Glushkov av., 03187, Kyiv, Ukraine; e-mail: glanat@yahoo.com

Major Fields of Scientific Research: Ontological Analysis, Semantic Web Applications, Multiagent Systems, Big Data Analytics.



Serhii Pryima – *Dmytro Motornyi Tavria State Agrotechnological University, 18 B.Khmelnysky Ave, Melitopol, Zaporizhzhia obl., Ukraine;*
e-mail: pryima.serhii@tsatu.edu.ua

Major Fields of Scientific Research: *Ontological Analysis, Competence Analysis, Non-formal learning*



Oksana Strokan – *Dmytro Motornyi Tavria State Agrotechnological University, 18 B.Khmelnysky Ave, Melitopol, Zaporizhzhia obl., Ukraine;*
e-mail: oksana.strokan@tsatu.edu.ua

Major Fields of Scientific Research: *Ontological Analysis, Competence Analysis, Non-formal learning*

A NEW APPROACH TO AUTOMATIC ONTOLOGY CREATION FROM UNTAGGED TEXT ON NATURAL LANGUAGE OF INFLECTIVE TYPE

Anna Litvin, Vitalii Velychko, Vladislav Kaverinsky

Abstract: *A new approach and its software implementation was presented to build a deeply semantically structured OWL ontology using as an input just a plain natural language text without regular structure or any previous targeting. The approach is primarily aimed to inflected languages; especially it is already implemented for Ukrainian. The created OWL ontologies could be easily converted to other graph databases formats, for instance – Neo4j and were successfully tested as valid using Protégé, RDFlib and Neo4j GUI environment.*

Keywords: *NLP, NLU, ontology, automatic ontology creation, OWL.*

ITHEA Keywords: *H.3.4 Systems and Software.*

Conference and topic: *Applied Aspects of Software Engineering: Software Designing.*

Introduction

The boring and time and labour consuming process of manually database (DB) creation inspires finding of new techniques to automate it. So that only a slight manual correction might be needed to make the DB more perfect. The modern level of computation and software technics development allows us to solve this problem. Completely automatic ontologies creation could be treated as a particular type of machine learning which consists in semantic structuralization of large amounts of text information in a way acceptable for the subsequent machine manipulation. This could be, for example, automatic gathering of information from the word web for a certain topic.

Because one of the popular languages for ontologies construction is OWL [OWL], which is actually a standard, it was considered as the basis for an ontology creation in this work. Moreover, an OWL ontology could, if you like, be easily converted to a different graph DB format, for instance – Neo4j.

Since the ontology creation automation is highly desirable it has been considered in many works for probably quite a long time, for instance, in [Young, 2018; Elnagar, 2020; Balakrishna, 2008]. However, the building technique and structure of ontology are highly dependent on its goal and destination as well as on the initial natural language the program should deal with. It should be noticed, that ontology creation atomization is not worked out enough for Slavic inflective languages (NL), especially Ukrainian. Our previous work [Litvin et al., 2021a] is devoted to this question. However, the approach considered there is only acceptable for previously tagged NL texts or texts sets with a predetermined regular structure (official documents, letters etc.). The presented here technique makes in the possibility to process merely plain texts without any previous targeting.

The basis of the proposed approach

In the basis proposed approach there is a rule based syntax-semantic analysis method. It is known that in the languages of inflective type the main role in words bounding in a sentence belongs to the combination of certain flections (changeable endings of words). A developed language system of highly varying flections combinations for different parts of speech allows expressing of considerable semantic information. Thus, merely analysis of words forms for compatibility check has the potential to bring might not all, however a significant part of semantic information.

In the considered approach we have proposed about 90 semantic types each of them actually could have several (up to more than a hundred) subtypes, which depend on such additional characteristics as gender, tense, number or a certain preposition for each of the words from the considered pair. However, these additional subtypes are not actually used in an ontology creation, they may be useful in the future for even more deeper and precise information structuration. Moreover, they come directly from the so called “correlators”, which are actually the part of the analyzer, that give the programs the options in what a way one or another semantic type could be expressed. There exist

actually different systems of semantic types. The set considered here could be derived from the astigmatic basis set by the method proposed in [Litvin et al., 2021b].

The ontology creation from a text includes two main operations performed by the separated program modules. The first and, probably, the main and the most resource consuming part consist in the syntax-semantic analysis of the initial text. The corresponding program module includes a manually created set of so called “correlators” and “determinates”.

The “determinates” represent combinations of possible flexions of words and prepositions between them (if any) and corresponding to each options of semantic subtypes. Also there is given whether the combination inverse and which of the words of the pair is to be considered as the main one. Names of the semantic subtypes are given as contentious symbols, for instance K1001, K4801, K6201 etc. Here is an example of a row from the file with “determinates” for Ukrainian language:

```
ім під ям    L    I    K6201K8644K8646
```

here we can see flexions for the first and the second of the possibly linked words, which are “ім” and “ям”. The preposition “під” is assumed to occur between them. The symbol “L” marks that the main word is the left one, as the flexions given and the symbol “I” says that the link is inverse. The possible semantic subtypes for the combination are designated as K6201, K8644, K8646.

The “correlators” represent the correspondence of each of the semantic subtypes to the possible options of parts of speech combinations, including their order in the pair. Also there are given the verbose names of semantic types (“macro types”). For each of this “macro types” could be several (up to more than a hundred) suitable subtypes. Here is an example of a row from the file with “correlators” for the Ukrainian language:

```
K3506      separability_of_an_action
```

```
S4S1;S4S6;S4S13;S4S5;S4S3;S4S10;S4S11;S4S12;S4S18;S4S22;S4S25;S4S2
```

Here we have “K3506” which is a name of a semantic subtype. Then there is “separability_of_an_action”, which is a verbose name of the corresponding semantic “macro type”. Then there is a sequence of possible parts of speech pairs given as contentious symbols. For example, S1 is “noun”, S4 is “verb”. The pairs are separated with semicolons.

Also the program has a dictionary in which are included word stems, lemmas and flections. The dictionary gives the correspondence between stems, lemmas and sets of flections. The dictionaries are stored in a special compact format and could be automatically created using open language data.

The main purpose of the program is to use these data and the initial text to find and typify the words in it, to determine their stems and flections, and recognize the links between the words and figure out their semantic types. The other result of such analysis is obtaining the linked words groups in the sentences. Formally the group is a fully connected graph. Practically such groups could correspond to a whole simple sentence, a part of a complex sentence or a participial sub phrase. The product of this modulus is two XML files: allterms.xml and parce.xml. They are used for the subsequent OWL-ontology creation.

The allterms.xml file is merely a list of the terms – nouns and name groups found in the analyzed text with some of their characteristics. It consists of the two main parts: <exporterms> and <sentences>. The first one includes the terms.

Here is an example of how a term is presented:

```
<term>
  <ttype>Noun_noun</ttype>
  <tname>книга людини</tname>
  <wcount>2</wcount>
  <osn>кни</osn>
  <osn>люд</osn>
```

```
<sentpos>1/1</sentpos>
```

```
<sentpos>1/2</sentpos>
```

```
<reldown>2</reldown>
```

```
<reldown>4</reldown>
```

```
</term>
```

Tag `<ttype>` represents the parts of the speech sequence that forms the term. Tag `<name>` is the text of the term as given. Tag `<wcount>` gives the number of the words in the term. Tags `<osn>` are given for each of the words and represent stems. Tags `<sentpos>` give the term's words positions in the text (the sentence number – from 0 / the word position in the sentence – from 1). Tags `<reldown>` and `<relup>` are optional. They show the relationship of the considered term to other terms from the file. Tag `<reldown>` points to the term of narrowing context – each of its words could be found in this term, but the considered one contains more. Tag `<relup>` points to the term of expanding context – contains all the words from this term and some others. Tags `<reldown>` and `<relup>` help to build the terms hierarchy in the creating ontology.

The part `<sentences>` merely includes the texts of all the sentences from the considered text in tag `<sent>` each.

The file `parce.xml` represents the syntax-semantic scheme of each sentence of the text. The sentences structures are given in container tags `<sentence>`. This container includes the following tags: several tags `<item>` representing the words and their characteristics; `<sentnumber>` – number of the sentence in the text (from 1); `<sent>` - the text of the sentence. Here is an example of a tag `<item>`:

```
<item>
```

```
<word>Книга</word>
```

```
<osnova>кни</osnova>
```

```
<lemma>книга</lemma>
```

```
<kflex>а</kflex>
```

```
<flex>ra</flex>  
<number>1</number>  
<pos>1</pos>  
<group_n>1</group_n>  
<speech>S1</speech>  
<relate>0</relate>  
<rel_type>K0</rel_type>  
</item>
```

Tag <word> contains the word's text as given in the considered text. Tag <osnova> represents the stem of the word. Tag <lemma> gives the lemma, the initial form of the word. Tags <kflex> and <flex> are flections. <kflex> - is just the ending, but <flex> is the whole word's part that could be changed. Tag <number> is the number of the word in the sentence. Tag <group_n> shows the belonging of the word to a linked group from the sentence. Tag <speech> contains the mark of the corresponding part of speech. Tag <relate> indicates the number of a word from which a semantic link is going to the considered word. If the word does not have any incoming links, like in the example above, its value is set to zero. And the tag <rel_type> is the type (subtype) of the semantic link. Value K0 means the absence of the relationship or its unknown type.

The ontology structure and creation technique

The obtained two files are used to create an OWL-ontology. Before describing its creation technique let us consider the appropriate ontology structure, which building is the goal.

The ontology has classes and properties. The main classes are: Action, Adjective, Adverb, Name, Number, Preposition, Term, Negation, UndefinedEntities. So we can see that ontology entities are sorted by their parts of speech. Subclasses of the class Term are the nouns groups and nouns with hierarchical structures. Class Name descendants are given names. Others entities' destination seems rather clear. The main properties are

the following: SentenceGroups, Groups and WordsLink. The descendant properties of WordsLink are the semantic types. In certain ontology all 90 of them may not be presented but only ones that appear in the considered text. The descendants of these properties are the certain links between entities (represented through classes). A “Domain” of such property is the main word in the linked pair, and the “Range” is a dependent one. The property Groups descendants represent the linked groups of sentences. They may have types given in their range field. It could be “Subordinate”, “Participial” or just nothing for other cases. The property SentenceGroups descendants represent the sentences, and descendants are subsequently linked words groups.

Having the two mentioned files and file with “determinates”, which include the semantic types verbose names, it is not hard to create the briefly described ontology. All the OWL entities are first created as the OOP representation objects. The root classes and properties are to be created first and obligatory. Then, using the file allterms.xml a hierarchical structure of the terms (nouns and name groups) is created. Then using the parce.xml are created the classes of others types of words and properties corresponding to the links between the words. At the same time properties of the type WordsLink are created, which represent the semantic types. Since in the file parce.xml only a certain subtype is given “determinates” are used to determine the semantic “macro type”. Because the words belong to the linked groups and they to sentences, this information is used to create correspondently Groups and SentenceGroups sub-properties. The Groups and SentenceGroups descendant properties are supplemented with a label that contains the text of the corresponding group or sentence. This contexts seem to be useful for more informative ontology answers. The typifying of the linked words groups is performed by the presence of certain words in the group: subordinating union, participle, gerund.

Testing of the ontology showed its possibility to obtain different information types on a given subject and the contexts where it is actually said. Thus the developed software system seems to be a useful instrument for automation of reference systems DB creation and the word web data aggregation.

Conclusion

A technique was proposed for automatic OWL-ontology creation from a natural language text. It is assumed that the language of the considered text is one of inflection type. Another peculiarity of the method is that it does not need either any previous text tagging or a regular structure.

The essence of the technique consists in rule-based syntax-semantic analysis methods and the fact that a great amount of semantic information in inflective languages could be obtained through the analysis of the combinations of different parts of speech flexions and prepositions. The ontology creation includes two stages: syntax-semantic analysis with intermediate XML-files creation and building based on them the corresponding OWL-file.

The proposed method was implemented as software and parameterized for Ukrainian language. It was tested on real texts, which showed its efficiency. The created ontologies appear correct and able to be processed through Protégé, RDFlib and Neo4j (using Neosemantics plug-in). The ontologies seem to be deeply semantically structured and at the same time rather simple and regularly organized. Thus, the developed software system seems to be a promising tool, which can significantly and effectively automate graph DB creation using just plain texts.

Bibliography

- [OWL] Web Ontology Language (OWL), <https://www.w3.org/OWL/>
- [Young, 2018] B. Young, A. JungHyen. Methodology for Automatic Ontology Generation Using Database Schema Information. *Mobile Information Systems*, 2018. pp. 1 – 13.
- [Elnagar, 2020] S. Elnagar, V. Yoon, M. A. Thomas. An Automatic Ontology Generation Framework with An Organizational Perspective. *Proceedings of the 53rd Hawaii International Conference on System Sciences*. 2020 pp. 4860 – 4869.
- [Balakrishna, 2008] M. Balakrishna, M. Srikanth. Automatic Ontology Creation from Text for National Intelligence Priorities Framework (NIPF). *Proceedings of 3rd International Ontology for the Intelligence Community (OIC) Conference*. 2008. pp. 8 – 12.

[Litvin at all, 2021a] A. Litvin, V. Velychko, V. Kaverinsky. Automatic Ontology Creation from Natural Language Text. International conference on software engineering "Soft Engine 2021". Kyiv, NAU. 2021 pp. 21 – 25.

[Litvin at all, 2021b] A. Litvin, V. Velychko, V. Kaverinsky. Approach to creating a system of semantic relations classification for the analysis of natural language texts. International Scientific Symposium (satellite) "INTELLECTUAL SOLUTIONS-C". 2021.

Authors' Information



Anna Litvin – Glushkov Institute of Cybernetics of NAS of Ukraine; postgraduate student. 40 Glushkova ave., Kyiv, Ukraine, 03187; e-mail: litvin_any@ukr.net

Major Fields of Scientific Research: natural language processing for analysis



Vitalii Velychko – Glushkov Institute of Cybernetics of NAS of Ukraine; Ph.D. Senior Researcher. 40 Glushkova ave., Kyiv, Ukraine, 03187; e-mail: aduisukr@gmail.com

Major Fields of Scientific Research: information systems with the processing of natural language objects



Vladislav Kaverinsky – Institute for Problems of Materials Science, Ph.D. Senior Researcher, 3 Krzhizhanovsky st., Kyiv, Ukraine; e-mail: insamhlaithe@gmail.com

Major Fields of Scientific Research: mathematical modeling

CODE STATIC ANALYSIS BY REVERSE ENGINEERING TOOLS

Mykyta Krainiy

Abstract: Paper proposes a conclusion of a review of possibilities of static code analysis for Altova UModel. Area of analysis – C# scripts created for Unity3D Games. Visualization tools – reverse engineering technique designing of Sequence diagrams from source code.

Keywords: Unity Script, Altova UModel, Sequence diagram, Component-Oriented Approach.

ITHEA Keywords: H.3.4 Systems and Software.

Conference and topic: Theoretical Aspects of Software Engineering.

Introduction

Today we have a large amount of software component, created for Unity3D. Creating of new game with reuse of existing components requires to “recover” and visualize algorithms of game hero acting. After visualization of these algorithms, for example using UML diagrams, developer may compare internal structure of the software components and functional requirements for the game.

Code Static Analysis of Unity3D Scripts

Lets consider an example showing peculiarities of static code analysis for Unity3D.

According to the approach of designing game applications

```
private void Gameplay()
{
    // every 10 seconds refer to new portion of words
    InvokeRepeating("NewWords", 10f,10f);
    InvokeRepeating("LaunchCubes",01f,02f);

    if (over == true)
```

```
{  
GameObject.Find("Placeholder").GetComponent<Text>().text = " Гру закінчено";  
    CancelInvoke();  
}  
}
```

The essence of the script - Some periodical actions are performed through timer and when all timers are stopped game is finished.

Let's see sequence diagram of this method, generated by Altova® UModel [Altova, 2021].

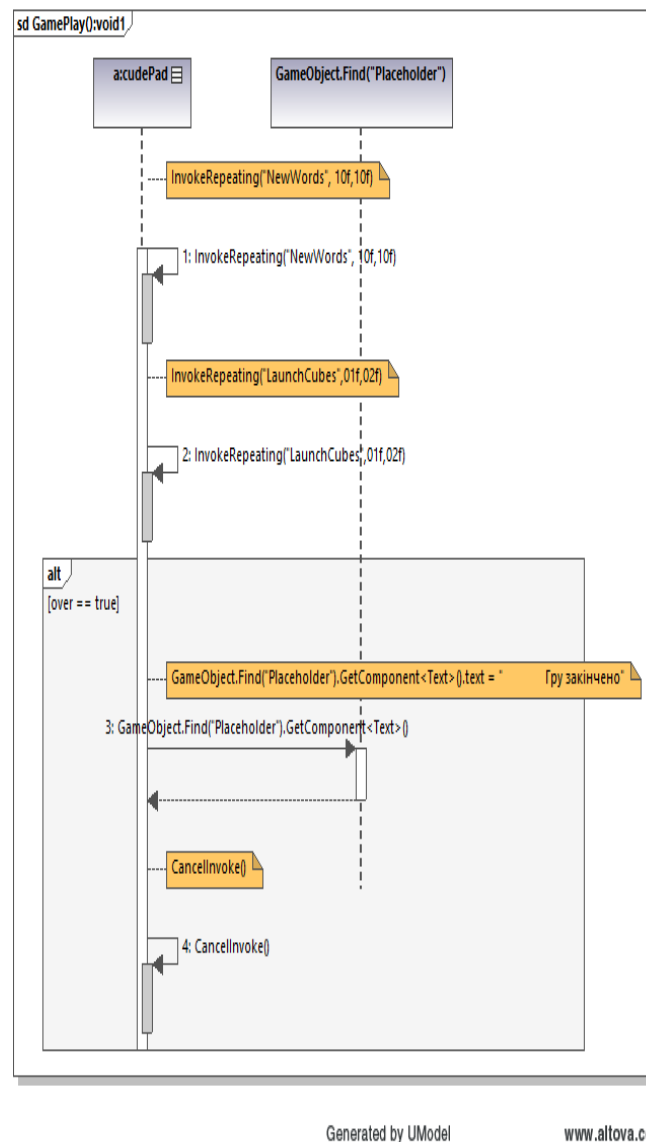


Figure 1: Sequence diagram of Gameplay

Analysis of these and many other sequence diagrams show that the depth of calls of Altova does not allow to consider real interaction between objects. For reverse engineering activities it will be useful to see interaction between objects (In unity terminology GameObjects) that are involved to functions NewWords() and LaunchCubes().

Proposed Approach for Static Analysis of the Code

In order to perform visualization of C# Unity scripts this it is proposed:

1. To design sequence diagrams of these methods and to perform more deep analysis of game algorithms.
2. Using analytical approaches to represent sequence diagrams from all methods that are invoked from method under research [Chebanyuk, 2018], using merging algorithms [Kolovos, 2015] compose common sequence diagram from several ones.
3. Perform translation from analytical approach to XML.
4. Using reverse software engineering tool “recognize” a XML as a sequence diagram.

Conclusion

In Component–Oriented IDE interaction between objects is done by means of finding other objects in gameScene and with further finding of corresponding components. Usually method that initiates finding operations stores referencies to other objects. That why approaches to generate Sequence Diagrams must consider objects inside of other methods (components) and messages between them.

Experimental results shows that settings of existing reverse engineering tools do not allow to perform deep static analysis of the code. In order to improve the situation more complex analytical should be involved to the analysis process [Chebanyuk at all, 2020].

Bibliography

[Altova, 2021] Altova <https://www.altova.com/es/umodel>

-
- [Chebanyuk O., Palahin O., Markov K., 2020] Chebanyuk O., Palahin O., Markov K., 2020. Domain engineering approach of software requirements analysis. Проблеми програмування. №2-3,2020, 164-172. DOI: <https://doi.org/10.15407/pp2020.02-03.164>
- [Chebanyuk O., 2018] Chebanyuk O., 2018. An Approach of Text to Model Transformation of Software Models. In Proceedings of the 13th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2018), 432-439
- [Kolovos at all., 2009] Kolovos D., Paige, R., Rose, L., Polack, F.: The Epsilon Book. University of York, York (2009)
-

Authors' Information



Mykyta Krainiy – Assoc. Prof. of the Department of Software Engineering, National Aviation University, Kyiv, Ukraine.
e-mail: nickkhryne@gmail.com

Major Fields of Scientific Research: Reverse software engineering, Software modelling.

SECTION: APPLIED ASPECTS OF SOFTWARE ENGINEERING
**DESIGN PRINCIPLES FOR SOFTWARE USING TECHNIQUES OF BIOLOGIC
INFORMATION PROCESSING**

Karl Javorszky

Abstract: We propose to use a technique of counting ‘in stereo’, by using **two** counting systems concurrently. Research in number theory has brought up a very slight inner discongruence within the numbering system. The proposal is to root the definition of the term ‘information’ in properties of natural numbers, with the refinement that the extent of information is equal to the extent of being otherwise. The property of being otherwise can be numerically defined on a collection of pairs of natural numbers. An etalon collection of pairs (a,b) is used to demonstrate the concept of being otherwise. The property of being otherwise is based on the numeric fact, presented in oeis.org/A242615, according to which there are **two** upper limits of the number of logical relations existing among members of collections, namely sentences which state the existence of **similarity**, ‘**n?**’ contrasted to sentences that describe **differences**, ‘**n!**’, among the members of one and the same collection. The values **f¹(number of different relations)** point back to slightly differing values of **n**, and this in the form of $\Delta (f^1(n?), f^1(n!)) \{=, <, =, >, >>\}$ for $n \{1, 11, 32, 66, 97, 140\}$. The numeric incongruences appear to play important roles in Physics, Chemistry, Theoretical Biology. Here, in the context of software design, the practical relevance of being able to switch between **two** methods of describing one and the same state of a set, becomes apparent in the **increased efficiency** in the transmission of messages. Recognition of forms and patterns in (eg. optical) multitudes becomes possible by using the properties of **cycles**, which are an artefact of periodic changes which sender and receiver subject the collection of tokens to, in a pre-agreed (or, in Nature, a-priori existing) algorithm. The theoretical foundations for the possibility of recognition of Gestalts are presented to be implications of the inner incongruence of the numbering system, as shown by oeis.org/A242615. There exist **shortcuts**, abbreviated names for **patterns**. This is equivalent to a condensation of information.

Keywords: Information, fundamental definition of; memory; learning; AI; enumerations of states of sets.

ITHEA Keywords: F.4.1, H.2.4, I, I.2.11, I.6.5, H.1, I.4.8, J.3, A.1, H.3.3, E.1, D.2.2, C.2.4, I.5.1, D.2, A.0

Conference and topic: SE 2022 SoftEngine International Conference on Software Engineering <http://www.ithea.org/softengine/>

Introduction

We ask the reader to conceptualize an inbuilt quirk in the numbering system. The quirk itself is pointed out in oeis.org/A242615.

In essence, we make use of the inbuilt relative inexactitude to create a counting and enumerating system which is more detailed, more exact **by orders of magnitude** compared to the counting – enumerating system presently in use. One may hope that the fineness of units used in quantum-based models can be achieved by using the model presented here, which creates fractions of units.

Example:

We use *two* eyes to perceive the world. Of the *two* pictures delivered by perception, cognition creates *one*, integrated picture. It is a culturally anchored, common-sense approach that we speak about the form and content of that *one* picture, which has been assembled, if we talk about the world. As senders of messages, we rely that the receiver has a comparable neurological apparatus at their disposal. If we say “House H is located at the corner of Streets A, B”, we think of *one* landscape or system of coordinates, where we can point out a specific location. We know that the recipient of this message will similarly employ *one* mental or physical map, on which they shall identify a location.

It would be quite unusual, if we said: “As seen by the **left** eye, House H is located at corner of Streets A, B, *displaced by the perspective bias ‘left’* by λ %; as seen by the **right** eye, House H is located at corner of Streets A, B, *displaced by the perspective bias ‘right’* by σ %.” Aside of being unusual, there is nothing to say against this way of messaging the coordinates. The sentence is grammatical and we do have the technical means to establish exactly what we refer to by the terms (λ, σ) .

The main hypothesis is that it is **worth while to refrain from integrating**, merging, fusing the two underlying pictures into one overriding picture, which we then by cultural tradition call The Picture. We do have the accounting tools and computer capacities to

categorize and enumerate the slight deviations caused by factors (λ, σ) . The slight fragments that distinguish the left picture from the right one are a nuisance in our current understanding. We do not question Nature's ability to merge the two pictures into one within our brain. It can be done, and it is very practical that it is done, and common sense requires that if we talk about pictures we mean the net result, and not the underlying additions and subtractions of some small details that are artefacts of factors (λ, σ) .

The question arises naturally, what benefit would it bring if we transmit the same coordinates in a more circuitous way, by using two reference points instead of one to build up the picture. There is no sense in describing a picture more exactly, if it is and remains the same picture about the same state of the world.

The surprise is that there are additional benefits to using two reference systems instead of one, further and beyond of being exact for the exactitude's sake. In the example with the two optical pictures coming from two eyes, the focus is on that part of the two pictures which are congruent, translatable into each other and refer to one and the same state of the world.

Using the two metrics in a fashion that is not optimized towards congruence, one sees interdependences in the system which are best approached in a context of economics.

It is more *efficient* to transmit a message in batches of 66 tokens which can be in any order (think of beads in a sack) than have 66 tokens in a sequential order (think of beads forming a necklace).

It is more *efficient* to transmit a message in batches of $6 * 11$ tokens which are sequenced than have 66 tokens which can be in any order. in a sequential order (think of beads forming a necklace).

The term *efficient* refers to the number of tokens needed to carry a given number of messages. The question: "*How many logical relations make up one object?*" can now be numerically approached.

The stereo approach to counting is logically solid, is nothing but an added chapter to arithmetic. Its basic concepts run against deep-seated neurological and cultural conventions and habitudes, but not against rules of computing.

The general theory underlying the model presented here is presented in the form of a logical treatise in [Javorszky, K., 2016].

Formal Introduction of the Model

1. What is the problem?

The meaning of the term 'information' needs to be rooted in the numbering system. Information is a phenomenon of Nature and also of our mental creations. It has to be understood as a fundamental property of our numbering system. The numeric facts shown in A242615 allow for using *two* reference points on **N**. Every measurement (statement, result) can be reformulated using principles and techniques known from interferences.

A second problem is, that in the theory in force until presently, it is not possible to *compress and unfold* information, like we observe Nature doing it in the genetic information transfer and in the memory. If one has only *one* **N**, one cannot collapse and recreate the rules that enforce properties of material assemblies. Having in addition to **N** further two variants of **N**: **N λ** , **N σ** allows protracted exchanges of *relation* \leftrightarrow *fragment of object*.

2. Why is it interesting and important?

It is high time to update the philosophy behind the fundamental axioms that restrict our ways of thinking about counting. In the 21st century, in the midst of a social revolution regarding individuality, we should switch from the idea of the unit being of one standard kind, indistinguishable from any other units, to the idea of the unit being a cohort of individuals, which are by their nature similar and different to the other units of the cohort.

The time of infallibility and persuasion by the sword has passed her best days. It is time to allow for Truth to be not the one-and-only variant but to acknowledge the existence of alternative truths in the abstract fields, too. Using a model in which alternative truths exist

requires multidimensional thinking, which the human brain is not as good at as are computers. A world view built by computers needs some getting used to, but its utter boring rationality gives one a solid halt in a world that is full of logical rivalries, contradictions, local breakdowns. The picture is not as unified and harmonic as previously, but is much closer to reality.

3. Why is it hard?

European culture, rooted in Judeo-Christian concepts of monotheism, has been maintaining that its concepts fit into each other seamlessly, perfectly, free of contradictions and ambiguities. The Cathars, who have evolved a system of concepts of duality, have been put to the sword as heretic in the 14th century. Their fate has been a strong warning for anyone experimenting with epistemological concepts of dialectic and tensions within the system model.

In our schools, children are taught, that $a+b=c$ as a narrative has only one valid interpretation, namely that of c , where all differences that are present in the version $a+b$ will **magically disappear** in the process of addition. While a,b can reasonably be assumed to be *two* ideas, c is imagined to be made up of *one* kind of units. The unification and regularization of basic units offers many advantages, yet the idealized version negates the existence of differences that were present in the form of a,b . Our children are offered a sanitized, family-friendly version of the relations between the parts and the whole. In that version, taught up into our days, everything adds up finely, and there are neither rivalries nor conflicts among the parts of a whole.

We have been taught to use *one* sheet of millimeter paper as the background of any of our mental creations and explanations. This attitude may be practical and easy to administer, but it fails to take into account the actual reality of our living in a world that is based on *duality*. Whether it is *proton* ↔ *neutron*, *DNA* ↔ *RNA*, *female* ↔ *male*, *foreground* ↔ *background*, *sequential* ↔ *contemporary*, *breathing in* ↔ *breathing out*, *etc.*, we are confronted with empirical observations that call out our time-honored, traditional, idealized and simplified fundamental concept of *mono-perspectivism*.

We know of the practical advantages of using interferometry. Triangulation as a technique is a simple implication of using *two* sets of references (as opposed to the present system of ideas, in which all measurements are related to that one-and-only reference point of *0, Zero*).

As the example with the picture in the brain, and in our communications among each other, shows, having a convention that the *one, merged* picture is the only one to be referred to, brings forth advantages of uncomplicated and easy design principles for any system of thoughts that builds a model to depict Nature with. The unfortunate fact is, that Nature builds from the bottom up and not from the top down. Before the merged picture can evolve, there must exist *two* versions of it, which become merged by faculties of our brain. The *two, slightly deviating* pictures are **prior to** the *one, perfectly fitting picture*. Nature builds on the differences, uses the differences. In the organizational level of adaptive, self-correcting feedback loops (our neurology, eg), the remainders and rests of the two pictures are recycled. Those fragments of the two basic pictures, which after having served their usefulness as constituents of the merged picture, remain as fragments (like snippets of pieces of cloth in a tailor's shop) and are recycled by the brain (we encounter them in the form of dreams); unrecycled these rests and remainders perform the procedure of *entropy*.

4. Why do naive approaches fail?

For psychologists, it is a natural and fundamental truth that there exist several versions of truth, each equally legitimate and valid. In the STEM environment, the concept of multicausal genesis is not yet rooted.

The a-priori existence of *patterns* is an undisputable fact of Nature. The classical mathematical philosophy does not foresee any meaningful relations among natural numbers as they stand. The importance of ***cycles*** as porters of a-priori logical relations between amounts, places and time has not yet been recognized.

The *upper limits* for the number of sentences that can be said about a collection of *n* members have only been dealt with in the form of counting sentences that describe

differences among the members before a background of similarities. The parallel measure, the maximal number of distinct sentences that describe similarities among the members before a background of diversities, has not been addressed yet. *The existence* of upper limits of relations on objects $f(n)$ paves the way to ideas of thresholds, quality transformations, saturation: these ideas are fundamental in biology, but not yet working day concepts in the technical sciences. *The extent* of upper limits – as we contrast the maximal similarity among objects with the maximal diversity among objects – opens up a completely new area of arithmetic, geometry and information transmission. The terms of trade in the *three-way Bazar* are **sliding $f(n)$** . There exists a numeric table for the terms of trade among *{number of similarity relations, number of diversity relations, number of objects}*.

5. Why hasn't it been solved before?

Reason 1.: Ideologic, theological reluctance to depart from the unified concept. This is well reflected in the underlying idea of the Shannon concept of information: that messages transmission is identifying 1 of elements of \mathbf{N} and communicating the algorithm to the receiver, who has at his disposal the same \mathbf{N} .

The idea to use *two* measuring rods that are calibrated slightly differently, and use the immanent interference pattern between the two to identify 1 of elements of \mathbf{N} as the background, and communicate the algorithm to the receiver, who has at his disposal the same interference pattern established by the slight relative mis-calibrations of the two measurement rods, this idea introduces concepts and techniques of *accounting*, with two different ledgers and one consolidated one. Maybe the idea is too much practical for number theory.

Reason 2.: Ehrenfels inhibition is the name for the observation, that humans' observational capacities are more challenged if the stimuli to be remembered are irritatingly close (similar) to each other. There appears to be a requirement of balance among similar and diverse so that our neurology functions best. We are, our neurology is optimized for an environment that contains in a relatively stable proportion input and

impressions of a similar and of a different nature. We would not be optimized for a feature of Nature if that feature was not there. So there exists in Nature, by itself, a proportion between that what remains similar and that what changes. Tradition has it that one speaks about that, and uses as a background that, what remains and is the same. Relative to this background, processes can be observed to take place. Exchanging the perspectives and inquiring what the picture looks like if changes and rearrangements are axiomatic and are used as background, before which we can observe that what happens regularly, inevitably, predictably is a mental dexterity which salesmen, diplomats, lawyers and psychologist are trained in, but not those educated in the technical sciences. The background and the foreground are of course merged in our cognition, and they are irritatingly similar to each other. Dissecting their properties needs some returning attention.

Reason 3.: Availability of computers is a key factor. It is absolutely impossible to gain insight into the interactions of (a,b) in the adult version of the story of their romance, without a computer.

6. What are the key components of my approach and results?

The building blocks of the model are:

a. the actors

We conduct a mental experiment with a cohort of pairs of natural numbers (a,b) . The cohort we prefer to use contains pairs of members which each can have 16 different variants, yielding a cohort size of 136 so-called *logical primitives*. (© M. Abundis).

b. the habitat

The logical primitives undergo **periodic changes**. We use *aspects* of the members, like $a+b$, $a-b$, $b-2a$ to create **sequential orders**. When 2 aspects are concurrently the case (two periodic changes happen concurrently, like the Earth's rotation and the Moon's position), there exists a coordinate for a **planar place** which is given by the two relevant sequential ranks, for each primitive. Using such planes that can be arranged

perpendicularly, by having common axes, we find *two* Euclid-type spaces. The primitives have by this method found **a spatial position**, actually 2 of such.

c. the paths in the habitat

The spatial grid itself is built up by ***cycles***. Cycles appear as constituents of a procedure that is known as 'reordering'. The topography of the habitat is rooted in the properties of the *similarity-related* cycles, the content of the habitat is determined by the cycles that are present due to the *diversities* among the members.

Using these building modules of the model, **we observe:**

a. agglomerations

The collection being subject to more than two reorders at the same time, it is inevitable, that on specific spatial coordinates agglomerations of amounts come into existence. Like in traffic simulation: if sufficiently many cars are on the road traveling to sufficiently diverse and similar goals, pileups are unavoidable. The pileups come in ***types***. (Like traffic jams can be differentiated on the number of cars involved, their momentum, composition of types of cars, etc.) As these agglomerations appear out of nowhere in our perception, although they had always been there, only we have not knowingly recognized them, like archetypes, we propose to name the various types of material agglomerations due to periodic changes that elementary logical symbols undergo, ***logical archetypes***. The concept of logical archetypes should correspond to the concept of ***chemical elements***.

b. regularity

It is in the nature of periodic changes that they are a sequence of recurring states. The regularity allows a concept of ***evolutionary adaptability***. Whether self-organization is an immanent property of Nature, can be decided by investigating the following hypothesis:

Any random arrangement of the logical primitives is closer (more similar, easier to transform into) to an order based on properties of logical primitives, than to a different arrangement which is not based on properties of logical primitives.

In our man-made habitat, it is easy to see that periodic changes impose a *rhythm* on the assembly.

c. predictability

Recurring states can be predicted.

A state can be identified/predicted by the appearance of such members of cycles which are known to be predecessors in cycles containing the constituent of the state to be predicted/identified. Thy cycles bring forth signs of change.

d. alternatives

The whole model is – but for the logical archetypes – a collection of alternatives. The alternatives have a numeric property of their extent of being otherwise. It is a huge accounting challenge, but it is theoretically possible to classify and tabulate all possible alternatives. Of this lot, only such need to be used, which are included in all ways of describing the collection: as a web of relations similarity-based, as a web of relations diversity based, as a content that is expressible in both descriptions, as a process that creates and reduces <so-and-so much> discongruence.

Closing remarks

The present paper offers an invitation. Why don't you explore the completely new landscape, which gives a formidable background for quite many of established concepts from Physics and Chemistry?

The ideas presented here make the impression of being no less than revolutionary. This reflects less on the novelty of the new algorithms, but rather on the conservatism and rigidity required in the technical sciences. There is nothing extraordinary in the idea to assemble an etalon cohort of simple symbols and subject them to periodic changes. The question is what one reads out of the observations and whether one dares believe that one has proceeded logically and rationally while attaining the observed results.

In the context of software, basic design principles of, one can rephrase the general concept in terms of search strategies. We can address 1 member of **N** by sequential

methods or by methods of queries (contemporary searches, based on categories of commutative symbols). It has been shown (cf A242615) that query-based searches are ~ 350 % **more efficient** relative to sequential searches if the batch of information carrying vehicles numbers ~ 66.

Using the cycles as indexes of relationships, one finds a **lexicon** which is by its immanent features **already cross-referenced**.

Summary

It is unavoidable for the Reader, if they wish to continue investigations into cohesive counting, that they generate, program, set up their own *tautomat*. This extensive system of tables is comparable to the trigonometric tables, inasmuch as relations among numeric values are explicated. Instructions on how to build your own tautomat are eg in *Picturing Order* [Javorszky, K., 2018]. The back-of-the-envelope laboratory prototype at www.tautomat.com is also available.

For a research institute which has insight, resources, circumstances to set up a public version of The Tautomat (like the Online Encyclopedia of Integer Sentences, a common resource), the advantages would be formidable. If you host the etalon laboratory animal, you will be the first to know, which variations and mutations that system of implications will more create, aside the few mentioned here. This could be a massive competition advantage. One cannot lose a bet on the idea that ordering and reordering simple logical tokens will turn up typical patterns and that these manifold typical patterns will be of interest to Physics, Chemistry and Physiology.

Bibliography

- [Javorszky K., 2016] Javorszky Karl. Natural Orders De ordinibus naturalibus, Morawa, Wien 2016, ISBN: 978-3-99057-137-8 <https://www.facultas.at/item/32109801>
- [Javorszky K., 2018] Javorszky Karl. Picturing Order, itsrcp18_txt46, Cracow, 2018 https://www.researchgate.net/publication/329275837_Picturing_Order

Authors' Information



Karl Javorszky – *retired clinical psychologist; Loeblichgasse 13/16, A-1090 Wien, e-mail: karl.javorszky@gmail.com*

Scientific Interest: Memory, how registering and reconstructing information is actually done.

SECTION: SOFTWARE ENGINEERING APPLICATION DOMAINS**APPLICATION OF SOFT SETS TO ASSESSMENT OF PROBLEM SOLVING SKILLS****Michael Gr. Voskoglou**

Abstract: *Volumes of research have been written about problem solving, which is one of the most important components of the human cognition that affects the progress of the human society for ages. In this work soft sets are used as means for obtaining a qualitative assessment of problem solving skills. The concept of soft set is a generalization of Zadeh's fuzzy sets introduced by Molodstov in 1999 as a new mathematical tool for dealing with the existing in real world uncertainty in a parametric manner.*

Keywords: *Soft Sets, Fuzzy Sets, Problem Solving, Assessment Methods.*

ITHEA Keywords: *G Mathematics and Computing.*

Conference and topic: *International Software Engineering Conference, National Aviation University, Kyiv, Ukraine, 2022*

Introduction

Until the middle of the 1960's *probability theory* used to be the unique tool in hands of the experts for dealing with the existing in real life and science situations of uncertainty. Probability, however, based on the principles of the bivalent logic, has been proved sufficient for tackling problems of uncertainty connected only to randomness, but not those connected to imprecision or to incomplete information of the given data.

The *fuzzy set theory*, introduced by Zadeh in 1965 [Zadeh, 1965], and the connected to it infinite-valued in the interval $[0, 1]$ *fuzzy logic* gave to scientists the opportunity to model under conditions of uncertainty which are vague or not precisely defined, thus succeeding to mathematically solve problems whose statements are expressed in the natural language. Through fuzzy logic the fuzzy terminology is translated by algorithmic procedures into numerical values, operations are performed upon those values and the outcomes are returned into natural language statements in a reliable manner.

Fuzzy systems are considered to be part of the wider class of *Soft Computing*, also including *probabilistic reasoning* and *neural networks*, which are based on the function of biological networks [Papilinski, 2005]. One may say that neural networks and fuzzy systems try to emulate the operation of the human brain. The former concentrate on the structure of the human mind, i.e. the “hardware”, and the latter concentrate on the “software” emulating human reasoning.

Let U be the universal set of the discourse. It is recalled that a fuzzy set A on U is defined with the help of its *membership function* $m: U \rightarrow [0,1]$ as the set of the ordered pairs

$$A = \{(x, m(x)): x \in U\} \quad (1)$$

The real number $m(x)$ is called the *membership degree* of x in A . The greater is $m(x)$, the more x satisfies the characteristic property of A . Many authors, for reasons of simplicity, identify a fuzzy set with its membership function.

A crisp subset A of U is a fuzzy set on U with membership function taking the values $m(x)=1$ if x belongs to A and 0 otherwise. In other words, the concept of fuzzy set is an extension of the concept of the ordinary sets.

It is of worth noting that there is not any exact rule for defining the membership function of a fuzzy set. The methods used for this purpose are usually empirical or statistical and the definition is not unique depending on the personal goals of the observer. The only restriction about it is to be compatible to the common logic; otherwise the resulting fuzzy set does not give a reliable description of the corresponding real situation.

For example, defining the fuzzy set of the young people of a country one could consider as young all those being less than 30 years old and another all those being less than 40 years old. As a result they assign different membership degrees to people with ages below those two upper bounds.

A lot of research has been carried out during the last 60 years for improving and extending the fuzzy set theory on the purpose of tackling more effectively the existing uncertainty in problems of science, technology and everyday life. Various generalizations of the concept of fuzzy set and relative theories have been developed like the type-2

fuzzy set, the intuitionistic fuzzy set, the neutrosophic set, the rough set, the grey system theory, etc. [Voskoglou, 2019b].

In 1999 Dmtri Molodstov, Professor of the Computing Center of the Russian Academy of Sciences in Moscow, proposed the notion of *soft set* as a new mathematical tool for dealing with the uncertainty in a parametric manner [Molodstov, 1999].

Let E be a set of parameters, let A be a subset of E and let f be a mapping of A into the set $\Delta(U)$ of all subsets of U . Then the soft set on U connected to A , denoted by (f, A) , is defined as the set of the ordered pairs

$$(f, A) = \{(e, f(e)) : e \in A\} \quad (2)$$

In other words, a soft set is a parametrized family of subsets of U . Intuitively, it is "soft" because the boundary of the set depends on the parameters. For each e in A , $f(e)$ is called the *value set* of e in (f, A) , while f is called the *approximation function* of (f, A) .

For example, let $U = \{H_1, H_2, H_3\}$ be a set of houses and let $E = \{e_1, e_2, e_3\}$ be the set of the parameters $e_1 = \text{cheap}$, $e_2 = \text{beautiful}$ and $e_3 = \text{expensive}$. Let us further assume that H_1, H_2 are the cheap and H_2, H_3 are the beautiful houses. Set $A = \{e_1, e_2\}$, then a mapping $f: A \rightarrow \Delta(U)$ is defined by $f(e_1) = \{H_1, H_2\}$, $f(e_2) = \{H_2, H_3\}$. Therefore, the soft set (f, A) on U is the set of the ordered pairs

$$(f, A) = \{(e_1, \{H_1, H_2\}), (e_2, \{H_2, H_3\})\} \quad (3)$$

A fuzzy set on U with membership function $y = m(x)$ is a soft set on U of the form $(f, [0, 1])$, where $f(\alpha) = \{x \in U : m(x) \geq \alpha\}$ is the corresponding α -cut of the fuzzy set, for each α in $[0, 1]$. The concept of soft set is, therefore, a generalization of the concept of fuzzy set.

An important advantage of soft sets is that, by using the set of parameters E , they pass through the existing difficulty of defining properly the membership function of a fuzzy set.

The theory of soft sets has found many and important applications to several sectors of the human activity like decision making, parameter reduction, data clustering and data dealing with incompleteness, etc. One of the most important steps for the theory of soft sets was to define mappings on soft sets, which was achieved by A. Kharal and B.

Ahmad and was applied to the problem of medical diagnosis in medical expert systems [Kharal & Ahmad, 2011]. But fuzzy mathematics has also significantly developed at the theoretical level providing important insights even into branches of classical mathematics like algebra, analysis, geometry, topology etc.

Problem Solving in Mathematics Education

Volumes of research have been written about problem solving, which is one of the most important components of the human cognition that affects the progress of the human society for ages. In [Voskoglou, 2011] we have examined the role of the problem in learning mathematics and we have attempted a review of the evolution of research on PS in mathematics education from the time of Polya until today.

Polya laid during the 50's and 60's the foundation for exploration in heuristics for Problem Solving (PS) since he was the first who described them in a way that they could be taught. The failure of the introduction of the "New Mathematics" in school education placed the attention of specialists during the 80's on the use of the problem as a tool and motive to teach and understand better mathematics. A framework was created describing the PS process and reasons for success or failure in PS, which was depicted in Schoenfeld's *Expert Performance Model (EPM)* for PS [Schoenfeld, 1980].

While early work on PS focused on describing the PS process, more recent investigations during the 2000's focused on identifying attributes of the problem solver that contribute to successful PS. Carlson and Bloom drawing from the large amount of literature related to PS developed a broad taxonomy to characterize major PS attributes that have been identified as relevant to PS success. This taxonomy gave genesis to their *Multidimensional PS Framework (MPSF)* [Carlson & Bloom, 2005], which includes the following steps: *Orientation, Planning, Executing* and *Checking*. It was observed that, when contemplating various solution approaches during the planning step of the PS process, the solvers were at times engaged in a *conjecture-imagine-evaluate* sub-cycle. It is of worth noting that a careful inspection of the two PS models shows that the steps of MPSF are in one-to-one correspondence to the steps of Schoenfeld's EPM. However,

there exists a basic qualitative difference between the two models: While in MPSF the emphasis is turned to the solver's behaviour and required attributes, the EPM is oriented towards the PS process itself describing the proper heuristic strategies that may be used at each step of the PS process.

Use of Soft Sets for Assessing Problem Solving Skills

Quality is a desirable characteristic of all human activities. This makes assessment one of the most important components of the processes connected to the application of those activities. The present author has developed in earlier works several methods for assessing human-machine performance under fuzzy conditions, including the measurement of uncertainty in fuzzy systems, the use of the Center of Gravity (COG) defuzzification technique, the use of fuzzy or grey numbers, etc. All these methods have been reviewed in [Voskoglou, 2019a]. Here a method using soft sets is developed for the assessment of PS skills in a parametric manner. Such kind of methods are very useful when the assessment has qualitative rather than quantitative characteristics.

Assume that a mathematics teacher wants to assess the PS skills of a set $U = \{S_1, S_2, \dots, S_n\}$ of students. Let $E = \{e_1, e_2, e_3, e_4, e_5\}$ be the set of parameters e_1 =excellent, e_2 =very good, e_3 =good, e_4 =mediocre and e_5 =unsatisfactory. Assume further that the first four students demonstrated excellent performance, the next five very good, the following 7 good, the next eight mediocre and the rest of them unsatisfactory performance. Let f be the map assigning to each parameter of E the subset of students whose performance was assessed by this parameter. Then, the overall student performance is represented mathematically by the soft set

$$(f, E) = \{(e_1, \{S_1, S_2, S_3\}), (e_2, \{S_4, S_5, \dots, S_8\}), (e_3, \{S_9, S_{10}, \dots, S_{15}\}), (e_4, \{S_{16}, S_{17}, \dots, S_{23}\}), (e_5, \{S_{24}, S_{25}, \dots, S_n\}) \quad (4)$$

The use of soft sets enables also the representation of each student's individual performance at each step of the PS process. In fact, let S_1 =orientation, S_2 =planning, S_3 =conjecture-imagine-evaluate, S_4 =executing and S_5 =checking be the steps of the previously mentioned MPSF. Set $V = \{S_1, S_2, S_3, S_4, S_5\}$, consider a particular student of

U and define a map $f: E \rightarrow \Delta(V)$ assigning to each parameter of E the subset of V consisting of the steps of the PS process assessed by this parameter with respect to the chosen student. For example, the soft set

$$(f, E) = \{(e_1, \{S_1, S_3\}), (e_2, \{S_5\}), (e_3, \{S_4\}), (e_4, \{S_2\}), (e_5, \emptyset)\} \quad (5)$$

represents the profile of a student who demonstrated excellent performance at the steps of orientation and conjecture-imagine-evaluate, very good performance at the step of checking, good performance at the step of executing and mediocre performance at the step of planning (he/she faced difficulties, but he/she finally came through).

Conclusion

The discussion performed in this work leads to the conclusion that soft sets offer a potential tool for a qualitative assessment of human-machine performance in a parametric manner. Due to its general texture, this soft-set assessment method could be also applied to a variety of other cases for assessing human and/or machine activities and this is an interesting subject for future research.

References

- [Carlson & Bloom, 2005], Carlson, M.P., Bloom, I., The cyclic nature of problem solving: An emergent multidimensional problem-solving framework, *Educational Studies in Mathematics*, 58, 47-75, 2005.
- [Kharal & Ahmad, 2011], Kharal, A. & Ahmad, B., Mappings on Soft Classes, *New Mathematics and Natural Computation*, 7(3), 471-481, 2011.
- [Molodtsov, 1999], Molodtsov, D., Soft Set Theory-First Results, *Computers and Mathematics with Applications*, 37(4-5), 19-31, 1999.
- [Papilinski, 2005], Papilinski, A.P., *Neuro-Fuzzy Computing*, Lecture Notes, Monash University, Australia, 2005.
- [Schoenfeld, 1980], Schoenfeld, A., Teaching Problem Solving Skills, *Amer. Math. Monthly*, 87, 794-805, 1980.
- [Voskoglou, 2011], Voskoglou, M. Gr., Problem-Solving from Polya to Nowadays: A Review and Future Perspectives. In *Progress in Education*, R.V. Nata (Ed.), Nova Science Publishers: NY, Vol. 22, Chapter 4, 65-82, 2011.
- [Voskoglou, 2019a], Voskoglou, M.Gr., Assessing Human-Machine Performance under Fuzzy Conditions, *Mathematics*, 7, article 230, 2019.
- [Voskoglou, 2019b], Voskoglou, M.Gr., Generalizations of Fuzzy Sets and Relative

Theories, in Voskoglou, M.Gr. (Ed.), *An Essential Guide to Fuzzy Systems*, pp. 345-368, Nova Science Publishers, NY, USA, 2019.

[Zadeh, 1965], Zadeh, L.A., *Fuzzy Sets, Information and Control*, 8, 338-353, 1965.

Authors' Information



Michael Gr. Voskoglou – *Graduate Technological Educational Institute (TEI) of Western Greece; Professor Emeritus of Mathematical Sciences; 26334 Patras, Greece e-mail: mvoskoglou@gmail.com*

Major Fields of Scientific Research: *Algebra, Artificial Intelligence, Fuzzy Logic, Markov Chains, Mathematics Education*

ABOUT THE ISSUE OF JOINT PROCESSING OF IMAGES OF THE EARTH'S SURFACE, OBTAINED FROM DIFFERENT TYPES OF MEDIA

Valeriy Pysarenko, Julia Pisarenko, Ivan Varava, Olga Gulchak

Abstract: *The problem of joint processing of images of the Earth's surface obtained from spacecraft and unmanned aerial vehicles, which is associated with the need to combine frames of Earth sections during image preprocessing, is considered.*

Keywords: *spacecraft, unmanned aerial vehicles, Earth's surface*

Conference and topic: *ITHEA ISS SoftEngine2022, Image processing and computer vision*

Introduction

The procedure for remote sensing of the earth (ERS) from a spacecraft is widely used to assess the state of the terrain, for example, to detect the state of vegetation [Shunping Ji and all, 2020] or, very importantly, the consequences of destruction after technological and environmental or military events. But these methods do not provide detailed information about the state of the infrastructure. In this case, there is a need to support technologies for space survey of territories with additional methods of photographing a given surface of a land plot from heights from 10 m to 200 m using telephoto equipment placed on board unmanned aerial vehicles (UAVs).

Thus, the problem arises of joint processing of video and photographic materials obtained by UAV and spacecraft (SC) equipment. Joint processing of photographs or video materials from the spacecraft and UAV boards provides the most comprehensive information about the state of the studied territory of the Earth (Fig.1).

Below is a list of the most typical problems and tasks that arise in the development of methods for such joint processing of images of the earth's surface, including multi-temporal ones (which is sometimes necessary to identify the dynamics of changes in individual parts of the earth's surface).

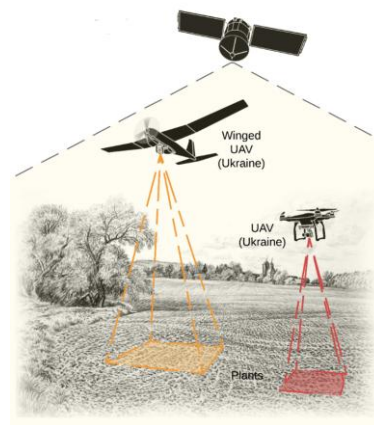


Fig.1: Illustration for the procedure of remote sensing of the earth's surface from various types of media

The solution of these problems occurs during the so-called pre-processing of images, after which it becomes possible to further process the prepared images according to the adopted program [Shunping Ji and all, 2020]:

The problem of compensating for the mismatch of different-time frames of remote sensing images.

The problem of combining images obtained from different aircraft.

The problem of the impact of atmospheric phenomena [A. Huete and H. Liu, 1994].

The problem of filtering noise in the image.

The problem of joint processing of images of the earth's surface with different resolutions.

Image processing will consist of the following steps:

- 1) The array under study (the area from which images have already been taken) with the coordinates (two-dimensional) of each "frame" is considered a template.
- 2) Geometric spatial references, key points, are used to identify the boundaries of each frame. And when the UAV passes the next time after a certain time, a picture of the same area is obtained. The time interval between successive studies of the territory is determined from the characteristics and objectives of the task.

-
-
- 3) The operation of combining the boundaries of these multi-temporal images is carried out (determining the "common" frame for two snapshots). (Images received from an unmanned aerial vehicle (UAV) in different periods of time have discrepancies relative to each other. The same problem arises when processing images from spacecraft is required, namely, an automated search for changes in satellite images at different times. Then it becomes necessary at the stage pre-processing, use methods for eliminating the discrepancy between frames of remote sensing images at different times).
 - 4) Mathematical processing to detect the correlation of these two images is possible after the following preparatory work:
 - a) elimination of errors due to scale discrepancies;
 - b) elimination of errors due to differences in the spatial resolution of images of the same area of the surface;
 - c) elimination of the influence of differences in the angles of inclination of the shooting cameras. Distortions arise when gluing microframes and divergent axes of devices, shooting angles that form multispectral images;
 - d) elimination of the influence of atmospheric changes [A. Huete and H. Liu, 1994];
 - e) creation of a panchromatic image (black and white) of a specific area using multi-zone imaging, if it is decided to use the methods of correlation analysis of panchromatic images. Otherwise, methods of applying the correlation processing of multispectral images can be applied;
 - f) If necessary, orthorectification of images.

Note: not all of the above preparatory work may be needed, it all depends on the specific requirements for monitoring land masses.

Below in fig.2-4 one of the operations of preliminary processing of images of the earth's surface - image alignment - is illustrated for the purpose of the next step of the technique:

comparison of images at different times. Various satellite images of one of the areas were processed using OpenCV, an open source computer vision library.

Fig. 2,3 - satellite images of one area of the territory, made with an interval of several days.



Fig. 2



Fig. 3



Fig. 4: Photo from the UAV superimposed on a satellite photo of the area as an example of the operation of the program for combining two photos using OpenCV (an open source computer vision library available for free use for scientific, academic and commercial purposes)

Conclusion

After combining the images, it becomes possible to separate 2 frames of images at different times for processing them by the methods of correlation analysis [V.T. Bickel and all, 2018] using computer processing programs and identify changes in the state of the terrain [Rizzo and all, 2005, Leprince, S. and all, 2007].

References

- Shunping Ji, Yu Dawen, Chaoyong Shen, Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks., February 2020, Landslides 17(6) DOI:10.1007/s10346-020-01353-2
- A. Huete, and H. Liu, "An error and sensitivity analysis of the atmospheric and soil correcting variants of the NDVI for the MODIS- EOS," IEEE Transactions on Geoscience and Remote Sensing, July 1994, vol. 32, pp. 897-905 [DOI:https://doi.org/10.1109/36.298018].
- V.T. Bickel, A. Manconi, and F. Amann, "Quantitative Assessment of Digital Image Correlation Methods to Detect and Monitor Surface Displacements of Large Slope Instabilities," Remote Sens, 2018, № 10, pp. 865-883 [DOI:https://doi.org/10.3390/rs10060865].
- F. Rizzo, B. Carpentieri, G. Motta, and J.A. Storer, "Low-complexity lossless compression of hyperspectral imagery via linear prediction," IEEE Signal Processing Letters, 2005, vol. 12, no. 2, pp. 138-141 [DOI:https://doi.org/10.1109/LSP.2004.840907].
- S. Leprince, S. Barbot, F. Ayoub, and J.-P. Avouac, "Automatic and Precise Orthorectification, Coregistration, and Subpixel Correlation of Satellite Images, Application to Ground Deformation Measurements," IEEE Transactions on Geoscience and Remote Sensing, 2007, Vol. 45, No. 6, pp. 1529-1558, July 2007 [DOI:https://doi.org/10.1109/TGRS.2006.888937]
-

Author's Information



Valeriy Pysarenko, V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, chief of dept. №265 of Mathematical Problems of Applied Informatics V.M.Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine, e-mail: pisarenkojv@gmail.com

Major Fields of Scientific Research: General theoretical information research, Multi-dimensional information systems



Julia Pisarenko, V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, Senior Researcher lab. №236 of Virtual Environment Systems for Organization of Scientific Research, Kyiv, Ukraine e-mail: jvpisarenko@gmail.com

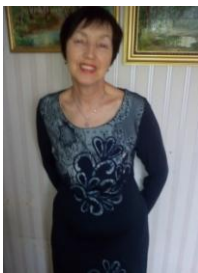
Major Fields of Scientific Research: Multi-dimensional information

systems, robototechnics



Ivan Varava, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Associate Professor at the Department of Automation of Power Processes and Systems Engineering National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” Kyiv, Ukraine, e-mail: ivan.varava@ukr.net

Major Fields of Scientific Research: Multi-dimensional information systems, robototechnics, image processing, programming



Olga Gulchak, V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, engineer of dept. №265 of Mathematical Problems of Applied Informatics V.M.Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine, e-mail: olga-gulchak@ukr.net

Major Fields of Scientific Research: robototechnics, image processing.

PRINCIPLES FOR SELECTING CHARACTERISTIC PEAKS OF SPEECH SIGNAL FOR SPEAKER IDENTIFICATION

Serhii Zybin, Yana Bielozorova

Abstract: *A method for isolating the parameters of the voice signal characteristic of the speaker by selecting the features of the extremums of the surface of the scalogram of the voice signal, built on the basis of wavelet coefficients. The structural model of the speech signal is described, the sequence of mathematical transformations of the information channel of speech communication is considered and generalized, on the basis of which the approach to selection of features of language structures is offered. An algorithm for selecting the characteristics of speech signal structures has been developed.*

Keywords: *voice signal, Morlet wavelet, special structures, fundamental frequency, speaker identification.*

Conference and topic: *SoftEngine 2022; Software Engineering Application Domains - Image Processing and Computer Vision.*

Introduction

The accuracy of speaker identification is based on the correct choice of identification features and the correct presentation of the voice signal. One of the most important characteristics and often determining is the frequency of the fundamental tone. Currently, there is no single approach to determining the fundamental frequency, so effective identification of the announcer is possible only on the basis of a comparison of the frequencies of the fundamental tone, performed on the basis of the same method of determining the frequency of the fundamental tone.

Related works

Nowadays, it is well known that vocalized fragments of a person's speech signal maintain the periodicity of voice oscillations. The speech signal should be considered in the form of pulsating oscillations of the air flow, where the period of repetition of the pulses of the speech signal is called the frequency of the fundamental tone. It has been studied that the fundamental tone determines the structure of the speech signal, and is also the main

parameter of the speech signal [Mallat, 1999]. The intonation contour of a person is a trajectory of changing the melody of the fundamental tone frequency.

Prosody of speech formation, one of the components of which is intonation, significantly distinguishes the acoustic signal of speech from written language. Therefore, the fundamental frequency carries a significant amount of information contained in the speech signal of the person. From the point of view of presenting this information, the process of allocating the fundamental frequency is of independent interest [Mallat, 1999, Muzy et al., 1991]. When calculating the fundamental frequency, it is also necessary to take into account both slow changes in the trajectory of the fundamental frequency, and its rapid changes and the very moments of inclusion/deactivation of the speech signal. However, when constructing systems for analysis, synthesis, compression or identification of a speech signal, it should be borne in mind that the fundamental frequency is usually used as one of the main features needed for improved description of the speech signal. When it is necessary to perform calculations of the fundamental frequency from a real speech signal, it is necessary to pay attention to the following points [Muzy et al., 1991]:

- 1) the ability to work at different noise levels;
- 2) estimation of the fundamental frequency should be performed with minimal error;
- 3) the minimization of error in a wide range of changes in the frequency of the fundamental tone, emotional and other changes in the language of the person, leading to variations in the frequency of the fundamental tone;
- 4) must perform all previous signal transformations;
- 5) effectively divide the signal into vocalized/unvocalized fragments.

To get the maximum amount of information contained in the fundamental frequency loop, you should also pay attention to the additional point when you need to perform instantaneous calculations of the fundamental frequency value.

Materials and methods

If $|Wf(s, u)|$ - function of voice signal description - has no local maxima on a small scale, it is assumed that the function f describing the speech signal is locally smooth and the process of isolating self-similar structures of the speech signal function f can be constructed by determining the maximum values of the functions $|Wf(s, u)|$ on a small scale. It is taken into account that the scale parameters are limited by the parameters of the segmentation of the speech signal and its step.

After the selection of a self-similar structure, the next task is its classification. We use the approach proposed in [Pavlov A.V. et al., 2007, Lardiès J. et al., 2004]. Expression can be written in a similar form

$$\log_2 |Wf(s, u)| \leq \log_2 A + \left(\alpha + \frac{1}{2} \right) \log_2 s \quad (1)$$

Therefore, the smoothness parameters at point v are determined by the slope of the function $\log_2 s$ (and accordingly $\log_2 |Wf(s, u)|$ along the maxima line. The peculiarity of the lines of maxima is its construction on the basis of the points of maxima of the module, which is the curve $s(u)$ in the coordinates (s, u) .

The classification of self-similar structures of the speech signal will be performed using (1) as follows. We introduce the notation $O_v(s, u)$ – as a line of maxima of the wavelet transform of the speech signal converging to the point $u = v$, at $s \rightarrow 0$. For each point v we define the slope $\log_2 O_v(s, u)$ as a function $\log_2 s$ for $s \rightarrow 0$:

$$\log_2 O_v(s, u) = \log_2 A + \left(\alpha' + \frac{1}{2} \right) \log_2 s \quad (2)$$

We will consider that at the point $u = v$ we have a self-similar structure α' .

An effective solution to the problem of classifying the self-similar structure of the speech signal depends on the characteristics of the basic function of the wavelet transform ψ . It was found that performing a wavelet transform allows to obtain a set of wavelet

coefficients that describe the speech signal. The most popular method of presenting the results of wavelet transform is a scalogram, which allows you to visually represent and assess the location of the surface extremums, built on the basis of wavelet coefficients $W(a, b)$.

Of particular interest are the local extrema of the surface coefficients. Theoretically, the analysis of self-similar structures can be performed on the basis of scalogram parameters, but there are a number of statistical functions that allow the evaluation of spectrum characteristics more efficiently.

The general view of such statistical measures of the measure can be represented as

$$M(q, a) = \sum_{l \in L(a)} |W(a, t_l(a))|^q \quad (3)$$

where l – local maximum line, L – a set of lines of maxima of wavelet coefficient modules, $t_l(a)$ – maxima of wavelet coefficients related to lines l scale a .

According to [Lardiès J. et al., 2004], the dependence is executed

$$M(q, a) \sim a^{\tau(q)} \quad (4)$$

where $\tau(q)$ is determined for the value of q by calculating the slope $\ln(M(q, a))$ from $\ln a$, which is called the scaling exponential. Setting the value of q in (3) we obtain the dependence $\tau(q)$. The dependence $\tau(q)$ allows to obtain a multifractal spectrum of the speech signal based on the wavelet transform [Pavlov A.V. et al., 2007, Lardiès J. et al., 2004], which allows to describe the main characteristics of self-similar structures. The following dependence is used to obtain the multifractal spectrum

$$\begin{cases} D(h) = \min_q [qh - \tau(q)] \\ h = \partial \tau / \partial q \end{cases} \quad (5)$$

The stability of this method of obtaining the characteristics of self-similar structures is to use the frequency-time window, which automatically performs averaging operations, as well as to obtain modules of wavelet coefficients in the calculation. It should be noted that from the point of view of the energy approach to the analysis of self-similar structures,

the maxima of the wavelet transform coefficients at different levels of decomposition are the most significant.

The presented approach allows to expand the possibilities of speech signal analysis through the use of fractal and wavelet analysis. Unlike existing methods, this allows the analysis of non-stationary and short-lived signals.

Conclusion

An algorithm for selecting special structures in a speech signal has been developed and an informative feature for linguistic identification of a person who, unlike existing ones, uses the values of wavelet transform coefficients of speech signal on segments where extremes of fundamental frequency correlation are observed. who are responsible for the individuality of the speech signal, and achieve high accuracy of identification. An experiment on the selection of rational parameters of the wavelet transform for the implementation of the method is performed, the coefficients of the wavelet transform adjustment for the frequency bands are obtained.

As an assessment of the effectiveness of the study, a comparative experiment was performed to determine the frequency of the fundamental tone based on the proposed approach and the Pitch method. The option to determine the frequency of the fundamental tone, based on the proposed method, reduces the error of determination in all analyzed variants of the voice signal.

Bibliography

- [Mallat, 1999] Mallat S. A wavelet tour of signal processing, Courant Institute, New York University, 1999, 671 pp.
- [Muzy et al., 1991] Muzy J.F., Bacry E., Arneodo A. Wavelets and multifractal formalism for singular signals: application to turbulence data // Phys. Rev. Lett. 1991. V.67. P.3515–3518.
- [Pavlov A.V. et al., 2007] Pavlov A.V., Anishenko V.S. Multifractal signal analysis based on wavelet transform //Saratov University, 2007. №1. 3-25 pp.
- [Lardiès J. et al., 2004] Lardiès J., Ta M.N., Berthillier M. Modal parameter estimation from output-only data using the wavelet transform, Archive of Applied Mechanics, Vol. 73. 2004. 718-733 pp.

Authors' Information



Serhii Zybin – DSc, Professor, Head of Software Engineering Department, National Aviation University, Kyiv, Ukraine.

e-mail: zysv@ukr.net

Major Fields of Scientific Research: game theory, cloud computing, cyberspace, distributed nets, software engineering, information security



Yana Bielozorova – PhD, Associate Professor of Software Engineering Department, National Aviation University, Kyiv, Ukraine.

e-mail: bryukhanova.ya@gmail.com

Major Fields of Scientific Research: Speech Recognition Models, Wavelet analysis, Software Architecture

ABOUT THE ORGANIZATION OF REGIONAL SITUATIONAL CENTERS OF THE INTELLECTUAL SYSTEM “CONTROL_TEA” WITH THE USE OF UAVS

**Julia Pisarenko, Katerina Melkumian, Alexander Kryachek, Olga Gavriilyuk,
Alexander Koval**

Abstract: *The basics of the principles of creation and filling of the technopark of unmanned aerial vehicles (UAV) are offered. The business process of UAV registration in the technopark of the situation center robots is described. The use of attributes (tags) to the UAV will increase the efficiency of a suitable device selection with a complete set that meets the task. It is proposed to organize regional situational centers, which will coordinate the work of state systems, urban life support systems, private services on the one hand and UAV on the other as executors. An example of automatic UAV selection for operative solution of the set task is given. The system should be based on the classification of the UAV during registration and on the basis of the UAV compliance function to the specific task, key features of the UAV, the time required to complete the task, the distance to the task and the policy to involve the UAV in such tasks to select those devices, who can quickly perform tasks. The function of classification of a specific UAV by tasks is offered. The function takes into account the configuration of the UAV. For each task, experts and automation select weights that indicate how large the impact of certain typical modules to perform a particular task, and for each UAV in the database stores information about whether the UAV takes into account a certain characteristic, or there is a corresponding module. To update the approach to vocational training today, the task of creating interactive mock-up training systems in real physical space is one of the main ones. The composition of the basic laboratory stand, its complete set and control system of the mobile modular robot which promotes mastering of basic functions of the UAV or mobile work of unearthy base are described.*

Keywords: *UAV, situational center, traffic control, professional education.*

Conference and topic: *International Conference on Software Engineering 2022.*

Introduction

Obviously, the city of the future is a smart city, a city that in real time controls most of the processes that take place in its bowels automatically and does it optimally, a city where security is raised to a whole new level, and access to information, the benefits of civilization, city services, any necessary human things as simple as possible. And how

will this city regulate the work of rapidly developing UAVs at the moment? After all, it is obvious that with the increase in the number of UAVs, if they begin to take over the routine functions of employees, we need a management system that will ensure safety and minimal accidents, regulate their movement, pave special air routes, simplify UAVs, standardize them, harmoniously integrates UAVs into the structure of the city, puts them at the service of the common good.

Formulation of the problem

Existing issues:

- the need to certify the movement of UAVs [Melkumian et al, 2021] in order to regulate it and prevent unauthorized actions. As a result, it is necessary to build air routes with the formalization of their purpose, taking into account the adaptive management of traffic dynamics;
- the ability to automatically switch remote control in case of emergency: loss of signal from the UAV at the base [Pisarenko and Melkumyan, 2019], the actual transfer of UAV data about its pre-emergency condition, UAV interception, interference/protocol breakage, radio weather interference;
- automated data collection [Pisarenko and Melkumyan, 2014] in the monitoring process for learning neural networks and after gaining new knowledge for further forecasting, passive collection of information for specific purposes;

selection of the communication channel and data transmission protocol between the triad "ground station information storage" + "UAV board" + "decision maker".

Topicality

Let's take a closer look at the issues that we are already slowly facing in the use of UAVs. There is no need for long explanations of the fact that an unmanned vehicle equipped with a good camera and image transmission equipment hanging in front of someone's room is both bad and illegal, that an unmanned carrier with a heavy parcel on board should probably not carry it over kindergarten or over crowds, because in the

event of an accident, its consequences can be very serious. Therefore, UAVs in the transition to their widespread use require mandatory regulation of their movement, certification, accurate construction of routes and informing them about the services of a smart city in real time. Certified UAVs are likely to be able to quickly transfer control to a specialist or automated city system to deal with a variety of emergencies and assist city services. UAVs that are constantly scurrying here and there are likely to have to do some useful things for the city, directly serve the common good, such as passively or actively collecting data for monitoring, looking for people, garbage, something unusual that needs intervention, find something which can be difficult to find with conventional stationary cameras, to provide communication. And in fact, it would be strange not to use the capabilities of flying cameras that capture the city from new angles. The use of communication channels for UAV control, various protocols, special software must be clearly regulated, the components of a smart city must work as a coherent mechanism and not interfere with each other.

To ensure the solution of the described requirements, the authors of the study propose to create a network of regional situational centers (RCC), and, in fact, ground stations covering a certain area, which provide comprehensive control of UAV traffic. These stations should be subordinated to the Central Situation Center (CSC), which will house the decision-making center [Pysarenko et al, 2020] and the main information repository.

Bibliography

- [Melkumian et al, 2021] Melkumian K., Pisarenko J., Koval A. Organization of Regional Situational Centers of the Intelligent System «CONTROL_TEA» using UAVs // 2021 IEEE 6th International Conference on Actual Problems of Unmanned Aerial Vehicles Development, APUAVD-2021 Proceedings, 2021, P. 37-40. DOI: [10.1109/APUAVD53804.2021.9615416](https://doi.org/10.1109/APUAVD53804.2021.9615416)
- [Pisarenko and Melkumyan, 2019] Pisarenko J., Melkumyan E. The Structure of the Information Storage «CONTROL_TEA» for UAV Applications // 2019 IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments, APUAVD-2019 Proceedings, 2019, P. 274–277. DOI: [10.1109/APUAVD47061.2019.8943938](https://doi.org/10.1109/APUAVD47061.2019.8943938)

-
- [Pisarenko and Melkumyan, 2014] J. Pisarenko, E. Mekumyan. Some components of intellectual system «Control_TEP» // Komp'yuterni zasobi, merezhi ta systemy. – 2014. – № 13. – P. 165-174. <http://dspace.nbuiv.gov.ua/handle/123456789/84843>
- [Pysarenko et al, 2020] Pysarenko V., Gulchak O., Pisarenko J. Technology for Improve the Safety of Ships from Methane Emissions Using UAVs // 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control, MSNMC-2020 Proceedings, 2020, P. 93–96. DOI: [10.1109/MSNMC50359.2020.9255661](https://doi.org/10.1109/MSNMC50359.2020.9255661)
-

Authors' Information



Julia Pisarenko – V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine; PhD, Senior researcher, Laboratory №236 of Virtual Environment Systems for Organization of Scientific Research

E-mail: jvpisarenko@gmail.com

Major Fields of Scientific Research: Multidimensional information systems, robotics



Kateryna Melkumian – National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», faculty of Informatics and Computer Engineering; PhD, Associate Professor of the Department

e-mail: e.melkumyan@ukr.net

Major Fields of Scientific Research: Multidimensional information systems, robotics



Alexander Kryachek – V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, dept. №265 of Mathematical Problems of Applied Informatics; National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», faculty of heat power engineer; PhD, Senior researcher

e-mail: alexandrkiachok@gmail.com

Major Fields of Scientific Research: Multidimensional information systems, robotics



Olga Gavrilyuk – V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, dept. №265 of Mathematical Problems of Applied Informatics; PhD, Senior researcher

e-mail: jvpisarenko@ukr.net

Major Fields of Scientific Research: Multidimensional information systems, robotics



Alexander Koval – V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, dept. №265 of Mathematical Problems of Applied Informatics; National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», faculty of Informatics and Computer Engineering; assistant, Junior Researcher

e-mail: t70000s@gmail.com

Major Fields of Scientific Research: Multidimensional information systems, robotics

PHOTO SHARPNESS EVALUATION SYSTEM DESIGN

Oleksandr Dolhanenko, Mariya Shirokopetleva, Natalya Lesna

Abstract: *The problem of image sharpness evaluation in photos of people was addressed. General image sharpness detection algorithms fail to produce accurate results when dealing with naturally blurred background due to the shallow depth of field effect. An alternative and optimized way of image processing flow is required to evaluate photos of people. The approach to designing a system for evaluating the sharpness of photos is given. The designed system uses OpenCV as the main source of sharpness detection results using different methods. A sequence diagram representing the processing sequence design and technologies to be used for development are proposed. The basic steps required for determining the general image sharpness were explained.*

Keywords: *Software engineering, design, system, diagram, sharpness, photo.*

ITHEA Keywords: *I.4.8 Scene Analysis, I.4.9 Applications, D.2.10 Design*

Conference and topic: *17th International Software Engineering Conference SoftEngine 2022. Software Designing.*

Introduction

Being a professional photograph means spending hours upon hours analyzing the resulting thousands of images after a photo shoot. Multiple images can be very similar from the first glance, but differ in lighting and sharpness if looked at close up. Such work of comparing image quality among many similar images is very tiring and needs to be somewhat automated, giving suggestions to the user as to the sharpness rating of each image.

Of course, many factors and image characteristics influence the photographers decision for the final image selection, but most of them (lighting, composition, contrast) can be adjusted relatively easily and "cheaply" (sacrificing not much of the resulting image quality), whereas the image sharpness adjustments are quite "expensive" and usually are not preferred.

Image sharpness evaluation is a problem that can be straight forward on the one hand and extremely complicated on the other. There are different methods for evaluating

overall image sharpens, including mathematical models, training neural networks and more.

The purpose of this work is to analyze the technical capabilities of different platforms and algorithms for evaluating image sharpness.

Basic sharpness evaluation using OpenCV

One of the methods for evaluating image sharpness is using the Variance of the Laplacian filter, applying a kernel to the image row by row and essentially evaluating the level of contrast between the kernel start and end [O. van Zwanenberg et al, 2019]. This method of gradient evaluation can be defined with the following function (see formula 1).

$$\Delta f(x,y) = \text{div}(\text{grad}(f)) \quad (1)$$

To implement such pixel gradient analysis, a software library called "OpenCV" can be used. It contains all necessary means of setting up the algorithm and interpreting the results. As stated on the official website, OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library.

OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Multiple languages can be used with OpenCV, namely Python and Java. Comparing the speed of execution and selecting the better alternative among the two is a topic for further research. For now, the Kotlin programming language (which is an extension language for the JVM platform) will be used. An example of using OpenCV for calculating the variance of the Laplacian:

```
var imageMat = Highgui.imread(image.absolutePath,  
Highgui.CV_LOAD_IMAGE_GRAYSCALE)  
val dst = Mat()  
Imgproc.Laplacian(imageMat, dst, 3)  
val median = MatOfDouble()  
val std = MatOfDouble() Core.meanStdDev(destination, median, std)
```

```
val variance = Math.pow(std.get(0, 0).get(0), 2.0)
```

As seen from the example, the library operates with matrices representing the original image, loaded in grayscale format. The result of such calculation is a floating number which can not be interpreted without a reference point.

Evaluating sharpness of specific image areas

An image taken with a professional camera will most likely have a shallow depth of field, meaning the foreground will be in focus and sharp and the background will be very blurry, due to the open aperture of the camera lens. This effect can cause major inconsistency when evaluating the overall image sharpness, as most of the background would produce low results, which would lead to the image being classified as blurry, when in fact the subject is in critical sharp focus Nuutinen M. at all., 2012], [Wang Z at all., 2004].

The setting of the camera which influences the background blurriness effect is called the aperture. The aperture is most likely to be set to the lowest value possible during a professional photoshoot, to let as much light in as possible and create background separation, or DOF (Depth of Field). However, this does create massive naturally blurry areas which need to be excluded by the sharpness evaluation system.

To solve this issue a system needs to be designed which takes into account the possibility of background being blurry and evaluates only the areas of the image which were intended to be sharp and in focus.

The developed system will be able to detect faces in the frame and analyze the distances to each face, selecting the ones that were intended to be in focus and analyzing only them for the resulting sharpness level.

Sharpness evaluation system design

To successfully evaluate the image sharpness, a system must have a modular design, allowing easy replacement with new implementations and upgrades along the way. The

sequence diagram below (see figure 1) demonstrates the steps required to process an image and output a single floating number result.

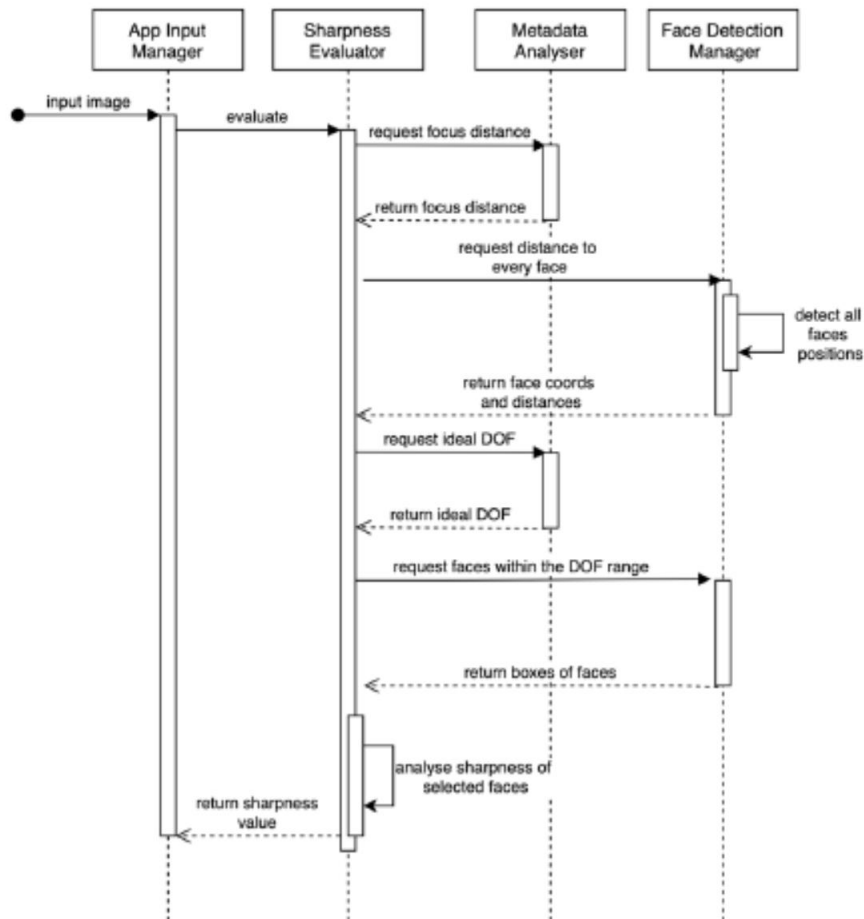


Figure 1: Evaluation Sequence Diagram

The system takes into account the photo metadata and relies on mathematical models to calculate the distances to each face visible in the frame. Having received this information, the sharpness evaluator module performs the analysis, applying the Variance of the Laplacian filter. The result is a value representing the average sharpness of faces in the frame which were intended to be in focus.

Conclusion

In this paper the problem of sharpness evaluation in photos of people was addressed. The issues with overall image analyses were stated and an optimized way of evaluation was proposed. The developed system design was presented via the sequence diagram.

The diagram shows all general steps for sharpness evaluation of areas that were intended to be in focus on the photo.

Acknowledgements

The paper is published with partial support by the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Bibliography

- [O. van Zwanenberg at all., 2019]. O. van Zwanenberg, S. Triantaphillidou, R. Jenkin and A. Psarrou. Edge Detection Techniques for Quantifying Spatial Imaging System Performance and Image Quality. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 1871-1879, doi: 10.1109/CVPRW.2019.00238.
- [Nuutinen M. at all., 2012] Nuutinen M., Orenius O., Saamanen T., Oittinen P. (2012). «A framework for measuring sharpness in natural images captured by digital cameras based on reference image and local areas», EURASIP Journal on Image and Video Processing, P. 4, doi: <https://doi.org/doi:10.1186/1687-5281-2012-8>.
- [Wang Z at all., 2004] Wang, Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. (2004). «Image Quality Assessment: From Error Visibility to Structural Similarity», IEEE Transactions on Image Processing, 13(4), P. 601, doi: <https://doi.org/doi:10.1109/tip.2003.819861>.

Authors' Information



Oleksandr Dolhanenko – Student at Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: oleksandr.dolhanenko@nure.ua

Major Fields of Scientific Research: Information systems, Software Architecture, Image Processing



Mariya Shirokopetleva – Senior lecturer of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: marija.shirokopetleva@nure.ua

Major Fields of Scientific Research: Artificial intelligence systems, Training systems, Data storage systems, Information systems



Natalya Lesna – *Professor of Software Engineering Department, Candidate of Technical Sciences (Ph.D.), Professor.*

Major Fields of Scientific Research: *Information systems, methods and models of distance learning*

REPRODUCING SOUND WAVES WITH AN ALTERNATIVE DIGITAL-TO-ANALOGUE CONVERTER AND EMITTERS

Stanislav Danylenko, Oleksander Vechur, Mariya Shirokopetleva

Abstract: *This paper considers the development of software-hardware complex that implements an alternative method of reproducing sound by air flows. This is achieved through the operation of fans, which are placed in a special way and are controlled by the supply of voltage by means of pulse-width modulation based on digital-to-analog conversion of data obtained from spectral analysis. The paper describes the details of the implementation of the complex: architecture, hardware components and their interaction, technologies used for software implementation and algorithm. It also describes the conduct of experiments with human participation and with the participation of another software system and presents their results. Based on them, conclusions are made and the next steps that need to be taken to achieve the goal are described.*

Keywords: *Software-hardware complex, digital-to-analog conversion, Fourier transform, spectrum analysis, pulse-width modulation.*

ITHEA Keywords: *C.3 Special-purpose and application-based systems, C.5.0 General, D.2.11 Software Architectures, H.5.5 Sound and Music Computing, J.7 Computers in other systems.*

Conference and topic: *Software Engineering Application Domains: Real-Time Systems.*

Introduction

Mankind has thoroughly studied sound and its characteristics and learned to use it for their purposes. However, there is a fairly large group of people from all over the world who have hearing problems and cannot get all the information from the outside world.

Existing methods of alternative sound transmission are almost always reduced to the use of hearing aids, but this is not always possible, so the challenge is to study an alternative method of reproduction and transmission of sound waves.

We are introducing a method of transmitting sound waves by means of air flows created by fans, which interact with human skin and allow to understand what the original sound was.

To do this, a test software-hardware complex with all the necessary software was created and experiments were conducted with the participation of other software systems and people.

Related work

In the last century, there was an attempt to develop an alternative way of transmitting sound waves directly to the human brain. This study was based on observations obtained during the war, but this method has not developed significantly and using this way it was possible to transfer only information about numbers from 1 to 10 [Frey, 1962].

Piezoelectric materials, on the basis of which hearing aids are built, are also widely studied to improve their efficiency [Pérez-Bellido et al, 2017].

Task and challenges

There are tasks in designing the architecture of the software-hardware complex, choosing the necessary hardware components: hardware platform, fans, computer stations and communication channels between them.

There is also the task of creating software for computer stations and hardware platforms that will allow you to receive information from the outside world and from recorded sounds and reproduce it with by means of fans.

Another task is to create the necessary digital-to-analog converter, which can be used to convert the received audio signal into an analog value of the current for fan rotation.

The body

The client-server architecture was chosen for implementation of the complex. This allowed to distribute the load between software modules and make it possible to place a module that processes information (client) and a module that communicates with the hardware platform (server) on different physical nodes that communicate with each other via HTTP.

The Arduino MEGA 2560 board, Delta Electronics ASB0412MA fans, the speed of which is controlled by Arduino PWM pin, was chosen as the hardware platform. Number of fans - 9. They are placed in one plane in a square shape, 3 vertically and 3 horizontally.

The software is implemented using JavaScript programming languages using the Web Audio API module and the C-shaped Arduino language [Rogers et al, 2021].

Algorithm is based on spectral analysis, namely the use of discrete Fourier transform with subsequent distribution of sound pressure values for frequencies by ranges [Tan and Jiang, 2019]. Each fan corresponds to a certain range of frequency points. They determine the most common value, which is greater than 0. Then it is converted to a scale of 1-10 and added to it the prefix (sequence number) of the fan, on which the value will be reproduced. The value is then passed to the server, where it will be sent to the Arduino by wired connection. On the Arduino, the prefix is detached from the value again, and based on the prefix value, the corresponding fan rotates at the specified speed by means of pulse-width modulation, thus performing digital-to-analog conversion.

Two experiments were performed using the created complex: the first using a software system that will investigate whether the generated signal can be recognized by computer systems; the second was conducted with the participation of humans and showed the possibility of classifying sounds by humans.

The experiment with the software system was based on the use of a decision tree. The model was trained on the basis of data obtained from the complex for 5 classes of sounds. Then attempts were made to classify sounds, based on other input data

In an experiment involving humans, air flows were directed at their faces at distance near 5-7 sm. Five people took part in it. Sounds of 14 classes were used: piano, evening field, table tennis, electric trimmer, internal combustion engine operation, firearm shots, a cappella female vocals, sounds from a chicken coop, milk transfusion in a glass jar, ukulele, traffic jam, kettle whistle, train movement, ambulance siren. All sounds had 3 different versions in duration: 2 s, 5 s and 10 s. A separate experiment was performed for each combination of duration and count of sound classes – 4, 9 and 14. Each participant

had 2 attempts. Then the results were formed, which considered various aspects. Based on them, the resulting tables and figures were created to help interpret the results.

The lower threshold for the percentage of correct identifications for the experiments was set at 75%.

Case study or implementation of results

The experiment with another software system shows a result of 0%, which indicates the impossibility of using this method of recognizing the generated information.

Of the results obtained from human experiment, only one configuration overcame the threshold level – 4 files with a duration of 5 seconds (82,50%). Two more configurations scored close to the threshold: 4 files and sound duration of 2 s and 10 s (75.83%).

Interval sounds such as shot and table tennis sound are more likely to be correctly identified.

The results do not reveal any dependence on a particular participant or attempt to participate in the experiment.

Conclusion

Sets of characteristics were identified for which testing exceeded the lower threshold level. This indicates a positive result and makes it possible to conclude that the method under study is effective and can be used for its intended purpose.

Further work

Since positive results have been obtained for a certain set of file duration characteristics and their count, additional research is needed. The purpose of them is to determine the minimal required number of fans, their location, size and power that allow to exceed the lower threshold level.

Additional research is also needed to test the ability to recognize reproduced information using another software system using more sophisticated methods and models, such as machine learning.

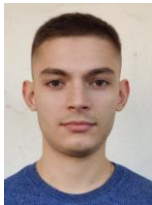
Aknowledgements

The paper is published with partial support by the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Bibliography

- [Frey, 1962] Frey A.H. Human auditory systems response to modulated electromagnetic energy, *Journal of Applied Physiology*, 17(4). 1962. p. 689-692. doi:10.1152/jappl.1962.17.4.689.
- [Pérez-Bellido et al, 2017] Pérez-Bellido A., Barnez K. A., Crommett L. E., Yau J. M. Auditory Frequency Representations in Human Somatosensory Cortex, *Cereb Cortex*, 28. 2017. p. 1–14. doi:10.1093/cercor/bhx255.
- [Rogers et al, 2021] Rogers C., Wilson C., Toy R., Adenot P., Choi H. Web Audio API. 2021. <https://www.w3.org/TR/2021/REC-webaudio-20210617>.
- [Tan and Jiang, 2019] L. Tan, J. Jiang, *Digital Signal Processing (Third Edition)*, Elsevier, 2019.
-

Authors' Information



Stanislav Danylenko – student of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: stanislav.danylenko@nure.ua

Major Fields of Scientific Research: Information systems, Real-time systems, Corporate systems.



Oleksander Vechur – Assoc. Prof. of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: alexander.vechur@nure.ua

Major Fields of Scientific Research: Information systems, Intellectual analysis of the text, Detection of plagiarism.



Mariya Shirokopetleva – Senior lecturer of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: marija.shirokopetleva@nure.ua

Major Fields of Scientific Research: Artificial intelligence systems, Training systems, Data storage systems, Information systems.

TRIP PLANNING SOFTWARE SYSTEM DESIGN

Mariia Kryvoruchko, Mariya Shirokopetleva, Oleksander Vechur

Abstract: *The subject of this article is a software solution design for tourist trips planning. The goal is to design and present an algorithm to address an issue of precise suggestion and the software system design. The following methods were used: linear additive convolution with normalizing factors and weights, combinations, travelling salesman problem. The suggested algorithm uses all the mentioned methods in order to optimize the results. In order to create a maintainable system, a modular design was suggested. The main logic of trip generation and the separation of this logic into modules is presented on a created sequence diagram. The system implementation is defined as a further step.*

Keywords: *Software engineering, design, system, diagram, linear additive convolution, combinations, travelling salesman problem, trip, planning.*

ITHEA Keywords: *I.4.9 Applications, D.2.10 Design.*

Conference and topic: *17th International Software Engineering Conference SoftEngine 2022. Software Designing.*

Introduction

Nowadays people have more abilities to travel then ever had, so now it's only about the desire to see some new places. And one of the main questions here is where to travel. A typical person decides where to go based on experience, types of activities and price. Also, there are a lot of people who enjoy moving around the countries in search of some new experiences but planning such trip logistics can be very difficult. The goal is to implement a service to provide the ability to suggest the user a place or multiple places to visit connected logistically based on the individual preferences. The more precise solution is required – the more preferences should be set by the user.

Evaluating Travel Routs

The first step is rating and choosing countries and their cities based on the preferences set by the user. There are a lot of various criteria to meet: is it necessary to have a sea nearby or maybe mountains are preferred or is it a city-oriented vacation, which transport

is available, activities, climate, etc. Linear additive convolution with normalizing factors and weights should be used to resolve this multicriteria optimization task. In order to meet user's preferences – weights are used. And normalizing factors are required as all the values will be in different units.

The second step is to combine the possible countries and then the chosen countries' cities, having as a result the combinations of cities from different countries to visit.

The third step is to access every combination and directions within it. It is required to consider both – price and distance between the cities. This task is a kind of a travelling salesman problem which involves finding the most profitable route passing through the selected cities at least once, and then returning to the original city.

The last step is the assessment of the prepared combinations using normalizing factors in order to receive the final rate for each option.

Travel Routs Suggestions System Design

In order to create a scalable and easily maintainable system, a modular design must be used. The sequence diagram of processing the routs in separate modules is displayed on figure 1.

The user setting the preferences and asking to generate a route triggers the algorithm execution. At first, the Route Builder uses an Options Rater to rate countries and cities, then the options are combined using an Options Combinator and then the routes are accessed and the best directions within every combination are chosen within a Directions Chooser. As the last step before showing the result to the user, the routs are accessed one more time considering all the previous rates.

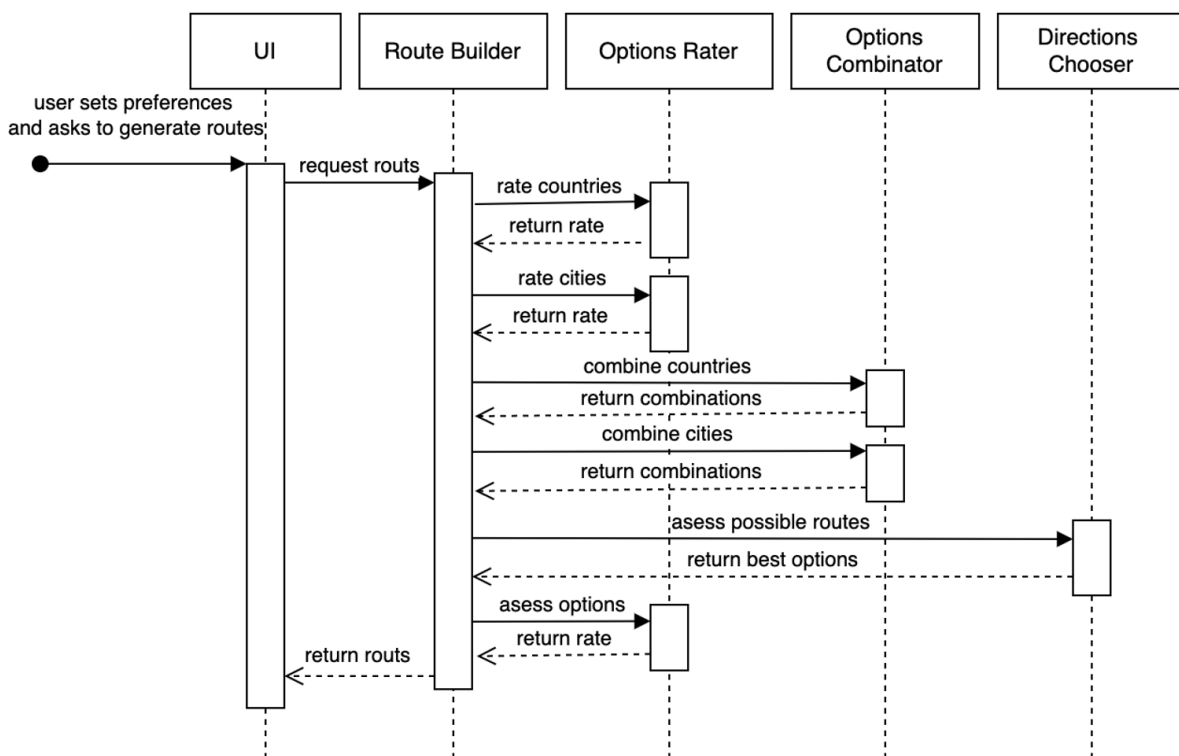


Figure 1: Evaluation Sequence Diagram

In cases where only one destination needs to be chosen – the algorithm will stop before the combining step.

Conclusion

In this paper the problem of the trip suggestion was addressed. The design of the system that needs to be developed was presented in the sequence diagram, where the steps and the logical modules are shown. The further step is the system implementation.

Acknowledgements

The paper is published with partial support by the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua/)

Authors' Information

Mariia Kryvoruchko – Student at Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: mariia.kryvoruchko.cpe@nure.ua

Major Fields of Scientific Research: Information systems, Software Architecture

Mariya Shirokopetleva – Senior lecturer of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: marija.shirokopetleva@nure.ua

Major Fields of Scientific Research: Artificial intelligence systems, Training systems, Data storage systems, Information systems

Oleksander Vechur – Assoc. Prof. of the Department of Software Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. e-mail: alexander.vechur@nure.ua

Major Fields of Scientific Research: Information systems, Intellectual analysis of the text, Detection of plagiarism.

PROCEEDINGS

International Conference on Software Engineering

Technical editor: – Yana Bielozorova

Copyright © 2022 All rights reserved for the publisher and all authors