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

International Conference on Software Engineering

April 13 – 15, 2020

Proceedings

Kyiv 2020

Conference organizers and partners

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

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 lwas 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, Ucraine, 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, Egypt);

Albert Voronin (National Aviation University, Ukraine);

Krassimir Markov (Institute of Information Theories and Applications, Bulgaria);

Oleksii Pysarchuk (National Aviation University, Ukraine);

Olena Chebanyuk (National Aviation University, Ukraine);

Vitalii Velychko (V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine).

Program committee

Antinisca Di Marco (University of L'Aquila, Italy);

Besik Dundua (Tbilisi State University, Georgia);

Eugene Machusky (National Technical University of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Ukraine);

George Totkov (University of Plovdiv, Bulgaria);

Juan Bautista Castellanos Peñuela (Universidad Politécnica de Madrid, Spain);

Krassimira Ivanova (University of Telecommunications and Post, Bulgaria);

Michael Gr. Voskoglou (Graduate Technological Educational Institute of Western Greece School of Technological Applications, Greece);

Mohammad M. Banat (General Chairperson of Mosharaka International Conferences, Jordan);

Mykola Guzii (National Aviation University, Ukraine);

Nagwa Badr (Ain Shams University, Egypt);

Shahnaz N. Shahbazova (Technical University, Azerbaijan);

Vladimir Talalaev (National Aviation University, Ukraine);

Yuh Wen Chen (Da Yeh University, Taiwan).

Organization committee

Tetiana Konrad (National Aviation University, Ukraine);

Valentyna Skalova (National Aviation University, Ukraine):

Yana Byelozorova (National Aviation University, Ukraine).

Table of contents

SECTION: THEORETICAL BASICS OF SOFTWARE ENGINEERING

DOMAIN MODEL TO PERFORM DEEP LEARNING NEURAL NETWORKS TO INVESTIGATE BRAIN TUMORS

Olena Chebanyuk, Abdel-Badeeh M. Salem

***Abstract**: Modern practices of domain model designing are concentrated in practical aspects of software engineering related to Software Product Lines. But implementing approaches of domain models designing in other areas allow to represent visual systematization of relationships between domain entities and information about their internal structure. Paper starts the research of smart health information system analysis and devoted to designing of domain model for investigation of brain humor.*

***Keywords**: AGILE, Smart Health Information System, Domain Model, Class Diagram.*

Introduction

In order to design quality model in Smart Health Information System it is necessary to spend some efforts to design well-structured model. Requirements for the stable and well-designed models are follows:

- ✓ consider peculiarities of human perception;
- ✓ full and strict representation of the system;
- ✓ possibility to analyze whole system as well as to investigate detailed algorithms or structure of some system component;
- ✓ representation of system functionality or structure in full and non-contradictory way;
- ✓ metamodeling notation for representation of metamodels must provide compact and detailed representation of metamodels in order to design quality models that would satisfy metamodeling notation as well as all stakeholders needs.

Designing analytical representation of domain model to perform deep learning neural networks to investigate brain tumors

Let's consider approach proposed in paper [Chebanyuk & Palahin, 2019] by means of designing domain model of problem domain "deep learning in investigation of brain tumors"

First task is to gather artifacts that contain initial information about problem domain:

Gathered artifacts: methodology, information about principles of processing images, shown different brain diseases, pictures shown photos of brain with different diseases.

Pictures shown diseases are templates for comparing health and weak brain. They given below (figure 1) Information, extracted from these pictures (Mohsen H. et al., 2018), represents names of brain diseases that are entities of problem domain: metanistic bronchogenic carcinoma, sarcoma, ghoblastoma.

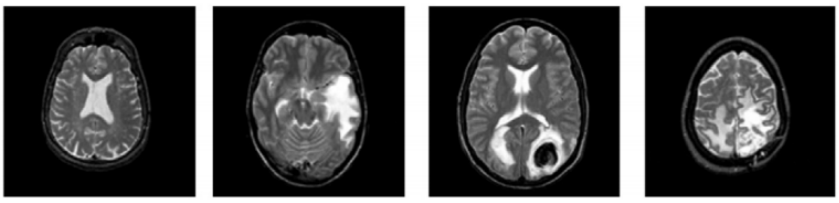

Figure 1. Pictures of brain diseases

Figure is taken from

https://reader.elsevier.com/reader/sd/pii/S2314728817300636?token=8A70AB99228A64BF0FB63C553EAEEEF7C530BBB879C49D81B44413D812CFE1DD51FF56276B9A0BE2AB691355B5363EEC

Let's analyze the metamodeld problem domain of "deep learning in investigation of brain tumors" [Chebanyuk & Palagin, 2019].

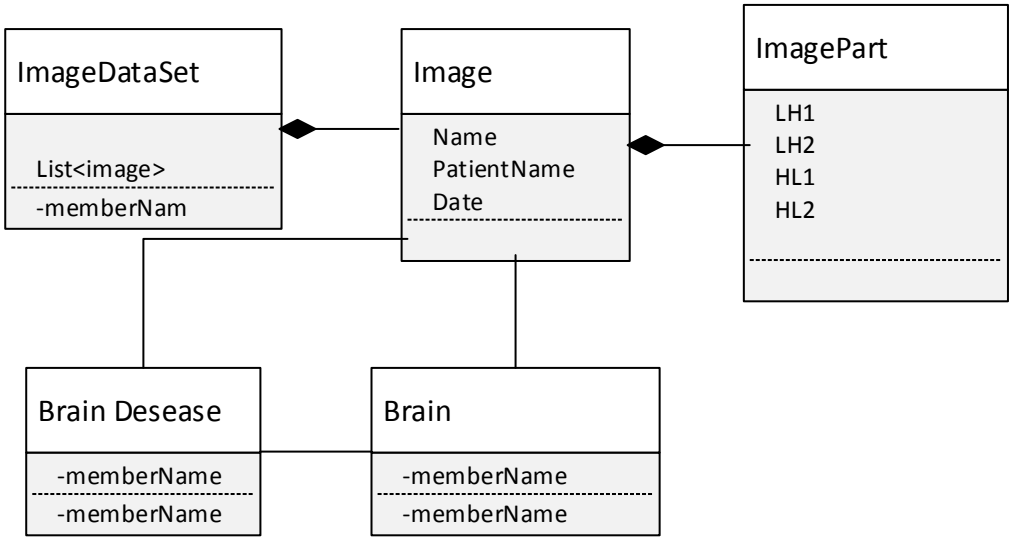

Figure 2. Model of problem domain of “Deep learning in investigation of brain tumors”

Figure is taken from [Chebanyuk & Palagin, 2019].

Analytical representation of domain model is prepared according to algebra, describing software static models [Chebanyuk E.V., 2013]: Class “Brain” is chosen as domain model central component [Heba et al., 2018].

$$\begin{cases} \{Brain* = BrainDesease(ass)Brain(ass)\,\mathrm{Im}\,age \\ \mathrm{Im}\,age^{*} = \mathrm{Im}\,agePart(comp)\,\mathrm{Im}\,age \\ \mathrm{Im}\,ageDataSet^{*} = \mathrm{Im}\,age(comp)\,\mathrm{Im}\,ageDataSet \end{cases}$$

Conclusion

Paper is devoted to designing of analytical representation of problem domain model “deep learning in investigation of brain tumors” using algebra, describing software static models. Analytical representation of domain model provides tha basic for performing new operations use domain model as a template to improve structural characteristics. allow to the new specialists to aquent with the structure of domain model quickly and “download into memory” structural organization of problem domain in order to define the best diagnosis, analyze schemas of improving health when a patient get some brain disease as well as provide research collaboration in investigation sphere.

Acknowledgement

Abhors are grateful for Artificial Intelligence and Knowledge Engineering Research Labs for given material for investigation.

Bibliography

[Chebanyuk O. & Palahin O., 2019] Chebanyuk O. & Palahin O. (2019) Domain Analysis Approach. International journal "Informational Content and Processing". Volume 6, Number 2, 2019, 3-20.

[Chebanyuk E.V., 2013] Chebanyuk E.V. (2013) Algebra, describing software static models. International journal "Information Technologies and Knowledge", Vol. 7, №1, 2013, 83-93

[Heba et al., 2018] Heba Mohsen , El-Sayed A. El-Dahshan , El-Sayed M. El-Horbaty , Abdel-Badeeh M. Salem, "Classification using deep learning neural networks for brain tumors",ScienceDirect, Future Computing and Informatics Journal 3 (2018) 68- 71 http://www.journals.elsevier.com/future-computing-and-informatics-journal/

[Mohsen H. et al., 2018] Mohsen, H., El-Dahshan, E. S. A., El-Horbaty, E. S. M., & Salem, A. B. M. (2018). Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, *3*(1), 68-71.

Authors' Information

Olena Chebanyuk – Assoc. Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine. E-mail: chebanyuk.elena@ithea.org (chebanyuk.elena@gmail.com)

Major Fields of Scientific Research: Model-Driven Architecture, Model-Driven Development, Software Architecture, Mobile Development, Software Development

Abdel-Badeeh M. Salem Prof. Dr. Professor of Computer Science, Founder &Head of Artificial Intelligence and Knowledge Engineering Research Labs Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt: abmsalem@yahoo.com

Major Fields of Scientific Research: Intelligent Information Processing, Data Mining, and Expert Systems.2- Image Processing and Pattern Recognition,- Big Data Analytics,

Intelligent Education and e-Learning, Medical Informatics, Bio-informatics and e-Health, Knowledge Engineering and management, Intelligent software, Machine learning and Computational Intelligence.

CLASSIFICATION OF INITIAL INFORMATION FOR DOMAIN MODELS DESIGNING IN DEPENDEBILITY NETWORKS AREA

Mykola Guzii

Abstract: According to definition of domain analysis its aim to design new artifacts for reuse in further activities. Domain models serve as templates for refinement operations or for designing new artifacts that consider problem domain structure and reflect interconnections between problem domain entities.

Application of domain model in security system verification allows to improve structural characteristics if security system schemas. Thus, the important task is to provide initial quality information for domain models designing. Paper proposes a classification of sources of vulnerabilities in dependability networks area.

Keywords: Domain Engineering, Domain Analysis, Domain Model, Dependebility.

Introduction

Domain Analysis is the foundation for the future reusability of generated software development artifacts [Kyo et al, 1990]. An important improvement of the reuse process happens when we succeed in deriving common architectures, generic models or specialized languages by using domain analysis that helps the software development process in a specific problem area [Frakes, 2005].

Domain knowledges usually represented in informal way. Other problem that sources of knowledge and their representation are also different [J. Olson & G. Olson, 2015]. Many papers are devoted to the problem of preparing a structural representation of domain knowledge and organizing them into domain models. Review of domain analysis approaches according is represented in paper [Chebanyuk & Palahin, 2019]. In order to represent domain models languages and corresponding environments for their visual representation are used widespread. Examples of them are ECORE and other derivable languages from it as UML or its extensions [UML, 2017].

Approaches to domain models designing based on processing information that represented informally in different sources, namely controlled vocabularies, natural texts,

and other sources of information [Frakes et. al., 2005], [Beck et al, 2001] [Endres & Rombach, 2003] [Mangano et al, 2015]. [Asnina, 2006]

That why the task of prospecting initial information for domain models designing is important.

Area of designing of protected networks requires high precise procedures in such systems designing. The cost of mistakes is too high. Modern researches are directed on gathering the practical experience of protected network research. That why implementation of the formal methods of reusing lets gather experience of different users with the aim to collect the best practices

Classification of networks dependability

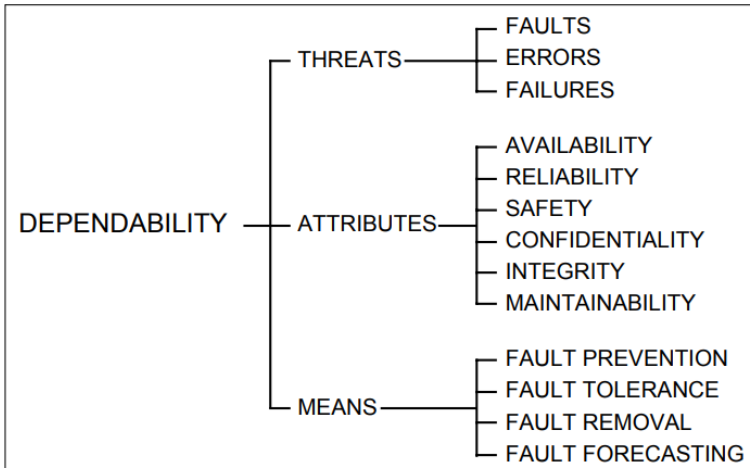

Figure 1 Classification of dependability

Figure is taken from

https://www.ppgia.pucpr.br/~alcides/Teaching/SistemasDistribuidos/TOF/avizienis01fundamental.pdf

Processing of the initial information for domain analysis requires gathering two sources namely fundamentals of dependability and template-based analysis of Quality of Service formal models (Avizienis A. et al, 2001), (Bruneo D et al., 2010) (Weber P. et al., 2012). Schema, containing the classification of dependability is represented in the figure 1.

Table 1 summarizes the results of analysis of the papers that provide the basic of classification of dependability and let to classify security risks and vulnerabilities according to levels of dependabilities.

Table 1

Matching security vulnerabilities to dependability

vulnerabilities		Sources of vulnerabilities	Prevention activities
Threats (Al-Kuwaiti M. et al., 2009)	Faults	Unqualified persons Spyers Keyboard sniffers Fault	Certification and company cafety policies Virus control
	Errors	Memory leaks	memory management
	Failures	Add extra information to data stream Viruses audio bookmarks	Hashing White noise
Attributes (Trivedi K. et al, 2009)	Availability	Drivers and network devices	Network ports management
	Reliability	Peer to Peer network vulnerabilities Software cash and cookies	KerBeRos Traffic encryption
	Safety	Network ports control Data channels control Un-plugging drivers of camera microphone and other devices	Proxies Physical observation Antispy software Process control Physically controlling of devices
	confidentiality	Network routes control	Routers settings State laws
	integrity	Trust data channel Control sum errors	Security testing Cryptography protocols
	Maintainability	Shell documentation errors	Software security components quality estimation

vulnerabilities		Sources of vulnerabilities	Prevention activities
Means (Bobbi et al, 2001)	Fault prevention	Virtual file system memory errors	Audit System logs analysis
	Fault tolerance	Security components conflicts	Firewall Estimation of security system risks Security testing of operating system kernel antiviruses
	Fault removal	Spy	Updates
	Fault forecasting	Hardware errors	Analysis of potential vulnerabilities in desktop software and mobile operating systems State laws

Conclusion

The contribution of this paper is to prepare initial information for domain analysis in the area of dependability. Table 1 contains the summarizing of faults that are relates to narrow sphere of dependability. Systematic representation of information related to dependability allows to design domain model and consider related entities in area of protected networks designing.

Bibliography

[Allen, 2015] Edward B. Allen. Design Artifacts are Central: Foundations for a Theory of Software Engineering. Technical Report MSU-20150420. Mississippi State University, Mississippi State, Mississippi 39762 April 2015. http://web. cse.msstate.edu/~allen/Allen15MSU20150420.pdf (accesed 01.09.2015)

[Al-Kuwaiti M. et al., 2009] Al-Kuwaiti, M., Kyriakopoulos, N., & Hussein, S. (2009). A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability. *IEEE Communications Surveys & Tutorials*, *11*(2), 106-124.

[Asnina, 2006] Asnina, Erika. "The Formal approach to problem domain modelling within model driven architecture." Proceedings of the 9th International Conference "Information Systems Implementation and Modelling"(ISIM'06). 2006.

[Avizienis A. et al, 2001] Avizienis, A., Laprie, J. C., & Randell, B. (2001). *Fundamental concepts of dependability* (pp. 7-12). University of Newcastle upon Tyne, Computing Science.

[Beck et al, 2001] K.Beck, M.Beedle, A.van Bennekum, A.Cockburn, W.Cunningham, M.Fowler, J.Grenning, J.HighUML diagramith, A.Hunt, R.Jeffries, J.Kern, B.Marick, R.C.Martin, S.Mellor, K.Schwaber, J.Sutherland, D.Thomas. Agile Manifesto Copyright 2001. Access mode http://www.agilemanifesto.org/ (accessed 02.09.2015)

[Bobbi et al, 2001] Bobbio, A., Portinale, L., Minichino, M., & Ciancamerla, E. (2001). Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering & System Safety*, *71*(3), 249-260.

[Bruneo D et al., 2010] Bruneo, D., Puliafito, A., & Scarpa, M. (2010, April). Dependability analysis of wireless sensor networks with active-sleep cycles and redundant nodes. In *Proceedings of the First Workshop on DYnamic Aspects in DEpendability Models for Fault-Tolerant Systems* (pp. 25-30).

[Chebanyuk O. & Palahin O., 2019] Chebanyuk O. & Palahin O. (2019) Domain Analysis Approach. International journal “Informational Content and Processing”. Volume 6, Number 2, 2019, 3-20.

[Frakes et al, 2005] Frakes, William B., and Kyo Kang. "Software reuse research: Status and future." IEEE transactions on Software Engineering31.7 (2005): 529-536.

[J. Olson & G. Olson, 2015] J. R. Olson and G. M. Olson. The Growth of Cognitive Modeling in Human Computer Interaction Since GOMS. University of Michigan. Access mode http://www.ics.uci.edu/~kobsa/courses/ICS205/03F/goms.ppt (accesed 01.09.2015)

[Endres & Rombach, 2003] A. Endres and D. Rombach. A Handbook of Software and Systems Engineering: Empirical Observations, Laws and Theories, Pearson - Addison Wesley, Harlow, England, 2003. ISBN 978-0321154200. 327 pp.

[Frakes, 2005] Frakes, William B., and Kyo Kang. "Software reuse research: Status and future." IEEE transactions on Software Engineering 31.7 (2005): p. 529-536.

[Kyo et. al., 1990] Kyo C. Kang Sholom G. Cohen James A. Hess William E. Novak A. Spencer Peterson Feature-Oriented Domain Analysis (FODA) Feasibility Study. Technical Report CMU/SEI-90-TR-21 ESD-90-TR-222 Software Engineering Institute Carnegie Mellon University Pittsburgh, Pennsylvania 1521 p. 1-148 https://resources.sei.cmu.edu/asset_files/TechnicalReport/1990_005_001_15872.pdf

[Mangano et al, 2015] N. Mangano, T. D. LaToza, M. Petre, and A. van der Hoek,“How Software Designers Interact with Sketches at the Whiteboard,” IEEE Transactions on Software Engineering, vol. 41, no. 2, Feb. 2015, pp. 135–156.

[Trivedi K. et al, 2009]Trivedi, K. S., Kim, D. S., Roy, A., & Medhi, D. (2009, October). Dependability and security models. In *2009 7th International Workshop on Design of Reliable Communication Networks* (pp. 11-20). IEEE.

[UML, 2017] About the unified modeling language specification version 2.5 https://www.omg.org/spec/UML/2.5/

[Weber P. et al., 2012] Weber, P., Medina-Oliva, G., Simon, C., & Iung, B. (2012). Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, *25*(4), 671-682.

Authors' Information

Mykola Guzii –PhD., Prof. of Computer Systems and Networks Department, National Aviation University, Kyiv, Ukraine. E-mail: mykola.guzii@gmail.com

Major Fields of Scientific Research: Computer Networks, Software Engineering

AN APPROACH OF ARCHITECTURAL SOLUTIONS DESIGNING TRANSFORMING USE CASE DIAGRAMS

Olena Chebanyuk, Olena Marchenko

Abstract: Software model represented as UML diagrams are central development artifacts in AGILE approach.

Paper proposes transformation approach that can be base for plug-ins development with the aim to perform in automatic mode. Peculiarity of the proposed approach it considers all paths of the initial diagram (UML Use case diagram) and allows to transfer a semantic to the resulting class diagram. In order to perform transformation graph representation of Software model is used.

Keywords: AGILE, Software Model Transformation, Use Case, Class Diagram, Graph.

Introduction

Software model transformation is key activity in AGILE approach [Kuske, S. et al., 2002]. It allows obtaining a requirement specification representation with necessary level of details [Bézivin, J. et al, 2006]. Without transformation operation software models are "islands" of requirement or architecture representation [Veitaite, I., & Lopata, A., 2020]. Application of a set of transformations allows providing a complex thread for performing all software development lifecycle operations in AGILE approach [Kerkouche, E. et al., 2020].

Transformation Approach

The essence of the proposed approach is to transform UML Use Case diagram into Class one using whole structure of Use Case. Involving existing transformation approaches often needs software architect efforts to gather UML Class Diagram structure in accordance with requirement specification semantics.

The contribution of this paper Proposed approach provides an analytic background for transforming UML Use Case diagram into Class one. Transformation is performed in two

steps, namely preliminary transformation with further class diagram elicitation. Preliminary transformation allows to generate BLOB class with set of public methods that matches to Use Case diagram precedents. Algebra, describing software static models [Chebanyuk, 2013] proposes to denote a class in the following way:

$$c(a) = \{\mathrm{A}, \mathrm{B}, \mathrm{X}\}$$
$$B = \{\beta_1, \beta_2, ..., \beta_n\}$$

where A – a set of class attributes, B – a set of its public methods, X – a set of class properties.

The first step of is representation of "one to one" transformation rules.

- ✓ Actor to class renaming rule: $a \to c(a)$, the name of the class c matches with name of the actor a. Example actor named User is transformed to class User. It is denoted by the following: class(user).
- ✓ Precedent to public method naming rule: $p \to \beta(p)$, precedent named p is transformed to public method of the class c(a) with the same name. Example precedent p, named "load data" is transformed to public method of class named "load data".
- ✓ Class naming rule: Denote a set of domain entities in controlled vocabulary, *ConVoc* as $ConVoc = \{c_1, c_2, ..., c_n\}, n = |ConVoc|$. A set of Use Case diagram precedents is denoted as follows: $P_{use\_case} = \{p_1, p_2, ..., p_k\}$ $p = (w_1, w_2, ..., w_t),\ p \in P_{use\_case}$ The rule of naming object is written in the following way:

$$ConVoc \cap p = \{w = c \mid w \in p, c \in Convoc\}$$
$$ConVoc \cap p \neq \varnothing \to class(name) = ConVoc \cap p \quad (1)$$

Proposed transformation approach is consists from the next steps:

1. Apply "actor to class renaming rule" and "class naming rule". Compose empty classes

2. Analyzing every precedent *p* according to (1) apply "Precedent to public method naming rule". If $ConVoc \cap p \neq \varnothing$ the method β is added to class:
$c(a) = \{A, B, X\}$
$B = \{(p_1 \to \beta_1(p_1)) \cup (p_2 \to \beta_2(p_2)) \cup ... \cup (p_n \to \beta_n(p_n))\}$

Conclusion and further research

Proposed approach composes a basic to development tools as well as practical approach to architectural solutions designing. The advantages of the proposed approach it allows to analyze whole Use Case diagram to design class diagram saving semantic of requirement specification (through Use case diagram). It requires less software architect efforts in comparison with knowing approaches in order to compose whole class diagram [Chebanyuk E, 2014].

Bibliography

[Bézivin, J. et al, 2006] Bézivin, J., Büttner, F., Gogolla, M., Jouault, F., Kurtev, I., & Lindow, A. (2006, October). Model transformations? transformation models!. In *International Conference on Model Driven Engineering Languages and Systems* (pp. 440-453). Springer, Berlin, Heidelberg.

[Chebanyuk, E., 2013] Chebanyuk, E. (2013). Algebra describing software static models. *International Journal "Information Technologies and Knowledge*, *7*(1), 1313-0455.

[Chebanyuk E, 2014] Chebanyuk E. (2014, January). An approach to class diagram design. In *2014 2nd International Conference on Model-Driven Engineering and Software Development (MODELSWARD)* (pp. 448-453). IEEE.

[Kerkouche, E. et al., 2020] Kerkouche, E., Khalfaoui, K., & Chaoui, A. (2020). A rewriting logic-based semantics and analysis of UML activity diagrams: a graph transformation approach. *International Journal of Computer Aided Engineering and Technology*, *12*(2), 237-262.

[Kuske, S. et al., 2002] Kuske, S., Gogolla, M., Kollmann, R., & Kreowski, H. J. (2002, May). An integrated semantics for UML class, object and state diagrams based on graph transformation. In *International Conference on Integrated Formal Methods* (pp. 11-28). Springer, Berlin, Heidelberg.

(Veitaite, I., & Lopata, A., 2020) Veitaite, I., & Lopata, A. (2020). Knowledge-Based Transformation Algorithms of UML Dynamic Models Generation from Enterprise Model. In *Data Science: New Issues, Challenges and Applications* (pp. 43-59). Springer, Cham.

Authors' Information

Olena Chebanyuk – Assoc. Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine. E-mail: chebanyuk.elena@ithea.org (chebanyuk.elena@gmail.com)

Major Fields of Scientific Research: Model-Driven Architecture, Model-Driven Development, Software Architecture, Mobile Development, Software Development

Olena Marchenko – Senior lecturer. of Software Engineering Department, National Aviation University, Kyiv, Ukraine. E-mail: marchenko.helene@gmail.com

Major Fields of Scientific Research: Database, Model-Driven Development, Mobile Development, SQL, Software Development Management, Software Development

EXTENDED TABLE ALGEBRA, EXTENDED MULTISET TABLE ALGEBRA, AND THEIR RELATIONSHIP

Iryna Lysenko

Abstract: The paper is focused on some theoretical questions of the table databases which should take into account during the development of software with relational (table) databases. Two mathematical formalisms such as extended table algebra and extended multiset table algebra are considered. Basic definitions referring to these formalisms are given. This paper is dedicated to the topic of the relationship between extended table algebra and extended multiset table algebra.

Keywords: relational databases, table algebra, multiset table algebra, extended operations.

ITHEA Keywords: H.2 DATABASE MANAGEMENT.

Introduction

In monograph [Redko, 2001] authors are considered a classical variant of table algebra. In monograph [Buy, 2016] it is done a generalization of table algebra, which means under relation any set of tuples, in particular, infinite. In addition, every table is correlated to a certain scheme. Multiset table algebra is also introduced in monograph [Buy, 2016]. In this case, under relation, we understand a multiset, in particular, infinite. Table algebra of infinite tables and multiset table algebra are also filled up by inner and outer join, semijoin, outer set operations, and aggregate operations to match the capabilities of SQL [Buy, 2016].

The purpose of this work is to discuss the question is whether extended table algebra a subalgebra of extended multiset table algebra with respect to the additional operations.

Basic definition

Let's introduce the basic concepts of table algebra of infinite tables and multiset table algebra in terms of monograph [Buy, 2016].

Let $\boldsymbol{A}$ be the set of attributes and $\boldsymbol{D}$ be the universal domain. An arbitrary finite set of attributes $R \subseteq \boldsymbol{A}$ is called scheme. The tuple of scheme R is a nominal set on pair R, $\boldsymbol{D}$. The projection of this nominal set for the first component is equal to R.

Under table algebra of infinite tables is understood an algebra $\langle \boldsymbol{T}, \Omega_{P,\Xi} \rangle$, where $\boldsymbol{T}$ is the set of all tables, $\Omega_{P,\Xi} = \{\cup_{\boldsymbol{R}}, \cap_R, \setminus_R, \sigma_{p,R}, \pi_{X,R}, \underset{R_1,R_2}{\otimes}, \div_{R_2}^{R_1}, Rt_{\xi,R}, \sim_R\}_{X,R,R_1,R_2 \subseteq A}^{p \in P, \xi \in \Xi}$ is the signature, P, Ξ are the sets of parameters.

A table of the scheme R ($R \subseteq \boldsymbol{A}$) is a pair $\langle t, R \rangle$, where t is a set (in particular infinite) of tuples of the fixed scheme R.

The operations of signature $\Omega_{P,\Xi}$ are defined in [Buy, 2016].

Under multiset table algebra is understood an algebra $\langle \Psi, \Omega_{P,\Xi} \rangle$, where Ψ is the set of all tables, $\Omega_{P,\Xi} = \{\bigcup_{All}^{R}, \bigcap_{All}^{R}, \setminus_{All}^{R}, \sigma_{p,R}, \pi_{X,R}, \underset{R_1,R_2}{\otimes}, Rt_{\xi,R}, \sim_R\}_{X,R,R_1,R_2 \subseteq \boldsymbol{A}}^{p \in P, \xi \in \Xi}$ is the signature, P, Ξ are the sets of parameters.

A table of the scheme R ($R \subseteq \boldsymbol{A}$) is a pair $\langle \psi, R \rangle$, where the first component ψ is an arbitrary multiset, the basis of which $\Theta(\psi)$ is the set of tuples of the same scheme and the second component R is a scheme of the table. The notation $Occ(s, \psi)$ denotes the number of duplicate tuple s in the multiset ψ.

Table algebra, multiset table algebra, and their relationship

The 1-multiset is multiset whose range of values is the empty set or a single-element set. Such multisets are the analogues of ordinary sets. Using this we can consider a table of table algebra of infinite tables as a pair $\langle t^1, R \rangle$, where t^1 is a 1-multiset (in particular infinite) of tuples of the fixed scheme R.

In this case obviously, that the set of all tables of table algebra of infinite tables $\boldsymbol{T}$ is a subset of the set of all tables of multiset table algebra Ψ, i.e. $\boldsymbol{T} \subseteq \Psi$.

In the paper [Glushko, 2018] it is found out able algebra of infinite tables does not form subalgebra of multiset table algebra. It is shown that table algebra of infinite tables is not

closed with respect to the union, projection and active complement operations of multiset table algebra.

Relationship between extended table algebra and extended multiset table algebra

The signature of extended table algebra and extended multiset table algebra contains also inner join (Cartesian join, natural join, join using attributes $A_1,...,A_n$ and join on predicate p) and outer join (outer left join, outer right join, outer full join and union join), semijoin and aggregate operations (*Sum*, *Avg*, *Max*, *Min*, *Count*). Formal mathematical semantics of these additional operations is defined in [Buy, 2016]. The universal domain $\boldsymbol{D}$ is replenished by the special element NULL for defining outer and aggregate operations.

Consider the case of inner join operations. Let us check Cartesian join $\underset{R_1,R_2}{Cj}$. Then $\left\langle t_1^1,R_1\right\rangle \underset{R_1,R_2}{Cj} \left\langle t_2^1,R_2\right\rangle = \left\langle t',R_1 \cup R_2\right\rangle$, where $\left\langle t_1^1,R_1\right\rangle \in \boldsymbol{T}(R_1)$, $\left\langle t_2^1,R_2\right\rangle \in \boldsymbol{T}(R_2)$. In other words, each tuple of $\left\langle t_1^1,R_1\right\rangle$ is paired with each tuple of $\left\langle t_2^1,R_2\right\rangle$, regardless of whether it is a duplicate or not.

Basis of multiset t' is defined as follows:

$$\Theta(t') = \left\{s' \mid \exists s_1 \exists s_2 \left(s_1 \in \Theta(t_1^1) \wedge s_2 \in \Theta(t_2^1) \wedge s' = s_1 \cup s_2\right)\right\}.$$

The number of duplicates is given by the following formula:

$$Occ(s',t') = Occ(s_1,t_1^1) \cdot Occ\,(s_2,t_2^1) = Occ(s' \mid R_1,t_1^1) \cdot Occ\,(s' \mid R_2,t_2^1) = 1,$$

where $s' \in \Theta(t')$. Thus, t' is a 1-multiset. Consequently, the set $\boldsymbol{T}$ is closed with respect to the Cartesian join $\underset{R_1,R_2}{Cj}$.

Check other inner join operations using a similar method we will see that each of them is closed with respect to the respective operations of extended multiset table algebra. This also applies to other additional operations.

Conclusion

This paper addresses the issue of the relationship between extended table algebra and extended multiset table algebra. It is proved that extended table algebra is closed with respect to the inner and outer join, semijoin and aggregate operations of extended multiset table algebra. The obtained results can be applied to the development of software with table databases.

Bibliography

[Redko, 2001] Redko, V., Brona, J., Buy, D. and Poliakov, S. Relation database: relation algebras and SQL-similar languages. Kyiv, Vydavnychyi dim Akademperiodyka Publ., 2001.

[Buy, 2016] Buy, D.B, Glushko, I.M. Calculi and extensions of table algebras signature. Nizhyn, NDU im. M. Gogol, 2016.

[Glushko, 2018] Glushko, I. About relationship between table algebra of infinite tables and multiset table algebra. In UkrPROG 2018, Proceedings of the 11th International Conference of Programming, pp.159–163. http://ceur-ws.org/Vol-2139/159-163.pdf

Authors' Information

Iryna Lysenko – Assoc. Prof. of Information technology and Data Analysis Department, Nizhyn Gogol State University, Nizhyn, Ukraine; e-mail: glushkoim@gmail.com

Major Fields of Scientific Research: Relational databases, Table algebras

PREPARING OF INITIAL INFORMATION TO DESIGN SOFTWARE VERIFICATION MODEL

Yuliia Bezkorovaina

Abstract: The quality of the software has to be confirmed. For that purpose, we are using verification during development. It is necessary to have information about the existing standards in this field and follow them in all phases of the software lifecycle. At this work, there overviewed standards with term verification definition, existing verification methods, and features of the software quality. This area is critical because the knowledge of software quality standards and their application contributes to the development of quality software that will be widely used and durable. Based on research was proposed the software verification model.

Keywords: Verification, model, software, measure, internal quality.

ITHEA Keywords: D.2.4 Software/Program Verification

Introduction

Nowadays, the number and complexity of software applications are increasing, and the importance of software quality also increased. The developers and managers software must have robust tools and methodology verification models for improving software quality.

There are different methods and standards for solving quality problems. Therefore, it is necessary to define the right characteristics that significantly affect software performance and quality.

To build quality software, we need to build models of quality.

A Short survey of related papers

The related papers aimed at verifying artifacts on separate phases lifecycle, compare standards of quality and verification software.

Proposed approach

Verification and validation are two close processes to each other and take effect together about quality software. The quality software model describes in [ISO9126, 1991], [ISO91261, 2001], [ISO91262, 2003], [ISO91263, 2003], [ISO91264, 2004] and [ISO25010, 2011]. The characteristics of software quality are functionality, reliability, usability, efficiency, maintainability, portability, security, and compatibility [ISO25010, 2011]. Verification ensures internal quality software [ISO91263, 2003], and validation ensures external quality software [ISO91262, 2003].

In this work, we will talk about only verification. There is a lot of definitions of term "verification". So, the first step for describing the verification model, we need to find out "what verification is?" and "where it is using?":

- ✓ do the thing right? [Boehm, 1981];
- ✓ verification is (a) the process of evaluating a system or component to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase, or in other hands, (b) formal proof of program correctness [IEEE610, 1990];
- ✓ verification: Confirmation by examination and provisions of objective evidence that specified requirements have been fulfilled [IEEE1012,1998];
- ✓ verification: (a) the same definition in [IEEE610, 1990]. (b) The process of providing objective evidence that the software and its associated products conform to requirements (e.g., for correctness, completeness, consistency, and accuracy) for all life cycle activities during each life cycle process (acquisition, supply, development, operation, and maintenance); satisfy standards, practices, and conventions during life cycle processes; and successfully complete each life cycle activity and satisfy all the criteria for initiating succeeding life cycle activities (e.g., building the software correctly) [IEEE1012, 2005];

- ✓ by verification we mean conformance checking of results of separate phases of software development to requirements and restrictions stated for these results on previous phases [Kuliamin, 2008];
- ✓ verification: confirmation, through the provision of objective evidence, that specified requirements have been fulfilled [ISO25010, 2011];
- ✓ verification: (a) the same definition in [IEEE610, 1990] and [IEEE1012, 2005]. (b) the same definition in [IEEE1012, 2005] but expands part “software” to “system, software, or hardware”. Verification of interim work products is essential for proper understanding and assessment of the life cycle phase product(s) [IEEE1012, 2017].
- ✓ depending on the phase of the software lifecycle, the internal quality software can ensure lots of verification methods. Those methods can be classified as review, static analysis, formal, dynamic, and synthetic methods [Kuliamin, 2008]. They include the next methods:
- ✓ review – management review, technical review, walkthrough, inspection, audit;
- ✓ static Analysis – correctness, patterns errors;
- ✓ formal methods - theorem proving, model checking, abstract interpretation;
- ✓ dynamic methods – runtime verification, online verification, passive testing, profiling, testing (unit, integration, system, regression, black or white box);
- ✓ synthetic methods - model-based testing, runtime verification, passive testing, extended static checking, abstract interpretation.

Each method can be used along with each other or separate on phases of the software lifecycle. For that purpose, they may be used tools or invite the specialist. The software lifecycle separates into phases [ISO12207, 2017]. So, review and formal methods are using for the whole software lifecycle; static and synthetic methods are using after the design phase, and dynamic are using after the testing phase to the end of the lifecycle [Kuliamin, 2008].

Conclusion

Research shows that the software verification model can help to improve the software development. It's including characteristics quality and their verification methods. This area is critical because the knowledge of software quality standards and their application contributes to the development of quality software that will be widely used and durable.

Further researches

In this work, we considered the software quality with the waterfall lifecycle and the verification methods of the artifacts its phases, so next will researched iterative models.

Bibliography

[Boehm, 1981] Barry W. Boehm. Software Engineering Economics. Prentice-Hall, Englewood Cliffs, NJ, 1981, 767 p.

[IEEE1012, 1998] IEEE Standard for Software Verification and Validation, in IEEE Std 1012-1998, 1998, 80 p.

[IEEE1012, 2005] IEEE Standard for Software Verification and Validation, in IEEE Std 1012-2004 (Revision of IEEE Std 1012-1998), 2005, 110 p.

[IEEE1012, 2017] IEEE Standard for System, Software, and Hardware Verification and Validation, in IEEE Std 1012-2016 (Revision of IEEE Std 1012-2012/ Incorporates IEEE Std 1012-2016/Cor1-2017) - Redline, 2017, 465 p.

[IEEE610, 1990] IEEE Standard Glossary of Software Engineering Terminology, in IEEE Std 610.12-1990, 1990, 84 p.

[ISO12207, 2017] ISO/IEC/IEEE 12207:2017 Systems and software engineering — Software life cycle processes. ISO, Switzerland, Geneva, 2017, 145 p.

[ISO25010, 2011] ISO/IEC 25010:2011. Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models. ISO, Switzerland, Geneva, 2011, 34 p.

[ISO9126, 1991] ISO/IEC 9126:1991. Software engineering — Product quality, 1991, 13 p.

[ISO91261, 2001] ISO/IEC 9126-1 Software engineering – Product quality – Part 1: Quality model. ISO, Switzerland, Geneva, 2001, 25 p.

[ISO91262, 2003] ISO/IEC TR 9126-2 Software engineering – Product quality – Part 2: External metrics. ISO, Switzerland, Geneva, 2003, 86 p.

[ISO91263, 2003] ISO/IEC TR 9126-3 Software engineering – Product quality – Part 3: Internal metrics. ISO, Switzerland, Geneva, 2003, 62 p.
[ISO91264, 2004] ISO/IEC TR 9126-4 Software engineering — Product quality — Part 4: Quality in use metrics. ISO, Switzerland, Geneva, 2004, 59 p.
[Kuliamin, 2008] V.V. Kuliamin. Metody verificatsii programmnogo obespetchenia. Institut systemnogo programirovania RAN, Moskva, 2008, 117 p.
[Sommerville, 2015] Ian Sommerville. Software engineering. Pearson, London, 2015, 811 p.

Authors' Information

Yuliia Bezkorovaina – Senior Lecturer of Software Engineering Department, National Aviation University, Kyiv, Ukraine; e-mail: yuliia.bezkorovaina@nau.edu.ua

Major Fields of Scientific Research: Software verification, software development, software engineering

SOFTWARE QUALITY: APPROACH TO QUALITATIVE TESTING AND THEORETICAL BASICS

VLADISLAV BEREZOVSKIY

Abstracts: one of the global software features is the quality of the product. Quality is an important aspect in our world that is used by well-known companies in the IT field. Thus, the well-known software testing tool ISO/IEC/IEEE 29119.

Keywords: software, quality, testing, ISO/IEC/IEEE 29119.

Introduction

To understand the essence and quality of software, it is necessary to substantiate frequently asked questions: what role does quality play?

Quality is one of the most demanding features. It is coercively used in public institutions. Software testing was also worth mentioning. Indeed, the satisfaction of entrepreneurs affects verification. Thus, new testing methodologies are being developed that should be improved by previous versions.

Definition of specification: quality, which is defined as a product or service, is obviously analyzed by the fixed characteristics of user satisfaction.

Consequently, quality affects the development of product trends and directly to create test cases.

Background

An internationally agreed set of standards that is often found in software testing methodologies includes ISO/IEC/IEEE 29119. They make it easy to write test cases in any organization. The table below complies with the interface-based modeling concept, which better measures software quality.

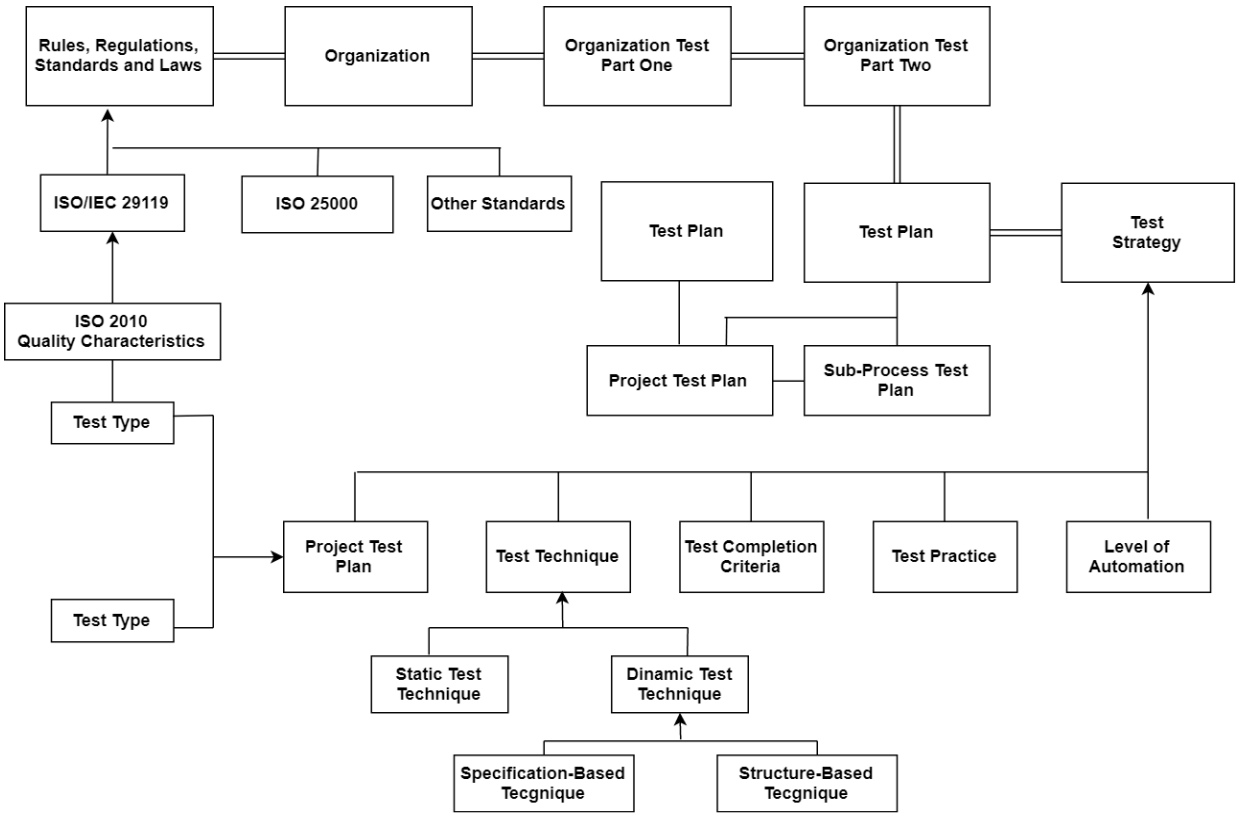

Figure 1. Interface Based Software Testing

To build this scheme, a standard for reviews of work products ISO/IEC/IEEE 29119-1, (2013), was defined.

According to the latest ISO/IEC/IEEE 29119 standards, several software testing standards that have stepped down have replaced: IEEE 829 documentation, BS 7925-2 test methods and BS 7925-1 software test glossary.

The target audience includes: testers, test managers, project managers, developers and people who are involved in the management or implementation of software testing.

Monitoring and Testing Management Process

The purpose of monitoring and testing management includes tracking the progress of testing in accordance with the test plan and with the testing specifications (for example, organizational testing policies). He also, as necessary, takes control actions and announces the necessary updates to the test plan.

This process is also used to track the progress of testing with higher-level test plans, such as a project test plan, and manage specific testing areas or certain types of testing.

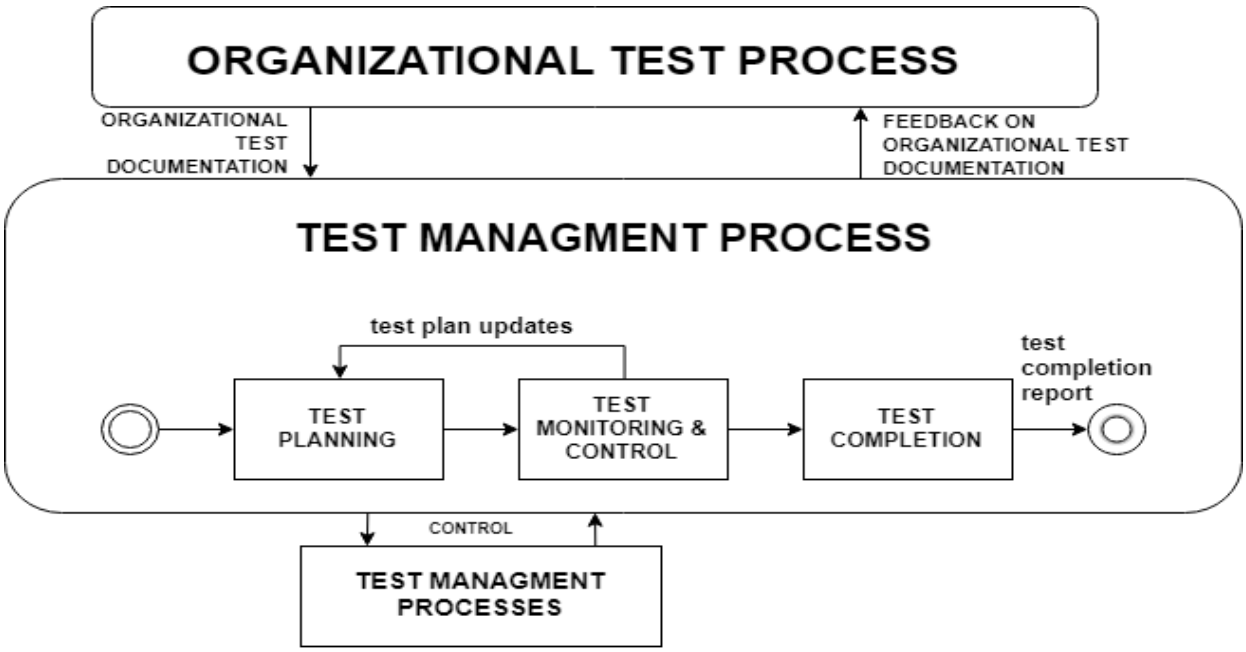

Figure 2. The role of Testing Management Process

Conclusion

The article examined important factors affecting the quality of software testing standards. The analysis of software standards was performed by certain criteria of quality work. Thanks to the table, we analyzed the concepts of modeling software testing based on the interface, which allows us to better measure quality, and also touched upon the role of the monitoring and testing management process.

Bibliography

Stuart Reid ISO/IEC/IEEE 29119 Software testing standards (2017, May) pp. 14-22.
GOST R 56922-2016 /ISO/IEC/IEEE 29119-3: 2013 System and Software Engineering. Software Testing. Part 3. Testing documentation (pp. 7-18).

Authors' Information

Vladislav Berezovskiy - Odessa National Polytechnic University, Odessa, Ukraine.
E-mail: vladinyo@ukr.net
Major Fields of Scientific Research: Informational Technologies, Quality Assurance Automation.

SECTION: APPLIED ASPECTS OF SOFTWARE ENGINEERING

AGENT SYSTEM FOR INTELLECTUAL CHATBOT

Anna Litvin, Vitalii Velychko, Vladislav Kaverinsky

Abstract: An agent system was developed to realize the interaction between chatbot components and the user’s interface. The system is oriented on using an ontology-driven knowledge base. The system allows conducting simultaneous interaction with multiple users by operating with asynchronous multithreading and temporary agents.

Keywords: agent system, chatbot, NLP, NLU, ontology, knowledge base.

ITHEA Keywords: D.2.11 Software Architectures.

Introduction

A multiagent system allows distributing tasks and functional among independent agents that are not controlled directly by a user [Nvana, 1996]. The agent's behavior is attempting to implement the BDI model. Such architecture has some advantages [Poslad, 2007]. A multiagent paradigm implies that the execution of the tasks could be distributed and it assumes a possibility for hardware load reduction. Each agent can operate as a microservice requested by other agents when needed. But unlike microservices, agents can make self-controlled decisions according to the situation that reflects in their “beliefs” and “desires”. One agent could be replaced by a new one or modernized without the need of rebuilding of others. If some agent crushes it does not lead to the whole system crash and an emergency could be more simply located and liquidated. One implementation of an agent system is an intellectual chatbot [Galitsky, 2019]. In this case functionality of knowledge base, interaction with users and system control is shared among agents.

A multiagent driven chatbot system

An agent system was developed to realize the interaction between chatbot components. Agents of the system are: Ontology Agent, Intermediary Agent, Register Agent, Control Agent, and Client Agents, which are temporary and exist at the same time in the quantity

equal to the number of users now interacting with the system. On figure 1 below a scheme of the developed agent system is given.

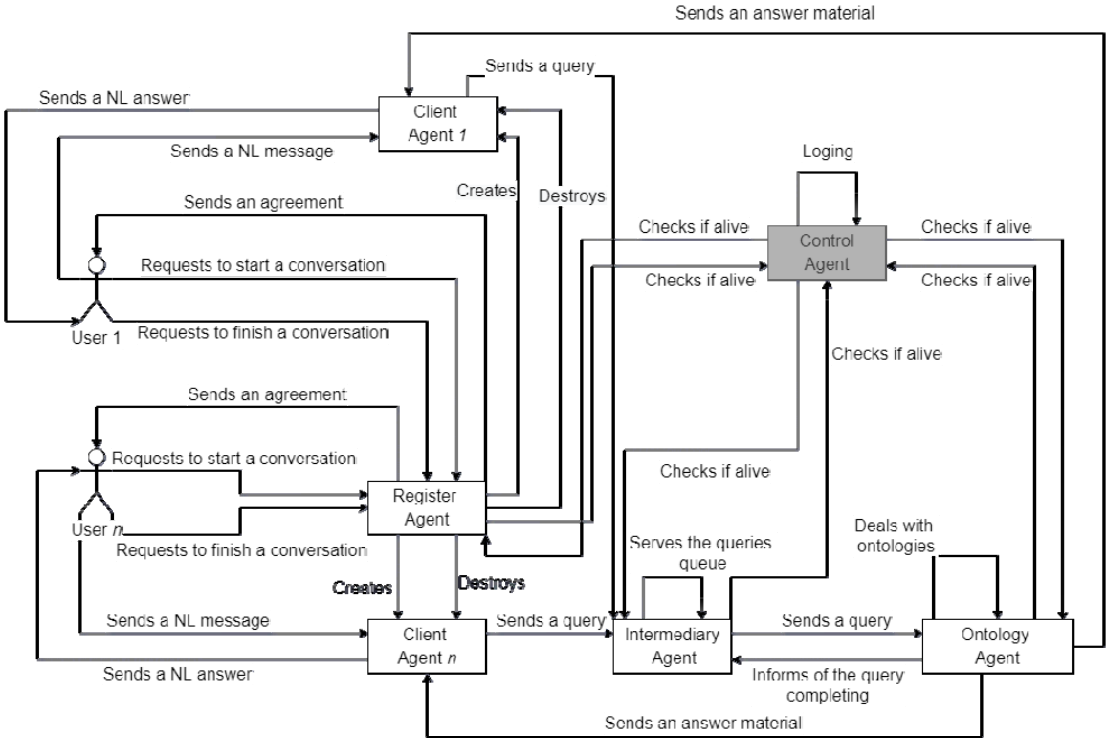

Figure 1. A scheme of the developed agent system (white components are already realized; gray ones are to be implemented in future)

Each agent of the system is an independent HTTP server interacting by FIPA protocol [Poslad, 2007]. The schema assumes that a user has an HTTP client that can interact with Register Agent and Client Agent API using FIPA protocol.

To begin a conversation a user sends a message to Register Agent with a request to start it. Then Register Agent asks Control Agent whether the system works properly. If the message is correct and there are no emergencies in the system, the Register Agent sends to the user a consent message and an address of the Client Agent it has created. The following interaction between the user and the system will be performed through this Client Agent except for the request of the conversation finishing, which will be addressed directly to the Register Agent. The presence of Client Agents allows the independence and asynchrony of conversations of various users in parallel. A Client Agent can contain a conversation and user's personal information. It performs NLP and NLU operations

with user’s messages and forms sets of SPARQL queries. These procedures are quite burdening and their performance takes a rather long time. Hence it seems a reasonable decision to make these agents multiple and working independently. Register Agent can have a possibility to run a Client Agent on the least loaded server, or it could be also beneficial to implement Client Agent on a user’s machine. Formed queries bounded to the conversation and user’s message IDs are sent to the Intermediary Agent that serves the request queue. Each message from a client is pushed by it to a queries queue. In a different thread the messages one by one are popped from the queue and send to an Ontology Agent. The main role of the Ontology Agent is to perform SPARQL queries to ontologies. Also there is preformed the most fitting ontology selection. To accelerate queries performance this process could be parallelized. After obtaining a set of results from ontology the following set of special SPARQL queries is performed to determine what these entities are in general. As one query package complete, Ontology Agent informs the Intermediary Agent. The obtained results are sending directly to the Client Agent. After receiving a message from the Ontology Agent the Client Agent forms a natural language message and sends it to the user. If the user is to finish the conversation, he gets his client to send a special message to the Register Agent with an Intention to stop the conversation. If the message is correct and there are no obstacles for the action the Register Agent destroys the Client Agent and informs the user that the conversation is over. Control Agent works all the time system operates and its main goal is to check whether agents work properly and trying to automatically recovery them if needed. If some agent does not receive a ping message from the Control Agent for some times in succession it pings the Control Agent and if needed trying to recover it.

Conclusion

In the paper a schema of an agent system for an ontology-driven chatbot is presented. The system is sharing typical tasks of such applications among Ontology Agent, Intermediary Agent, Register Agent, Control Agent, and Client Agents. Presented principles of the system allow independent and asynchronous work with multiple users, persistency and optimizing loading.

Bibliography

[Nvana, 1996] Nvana H. Software. Agents: An Overview. Knowledge Engineering Review, Vol.21, Issue 3. Cambridge University Press, 1996. pp. 205 – 244.

[Poslad, 2007] Stefan Poslad. Specifying Protocols for Multi-agent System Interaction. ACM Transactions on Autonomous and Adaptive Systems, Vol.4, Issue 4. 2007. pp. 15 – 39.

[Galitsky, 2019] B. Galitsky. Developing Enterprise Chatbots. Learning Linguistic Structures. Springer, San Jose, 2019.

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 Interests: 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 Interests: 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 Interests: mathematical modeling.

A MARKOV CHAIN MODEL FOR THE TEACHING PROCESS

Michael Gr. Voskoglou

Abstract: An absorbing Markov chain is introduced to represent mathematically the teaching process, when it is based on the ideas of constructivism for learning. Interesting conclusions are derived and a measure is obtained for the teaching effectiveness. An example on teaching the derivative to fresher university students is also presented illustrating our results.

Keywords: Constructivism, Transition Probabilities, Markov Chain (MC), Absorbing MC (AMC).

ITHEA Keywords: G Mathematics and Computing.

Introduction

Constructivism is a philosophical framework based on Piaget's theory for learning and formally introduced by von Clasersfeld during the 1970's. Constructivism argues that knowledge is not passively received from the environment, but is actively constructed by the learner through a process of adaptation based on and constantly modified by the learner's experience of the world [Taber, 2011]. The application of the ideas of constructivism in the teaching process has become very popular during the last decades. The steps of typical framework for teaching based on those ideas are the following:

- ✓ *Orientation* (S_1): This is the starting step which connects the past with the present learning experiences and focuses student thinking on the learning outcomes of the current activities.
- ✓ *Exploration* (S_2): In this step students explore their environment to create a common base of experiences by identifying and developing concepts, processes and skills.
- ✓ *Formalization* (S_3): Here students explain and verbalize the concepts that they have been explored and the instructor has the opportunity to introduce formal terms, definitions and explanations for the new concepts and processes and to demonstrate new skills.

- ✓ *Assimilation* (S_4): In that step students develop a deeper and broader conceptual understanding and obtain more information about areas of interest by practicing on their new skills and behaviors.
- ✓ *Assessment* (S_5): This is the final step of the teaching process, where learners are encouraged to assess their understanding and abilities and teachers evaluate student skills on the new knowledge.
- ✓ Depending on the student reactions in the classroom, there are forward or backward transitions between the three intermediate steps (S_2, S_3, S_4) of the above framework during the teaching process, the flow-diagram of which is shown in Figure 1.

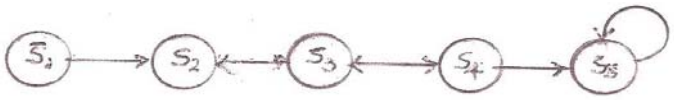

Figure 1: The flow-diagram of the teaching process

A Markov chain is introduced in this article on the above steps of the teaching process and through it interesting conclusions are derived for the teaching effectiveness.

The Markov Chain Model

A *Markov Chain (MC)* is a stochastic process that moves in a sequence of steps (phases) through a set of states and has a *one-step memory.* That means that the probability of entering a certain state in a certain step depends on the state occupied in the previous step and not in earlier steps. This is known as the *Markov property*. However, for being able to model as many real life situations as possible by using MCs, one could accept in practice that the probability of entering a certain state in a certain step, although it may not be completely independent of previous steps, it mainly depends on the state occupied in the previous step [Kemeny & Snell, 1963]. When the set of states of a MC is a finite set, then we speak about a *finite MC.* A state of a MC is called *absorbing* if, once entered, it cannot be left. Further a MC is said to be an *absorbing MC (AMC)*, if it has at least one absorbing state and if from every state it is possible to reach

an absorbing state, not necessarily in one step. For general facts on AMCs we refer to Chapter 2 of [Voskoglou, 2017] and for more details to the [Kemeny & Snell, 1976].

Here we introduce a finite MC having as states S_i , i = 1, 2,…, 5, the corresponding steps of the teaching framework described in the previous section. From the flow-diagram of Figure 1 it becomes evident that the above chain is an AMC with S_1 being its starting state and S_5 being its unique absorbing state. The minimum number of steps before the absorption is 4 and this happens when we have no backward transitions between the three middle states S_2, S_3 and S_4 of the chain. Denote by p_{ij} the *transition probability* of the MC from state Si to state Sj, for i, j =1, 2,…,5. Then the *transition matrix* A of the chain and its *canonical form A**, obtained by listing the absorbing state first are the following:

$$A = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & p_{32} & 0 & p_{34} & 0 \\ 0 & 0 & p_{43} & 0 & p_{45} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}, \text{ with } p_{32} + p_{34} = p_{43} + p_{45} = 1.$$

$$A^* = \begin{array}{c} \\ S_5 \\ - \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{ccccc} S_5 & & S_1 & S_2 & S_3 & S_4 \end{array} \\ \left[\begin{array}{c|cccc} 1 & 0 & 0 & 0 & 0 \\ - & - & - & - & - \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{32} & 0 & p_{34} \\ p_{45} & 0 & 0 & p_{43} & 0 \end{array}\right] \end{array} = \left[\begin{array}{c|c} I_1 & O \\ - & - \\ R & Q \end{array}\right].$$

Denote by I_4 the 4X4 unitary matrix. It can be shown that $I_4 - Q$ is always an invertible matrix [] and by applying standard techniques of the linear algebra one finds that

$$N = [n_{ij}] = (I_4 - Q)^{-1} = \frac{1}{1 - p_{34}p_{43} - p_{32}} \begin{bmatrix} 1 - p_{34}p_{43} - p_{32} & 1 - p_{34}p_{32} & 1 & p_{34} \\ 0 & -1 + p_{34}p_{34} & -1 & -p_{34} \\ 0 & p_{32} & 1 & p_{34} \\ 0 & -p_{32}p_{43} & -p_{43} & -1 + p_{32} \end{bmatrix}.$$

The matrix *N* is called the *fundamental matrix* of the AMC. It is well known ([Kemeny & Snell, 1976], Chapter 3) that the entry n_{ij} of *N* gives the mean number of times in state s_i before the absorption, when the starting state of the AMC is s_j , where s_i and s_j are non absorbing states. Therefore, since S_1 is the starting state of the above AMC, it becomes evident that the mean number of steps before the absorption is given by the sum

$$T = n_{11} + n_{12} + n_{13} + n_{14} = \frac{3 - 2p_{43}p_{34} - p_{32} + p_{34}}{1 - p_{34}p_{43} - p_{32}} \quad (1).$$

It becomes evident that the bigger is T, the more are the student difficulties during the teaching process. Another factor of the student difficulties is the total time spent for the completion of the teaching process. However, the time is usually fixed in a formal teaching procedure in the classroom, which means that in this case T is the unique measure of the student difficulties.

A Classroom Application

The following application took place recently at the Graduate Technological Educational Institute of Western Greece for teaching the concept of the derivative to a group of fresher students of engineering. The instructor used the instructional treatment that has been described in our Introduction.

It has been observed that the student reactions during the teaching process led to 2 transitions of the discussion from state S_3 (formalization) back to state S_2 (exploration). Therefore, since from state S_2 the chain moves always to S_3 (Figure 1), we had 3 in total transitions from S_2 to S_3. The instructor also observed 3 transitions from S_4 (assimilation) back to S_3. Therefore, since from state S_3 the chain moves always to state S_4 (Figure 1), we had 4 in total transitions from S_3 to S_4. In other words we had 3+3 = 6 in total "arrivals" to S_3, 2 "departures" from S_3 to S_2 and 4 "departures" from S_3 to S_4. Therefore $p_{32} = \frac{2}{6}$ and $p_{34} = \frac{4}{6}$. In the same way one finds that $p_{43} = \frac{3}{4}$ and $p_{45} = \frac{1}{4}$. Replacing the above values of the transition probabilities to equation (1) one finds that the mean number T of steps before the absorption of the MC is equal to 14. Consequently, since the minimum number of steps before the absorption is 4, the students faced significant

difficulties during the teaching process. This means that the instructor should find ways to improve his teaching procedure for the same subject in future.

Conclusion

In the paper at hands a mathematical representation of the teaching procedure of mathematics based on the principles of constructivism for learning was developed with the help of the theory of AMCs. This representation enables the instructor to evaluate the student difficulties during the teaching process, which is very useful for reorganizing the plans for teaching the same subject in future.

References

[Kemeny & Snell, 1963] Kemeny, J. G. & Snell, J.L. Mathematical Models in the Social Sciences, Ginn and Company, New York, USA, 1963.

[Kemeny & Snell, 1976] Kemeny, J.G. & Snell J.L., Finite Markov Chains, Springer - Verlag, New York, USA, 1976.

[Taber, 2011] Taber, K.S., Constructivism as educational theory: Contingency in learning, and optimally guided instruction, in J. Hassaskhah (Ed.), Educational Theory, Chapter 2, 39-61, Nova Science Publishers, NY, 2011.

[Voskoglou & Perdikaris, 1991] Voskoglou, M.Gr. & Perdikaris, S.C.,A Markov chain model in problem- solving, International Journal of Mathematical Education in Science and. Technology, 22, 909-914, 1991.

[Voskoglou, 2017] Voskoglou, M.Gr., Finite Markov Chain and Fuzzy Logic Assessment Models: Emerging Research and Opportunities, Createspace Independent Publishing Platform, Amazon, Columbia, SC, USA, 2017.

Author's 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

ANALYSIS OF EXISTING SOLUTIONS AND METHODS FOR SCANNING BARCODE

Yurii Milovidov

Abstract. A comparative analysis of existing solutions and methods for scanning barcodes using mobile devices was carried out. Each of these methods has its advantages and disadvantages. The use of a specific technology is determined by specific conditions.

Key words: mobile device, barcode, QR, Android, Zxing, Google Mobile Vision.

ITHEA keywords: D.2 Software engineering, D.2.1 Requirement/Specification, D.2.13 Reusable Software, D.2.2 Design Tools and Techniques, D.2.3 Coding Tools and Techniques.

Introduction

To develop a number of applications for mobile devices running on the Android operating system, it is necessary to select a tool for reading information from barcodes, in particular, QR codes that are used to store and transfer information such as links to web pages, mobile phone numbers, and geographical coordinates, contact details, etc.

Theoretical foundations

QR code (abbreviated from Quick Response Code) is the trademark for a type of matrix bar code first designed for the automotive industry. Today the QR code is widely used in all industries. QR-Code was developed in 1994. Its main advantage is simple scanning, which does not require special devices. A smartphone camera will be enough to scan. This feature has made it possible to spread QR-Code in areas such as logistics, trade and even manufacturing. This format is an image that has a two-dimensional matrix of black and white squares. The code itself can be quite small in size, so it's extremely convenient to place it even on small details or components. The information encrypted in the QR code is a string data type. Most often, QR codes are used to store and transmit information such as web page links, mobile phone numbers, geographic coordinates, contact information, and more. It should also be noted that the maximum number of

characters that can be stored in a single QR code is 4296 digits and letters or 7089 digits.

Related works

Factors affecting consumer usage of QR codes

This paper provides the results of a survey study of college students' awareness and use of QR codes. The findings indicate that the purpose of usage is significantly related to QR code usage rate. Practical users utilize QR codes more than experiential users and there is a positive relationship between electronic device ownership and QR code usage [Elif Ozkaya, H Erkan Ozkaya, Juanita Roxas, Frank Bryant & Debbora Whitson 2015].

QR Code Scanning app for Mobile Devices

In paper presented an implementation of an Android device using libraries is and combined algorithms in order to be able to scan any QR code fast accurate and easy. [Mircea Moisoiu, Andrei Negrău, 2014].

QR Codes: What Are They and Why Should You Care?

In this paper [Jason Coleman 2011]., it is described how individuals with smart phones can download free QR code reading applications and use their phone's camera to read the code; show several examples of creative ways libraries and other organizations are using QR codes.

Task

Among the requirements for the tool for reading information from special barcodes are the following:

- ✓ correct and trouble-free scan of barcodes with the subsequent transfer of information to the mobile application;
- ✓ ability to scan in poor quality lighting;
- ✓ scanner response speed when pointing the camera at the barcode;
- ✓ ability to get information when scanning from a certain distance;
- ✓ ability to scan various barcodes.

When choosing a tool, it is necessary to make sure that it is in accordance with its capabilities and requirements. The list of tools is limited to the specifics of the platform to which it will be applied. This platform is the Android operating system, so only those tools that are available for the selected platform will be considered.

Analysis of the ZXing Library

The libraries selected for comparison are the ZXing library and Google Mobile Vision. By the principle of work, they are similar. Both ZXing and Google Mobile Vision allow you to scan a variety of bar codes and transfer information to a mobile application for further work with it. These libraries are completely free to use and have open source code.

ZXing is an open source library. Its name can be deciphered as “zebra crossing”. It is a tool for reading barcode information of various formats.

In its arsenal, the ZXing library has a pretty impressive list of formats that can be read using it. Such a large list of formats supported by this library is its undoubted advantage.

The ZXing library has other competitive advantages. It is impossible not to indicate a high speed of recognition and reading information from a QR code. Of course, much depends on the device on which the mobile application with ZXing is installed. Also, the maximum distance, from which the reading of the QR code will be more dependent on the device itself.

It is worth noting another advantage of ZXing, namely the ease of use of this library from the point of view of a software developer. It is extremely integrated into the mobile application.

However, there are some drawbacks to its use. To begin with, it should be said that certain restrictions are imposed on its use. These restrictions include the minimum version of the Android operating system ZXing runs on. The minimum version of Android that this library works with is version 4.0.3 (API level 15). This means that the version of the application must be the same or higher.

Another disadvantage of using ZXing is the inability of the library to recognize QR-Code in low light. That is, under certain conditions in poor lighting in the room, the library loses the ability to recognize and read information from the bar code. However, it is difficult to call this fact a drawback. Each of the libraries reviewed works the same way.

Google Mobile Vision Library Analysis

The Mobile Vision Library is a library for reading barcodes that have been encrypted by various methods.

Mobile Vision does not lag behind ZXing in the list of barcode formats it is capable of working with.

The maximum distance at which Mobile Vision allows you to scan a barcode is about the same as ZXing.

The Mobile Vision Library has one interesting feature that is both an advantage and a disadvantage at the same time. These are the detailed settings the camera developers have when using it to read barcodes. Extra functionality is a tool that takes the barcode reader from the mobile application to the next level in the hands of an experienced software developer. However, in the hands of an inexperienced developer, this library is turning from a powerful tool into a burden.

Results

By comparing different scanning technologies and recognizing a variety of barcodes, was provided useful information to mobile device developers before deciding on the technology that is best for use.

Conclusions

The Google Mobile Vision library supports fewer barcode formats that it can recognize and decrypt compared to the ZXing library. However, ZXing has much more fans among developers, since it was developed much earlier than Mobile Vision and thereby managed to pass many years of testing and operation by developers. Despite the smaller number of settings that ZXing can demonstrate, it can be recommended for the development of mobile applications for QR code recognition.

Bibliography

[Oruç, A. and Ovatman, T. 2016] Testing of Web Services using Behavior-Driven Development, Proceedings of the 6th International Conference on Cloud Computing and Services Science (CLOSER 2016) - Volume 2, pages 85-92 ISBN: 978-989-758-182-3 – Access mode: https://www.scitepress.org/Papers/2016/59114/59114.pdf

[Oliveira, V., Julia, S., Passos, L. and Oliveira, K. 2017] A Linear Logic based Synchronization Rule for Deadlock Prevention in Web Service Composition, Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017) - Volume 2, pages 316-323 ISBN: 978-989-758-248-6 – Access mode: https://pdfs.semanticscholar.org/05a2/0725cbb574b6391ba9257730ab32803ced53.pdf

[Nacer, H. and Iggui, T. 2016] Semantic Networks – Based Approach for Web Services Management, Proceedings of the 11th International Joint Conference on Software Technologies (ICSOFT 2016) - Volume 2: ICSOFT-PT, pages 95-99 ISBN: 978-989-758-194-6 – Access mode: https://pdfs.semanticscholar.org/7832/5eb6408fecbf5c0eb65386c942b75f96212f.pdf

[Irum Rauf ; Inna Vistbakka ; Elena Troubitsyna 2018] Formal Verification of Stateful Services with REST APIs Using Event-B, Published in: 2018 IEEE International Conference on Web Services (ICWS) ISBN: 978-1-5386-7247-1 – Access mode: https://ieeexplore.ieee.org/document/8456341

Authors' Information

Yurii Milovidov, National University of Life and Enironmental Sciences of Ukraine, Kyiv, scientific Department of Programming Technologies. Senior Lecturer. Email: milovidov@email.ua

Major Fields of Scientific Research: Programming technologies, web-design, Internet – technologies.

FORMAL MODELS OF THE COMPREHENSIVE INFORMATION SYSTEM FOR THE MEDICINE-RELATED R&D SUPPORT

Vitalii Velychko, Kyrylo Malakhov, Oleksandr Shchurov

Abstract: The given paper describes design and development fundamentals (namely formal models) of the comprehensive problem-oriented information system for the scientific R&D lifecycle support in medicine. The R&D lifecycle support in medicine includes the processing of video/photo information (such as AI system for diagnosing the virus disease via processing of CT scans of patients' chests), the processing of medical data (such as Question answering medical expert system) and the NLP of medical corpora (such as processing of patient anamnesis, medical records and medical-related publications). The main results of our work are focused on enhancing the effectiveness of the scientist' R&D lifecycle in medicine.

Keywords: medicine-related R&D information system.

ITHEA Keywords: J.3 LIFE AND MEDICAL SCIENCES.

Introduction

One of the main ways to increase the effectiveness of solving a number of medical-related, social-related and economic-related problems in medicine is to computerize and to implement enterprise computing in the work (R&D lifecycle) of medical staff. These problems include the search for effective appliances that can provide an increase of three major health indicators: quality of care, patient safety, economic efficiency of medical care. The key direction to increase these indicators is the use in hospitals of modern clinical comprehensive problem-oriented information systems provided with decision support mechanisms in the particular domain. The goal of this research is to develop the fundamentals of the comprehensive problem-oriented information system (complex computer appliance approach) for the scientific R&D lifecycle support in medicine. This new information system will focus on R&D lifecycle support in three major medicine-related directions: the processing of video/photo information (a formal model of the complex computer appliance for the informative stigmas highlighting in the images, in

particular, CT scans and video records); the processing of medical data (a formal model of the Question answering medical expert system) [Sure, 2009]; the natural language processing of medical corpora (a formal model of the information system for the processing of patient anamnesis, medical records and medical-related publications or literature).

The medicine-related R&D information systems background

The essence of the comprehensive problem-oriented information systems (or complex computer appliance approach) is to receive and to process the results of the R&D of medical specialists of various medical domain. This approach will improve the quality and performance of laboratory diagnostic work (tests, analyzes, etc.). The systems can be divided into two categories. *Diagnostic and advisory systems*. These systems are used to diagnose pathological conditions, including the prognosis and making of recommendations on treatment methods for diseases of various profiles and types. Such systems are a special case of expert systems (ES). The core part of such systems is a software package which can analyze certain source data and is able to replace narrow-profile specialists in problem situations. For the ES, “knowledge” are represented in the form of knowledge bases (formalized collections of facts and rules of logic in certain domain) [Hartmann, 2009], which can be changed and supplemented. The search for a solution can be carried out using logical (aimed at creating expert systems with logical models), mathematical (simulation, analytical) and heuristic methods. *Automated R&D workstation environments* [Palagin, 2018] for medical researches – RDWEm (for the automation of the entire technological process in the relevant medical domain and to provide the support in making medical decisions). RDWEm contain a wide range of management, analytical, visual and design tools to implement various techniques, both clinical and R&D purposes. The RDWEm system can distinguish three main components: medical, hardware and software. Medical support of the RDWEm system includes methods for implementing the selected range of medical tasks that are solved in accordance with the capabilities of the hardware and software parts of the system. Medical support also includes techniques, measurement methods and direct

measurement of physiological parameters, determination of methods and allowable limits for the impact of the system on the patient. By the hardware we mean methods for implementing the technical part of the system, including means for obtaining biomedical information, methods for implementing therapeutic effects and computer equipment. The software includes mathematical methods for processing biomedical information, methods for visualizing the results, algorithms and the actual software that implement the functioning of the entire system.

General formal model of the RDWEm

The generalized formal model of the RDWEm system is represented by some finite set of software (software modules) [Palagin, 2014][Palagin, 2012][Broy, 1997]:

$$RDWEm = \sum_{i=1}^{n} \prod_{RDWEm_i} ,$$

where: $i = \overline{1,n}, n \in N$; n – is a number of software modules in the system; $\prod_{RDWEm}$ – some software module of the RDWEm system. At the same time, a mapping $G_{\prod_{RDWEm}}$ of the integration of functions of the set of RDWEm’ software modules into the generalized (target) function of the system is implemented:

$$G_{\prod_{RDWEm}} : S_{RDWEm} \to F_{RDWEm},,$$

where: S_{RDWEm} – is a set of functions of a particular software module of the RDWEm; F_{RDWEm} – is a generalized (target) function of the RDWEm.

Results and Discussion

The development of modern technologies increasingly covers the field of intellectual activity and, especially, in the field of scientific research and development. The new class of Current Research Information Systems and related intelligent information technologies have arisen that support the main stages of the scientific research and development lifecycle [Palagin, 2018]. A distinctive feature of such systems and technologies is the possibility of their problem orientation to various types of scientific research and development by combining on a variety of functional services and adding new ones

within the cloud-integrated environment. A typical representative of RDWE systems is the Personal Research Information System (PRIS) – the RDWE class system for supporting research in the field of ontology engineering (the automated building of applied ontology in an arbitrary domain area as a main feature), scientific and technical creativity. In accordance with the information model of the RDWE class systems, a new fundamental model of RDWEm was proposed.

Conclusion

Formal models of the comprehensive information system for the medicine-related R&D support is proposed. At the next stage of research, functional models of the RDWEm system will be developed using a variety of problem-oriented software modules.

Bibliography

[Sure, 2009] Sure Y., Staab S., Studer R. (2009) Ontology Engineering Methodology. In: Staab S., Studer R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg

[Hartmann, 2009] Hartmann J., Palma R., Gómez-Pérez A. (2009) Ontology Repositories. In: Staab S., Studer R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg

[Palagin, 2018] Palagin O.V., Velychko V.Yu., Malakhov K.S. and Shchurov O.S. Research and development workstation environment: the new class of current research information systems. In CEUR Workshop Proceedings, volume 2139, pages 255-269. CEUR-WS.org, 2018.

[Palagin, 2014] Palagin A.V., Petrenko N.G., Velychko V.Yu. and Malakhov K.S. Development of formal models, algorithms, procedures, engineering and functioning of the software system “instrumental complex for ontological engineering purpose”. In CEUR Workshop Proceedings, volume 1843, pages 221-232. CEUR-WS.org, 2014.

[Palagin, 2012] Palagin A., Petrenko N., Velychko V., Malakhov K. and Tikhonov, Yu., About development of an ontological toolkit. Problemy programmirovaniâ, No. 2-3. Publishing house "Academperiodika", 2012. pp. 289-298. ISSN: 1727-4907 (print version), http://pp.isofts.kiev.ua/ojs1/article/view/82

[Broy, 1997] Broy M., Mathematical methods in system and software engineering, NATO ASI series f computer and systems sciences, Vol.158. Springer verlag, 1997. pp. 271–312.

Authors' Information

Vitalii Velychko – PhD, Senior researcher. V.M. Glushkov Institute of Cybernetics of NAS of Ukraine, e-mail: aduisukr@gmail.com
Major Fields of Scientific Research: Artificial Intelligence, NLP.

Kyrylo Malakhov – Junior researcher. V.M. Glushkov Institute of Cybernetics of NAS of Ukraine, e-mail: malakhovks@nas.gov.ua
Major Fields of Scientific Research: Artificial Intelligence, NLP.

Oleksandr Shchurov – Software Engineer. V.M. Glushkov Institute of Cybernetics Of NAS of Ukraine, e-mail: alexlug89@gmail.com
Major Fields of Scientific Research: Artificial Intelligence, NLP.

SECTION: SOFTWARE ENGINEERING APPLICATION DOMAINS

PRACTICAL APPROACH TO BUILDING A SPEAKER IDENTIFICATION SYSTEM

Yana Bielozorova

Abstract: The rules for building a speaker identification system was described. The approaches to creation such system was analysed. The ways of creating such system was proposed.

Keywords: speaker identification system, wavelet, coding and information theory, pattern analysis.

Introduction

Parameters of involvant's individual voice characteristics are the basis of every voice search system. Parameters of individual voice characteristics in modern automatic systems of speaker identification are usually determined on the basis of two main factors - the main tone frequency and spectral characteristics [1]. Efficiency of such systems mainly depends on the methods for determination of these factors and stability of spectral characteristics and the main tone frequency.

Speaker Recognition Evaluation (SRE) carried out by the National Institute of Standards and Technologies (USA) since 1996 is the most objective and competent source of information as to capabilities of the modern systems and methods of automatic speaker identification.

SRE allows to obtain data on real capabilities of identification methods and systems including in comparison with others and to choose the most perspective trends of development.

For the period of the last ten years the test results show considerable progress in this field. However, modern systems of automatic speaker identification are essentially behind the effectiveness of the speaker identification implemented by human acoustical apparatus.

Traditionally NIST publishes impersonal test results by means of which it is impossible to determine what identification method or system is found to be the best one. Absence (with few exceptions) of the test results data at the test participants' websites usually points out the poor results or restricted developments which are often performed for authorities ensuring the state security or for commercial purposes.

At the same time achieved level of developments and progress in outlined perspective directions in this field allow to proceed with the stage of collective developments of such systems within the frames of the EU countries.

Complex long-term investigations of scientific and research groups in this area in Ukraine allowed creating an experimental model of modern system for search of involvants in the voice database [2]. Further, there is presented a physical and mathematic model based on which the investigations were performed and experimental model was developed.

Scientific novelty of investigations and development

In view of importance and commercial nature of investigations and developments of leading designers in the field of voice characteristics identification, there is no enough information about detailed principles of modern system operation in open access publications. However, based on the analysis of the many years investigations, it is possible to suppose that the methodology of investigations and developments, described in above sections, is a complex new scientific approach to tasks of voice identification.

Scientific novelty of the approach under consideration is an application of complex of the following mathematical methods and technologies of digital processing of acoustic audio information:

- ✓ complex discrete two-parameter Morlet transformation;
- ✓ non-orthogonal discrete transformations with discrete frequency step of 1 Hz;
- ✓ application of the wavelet transforms maxima approach for determination of the features of frequency characteristics;
- ✓ determination of connection of frequency characteristics features with the main tone frequency and other parameters of voice characteristics;

✓ analysis of voice information in small time intervals – 10-30 ms.

Essence of the innovative nature of the project

Confidential nature of the modern most advanced developments in the field of identification of the voice characteristics creates essential problems both for accelerated evolution of these systems and for wide application in the intergovernmental voice search databases.

The present development, if the practical efficiency is proved, can be distributed in any EU country with appropriate version of language localization.

Innovative nature of the proposed project consists of posing the task for creation of prototypes of more open modern systems for identification of voice characteristics within the frame of voice databases for specialized structures.

The second important moment ensuring innovative nature of the project from our point of view is new technologies and models for processing of discrete voice information.

Bibliography

[Solovyov, 2014] V. Solovyov. Spectral analysis and speech technology (Russian) // V. Solovyov, O. Rubalskiy / Journal of Kiev National University T. Shevchenko, Military special sciences, - Kiev, Vol, 42, p. p. 145-151, 2014.

[Rubalsky et al, 2014] O. Rubalsky, V. Solovyov, A. Shablja, V. Zhuravel. New tools to identify the person for voice of Data (Russian) Protection and Security of Information Systems: Proceedings of the 3rd International Scientific Conference (city. Lviv, 05 - June 6, 2014). - Lviv: - p.p.. 110 - 112.

Authors' Information

Yana Bielozorova – Senior Lecturer 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

DRIVING EDUCATION SOFTWARE PROJECT.

Plotnikov Nikita, Rudnichenko Nickolay, But Natalya

Abstract: In the era of information technologies, there is a big problem where drivers lack the experience on the road, due to the lack of training and poor quality of the roads. Due to the lack of a comprehensive understanding of the traffic rules and the lack of practical driving experience, a small number of training during training. The urgent task is the design and development of complete software that has a modular structure for the efficiency and flexibility of adding functional, which allows to provide an interactive process of learning the rules of driving motor vehicles of different types and types based on the current legislation. The introduction of such a system will allow to automate the processes of learning and control of knowledge, minimizing resources and the time spent on the learning of the traffic rules.

Keywords: simulator, transport, virtual reality,

ITHEA Keywords: object-oriented programming, design, software architectures

1 Introduction

Due to the constant increase in population the demand for the vehicles has increased, the automation of the training process in traffic rules for the new drivers is necessary. [Plotnikov, 2019]. Nowadays, most driving schools do not provide a high quality of driver training, thus creating a negative situation throughout the service market. This state of affairs asserts opinion, that all driving schools are equally bad and undermines their credibility of professional driving schools. [Romas, 2015].

2 Summary of Contributions

Currently, there are various systems and software applications on the market to support driving learning processes. For a more targeted comparison, the three most similar and popular projects are titled: «3D Instructor», «City Car Driving», «Sim Exam» [Plotnikov, 2019]. All considered analogues are paid and not cross-platform; in addition, they do not support virtual reality and new updates. Based on the analysis, a comparative table for functional feature was created (Table 1.1), where the proposed system is designated as IDTS.

Table 1.1- Comparison by analogues

Name	Virtual reality	Modern graphics	Correctly implemented traffic rules	Generated a map	Optimization
3D Instructor	-	+	+	-	-
City Car Driving	-	+	+	-	+
Sim Exam	-	-	-	-	+
IDTS	+	-	+	+	-

3 Task and challenges

Due these issues, the purpose of this work is to create a system that combines both the ability to test practical knowledge and skills in virtual form and in the form of tests. Users can choose the complexity of the process and the different types of tasks. For example, in a virtual simulator, the user can choose different transport and location. To implement this development not only within the framework of a simple application, but using modern virtual reality technology, which would allow to feel this driving process as close as possible to the real one.

Proposed approach

4.1 The basic concept

The proposed software is a desktop application that allows the user to drive on a virtual machine on a map of the city in which he resides, pass a knowledge test or repeat any rules. [Rudnichenko, 2019]. The system is intended as a replacement for existing obsolete applications in driving schools or for private training with VR support.

The main options for using the system by the user are: passing tests in the system to improve theoretical knowledge and test them; making changes to location maps and cities; look up current traffic rules to increase your chances of passing the exam; entering your personal information to start an account; create and change the route of the movement to give the task to students or for themselves; change the account

information; delete your account (verified); online tracking of task accomplishment; storing the results of passed tests.

4.2 Use of virtual reality technology

In our generation, the use of additional tools to improve the perception of certain processes is become popularity. One of these technologies is augmented reality. Thanks to it, the user can easily learn from the room. This technology will aid driving school in saving up the funds in driving lessons and driving skills without real practice. Implementation is possible through both a full helmet and using the phone. It is also possible to implement simple controllers that would replace the steering wheel or other control unit of the machine.

5 Case study or implementation of results

As a software development platform using Unity, language C #, database management system SQLite and development environment Visual Studio Code. Plugin MapBox was selected for geolocation. Mapbox for Unity is a developer platform used in various industries to create custom applications that address maps, data and spatial analysis. Mapbox tools are blocks that support every part of the webpage and mobile map creation process [Jambruno, 2002]. In the application, the user is able to set up a further trip. You can choose the city where you want to ride. In this case, the city generates the location that he chooses on the map. With MapBox, you can easily generate a building, such as a mall, school, or park. The user is able to select any suitable transport, weather or any other modifier from the available ones. You can also specify the route and find out the number of violations committed while doing so. When approaching any street, the player is notified of its name (fig 1). There was a success in testing of connectivity of BOBOVR Z4 VR, the system supports the data transfer, which makes it possible to further the development of complete desktop experience to VR.

Figure 1. Demonstration of the generated location

6 Conclusion

The software is designed for users who want to consolidate or improve their knowledge and for use in driving schools where instructors will be able to control and influence the learning process of the user. Using created solution with other software modules and virtual simulators, aimed at testing the practical driving skills of vehicles, will provide a comprehensive approach to automating the process of learning traffic rules and reduce the time of conducting lectures and practical classes.

7 Further researches

In the further perspective of the development software, it is possible to supplement the project of the developed module by providing support for virtual reality mode with additional equipment and many audio and text chat modes for messaging. Also, it is planned to introduce virtual advertising in the form of billboards, to develop a strategy of selling licensesfor different driving schools, to implement a more flexible steering wheel model for driving a car.

Bibliography

[Plotnikov, 2019] Plotnikov N. S. Development of the concept of a virtual simulator project to support the process of learning traffic rules / N. S. Plotnikov, S.S. Koval, I.V. Suzanskiy // All-Ukrainian Scientific and Practical Conference of Young Scientists,

Graduate Students and Students: Abstracts. Mykolaiv: Black Sea National University named after Petro Mogily, 2019. pp. 99-102.
[Romas, 2015] Romas Y. Driving schools / pitfalls on the way to rights, 2015, (online), https://www.autocentre.ua/ua/avtopravo/politsiya/avtoshkoly-podvodnye-kamni-na-puti-k-pravam-52228.html
[Rudnichenko, 2019] Rudnichenko N.D. Software module for assessing and controlling theoretical knowledge of the rules of the road / N.D. Rudnichenko, S.S. Koval, I.V. Suzanskiy // Materials of the Sixteenth All-Ukrainian Conference of Students and Young Scientists "Informatics, Information Systems and Technologies", Odessa: 2019. pp. 29-31.
[Jambruno, 2002] Jambruno M. 3D graphics and animation / M. Jambruno. - M.: Williams, 2002, pp. 640 c.

Authors information

Plotnikov Nikita – Odessa National Polytechnic University; student, Odessa, e-mail: supernikita45@gmail.com
Major Fields of Scientific Research: game development, virtual and augmented reality

Nickolay Rudnichenko – Odessa National Polytechnic University; assistant professor IT department: 65001, Odesa; e-mail: nickolay.rud@gmail.com
Major Fields of Scientific Research: data mining, software programming and design, machine learning

Natalya But – Odessa National Polytechnic University; post graduate IT department: 65001, Odesa; e-mail: murrderra@gmail.com
Major Fields of Scientific Research: artificial intelligence, software design

USING WEB-PARSING IN FORECASTING THE BEHAVIOUR OF FOREIGN EXCHANGE MARKET

YEVHEN MEDYANIK

Abstract: With the increase and demand of informational technologies in worldwide society it is being created a big data in different areas of human activities, especially in economical sector. Nowadays enormous size of information related to foreign exchange is available in World Wide Web and one of the most important issues is how to analyze and build effective software for building stable and reliable forecast. The aim of this article is to report about how to create software for these targets by the way of using web-parsing.

Keywords: C++, TinyXML, LibTidy, foreign exchange market, forecasting.

ITHEA Keywords: D Software.

Introduction

Only for the last 15 years the currency trading has opened up to the world. Obviously the popularity of internet has helped for making this market to be available for bigger part of the citizens that didn't used to retail on Forex. Nowadays nearly 99% of all transactions in the world were carried online. With this enormous scale it's important to be more prepared before start to work in this economic sphere. Taking into consideration all of these issues, in this article will be suggested one of the possible solutions of how to develop software for more accurate forecast using web-parsing and mathematics for continuous time models, in relation to the typical user.

Web-Parser on C++ base with relevant tools

Web-parser is one of the most powerful tools for get information related to foreign exchange market from internet web-pages for further mathematical calculations and building stable forecast. Taking this into consideration, the software needs to provide the high level of security and reliability, for this targets the best tool is C++ programming language. Of course, if we were using other programming languages such as Python, Perl or C# we would be able to achieve the destination a little bit faster, because of

existing framework for all of this programming languages, on the other hand we would be lost the high level of speed of work and security. In this situation, it’s needed to use libraries of C++, such as CURL, TinyXML and libTIDY, which can provide proper operation of all system.

CURL

This library provides the download necessary web pages for further transferring information in other parts of system. The main advantage of this library is supporting a bunch of different web-protocols such as HTTP, HTTPS, FTP, FTPS, DICT, GOPHIS and SSL certificate. From the other advantages of this library it’s important to highlight the high speed of downloading necessary information in developer’s console, which process the HTML code.

```
#define CURL_STATICLIB
#include <iostream>
#include <string>

#include "curl/curl.h"

static size_t my_write(void* buffer, size_t size, size_t nmemb, void* param)
{
    std::string& text = *static_cast<std::string*>(param);
    size_t totalsize = size * nmemb;
    text.append(static_cast<char*>(buffer), totalsize);
    return totalsize;
}

int main()
{
    std::string result;
    CURL* curl;
    CURLcode res;
    curl_global_init(CURL_GLOBAL_DEFAULT);
    curl = curl_easy_init();
    if (curl) {
        curl_easy_setopt(curl, CURLOPT_URL, "    ");
        curl_easy_setopt(curl, CURLOPT_WRITEFUNCTION, my_write);
        curl_easy_setopt(curl, CURLOPT_WRITEDATA, &result);
        curl_easy_setopt(curl, CURLOPT_VERBOSE, 1L);
        res = curl_easy_perform(curl);
        curl_easy_cleanup(curl);
        if (CURLE_OK != res) {
            std::cerr << "CURL error: " << res << '\n';
        }
    }
    curl_global_cleanup();
    std::cout << result << "\n\n";
}
```

Figure 1. C++ code for download web page using CURL library

TinyXML

Library TinyXML analyzes and parsers the XML document and builds Document Object Model that can be read, modified and saved. In dedication to economic sector theme it will save all necessary information in order to further calculation in continuous time models of market.

libTIDY

This library creates the console application to adjust the non-working hypertext markup language document, and searching for potential mistakes and errors in web-access. It enhances the layout and appearance of resulting layout created earlier in TinyXML library.

Mathematic continuous time models

All received data is needed to be entered in formulas of continuous time models:

$$b^{01} = r^{01} + \sigma^{01}\lambda^{0}.$$

Figure 2. Drift coefficient

$$\lambda^{0} = \frac{b^{01} - r^{01}}{\sigma^{01}},$$

Figure 3. Domestic market price of risk

$$r^{\mu\nu} = r^{\mu} - r^{\nu}.$$

Figure 4. Compound interest rate

$$\frac{dS_t^{01}}{S_t^{01}} = b^{01}dt + \sigma^{01}dW_t, \qquad S_0^{01} \text{ is given.}$$

Figure 5. Brownian Motion with drift

$$\frac{dS_t^{10}}{S_t^{10}} = b^{10}dt + \sigma^{10}dW_t = \left(r^{10} + \sigma^{10}\lambda^1\right)dt + \sigma^{10}dW_t,$$

Figure 6. The equations describing market evolution from the viewpoint of euro investments

Conclusion

The essence of this approach consists in development potential modern system dedicated to the building forecast in foreign exchange market and make the work on other economic markets easier, by the way of combination of parsing information and further mathematical calculations

Bibliography

Jim Brown. FOREX TRADING The Basics Explained in Simple Terms. 2011. pp. 28-30,
Alexander Lipton. Mathematical Methods for Foreign Exchange A Financial Engineers Approach. pp. 215-220.

Authors' Information

Yevhen Medyanik – Odesa National Polytechnic University, Odesa, Ukraine.
e-mail:medyanik. jenya@gmail. com
Major Fields of Scientific Research: Informational Technologies, Software engineering

DECISION SUPPORT SYSTEM CONCEPT FOR DATA MINING METHODS SELECTION.

Rudnichenko Nickolay, Gezha Nikita, Tishchenko Sophia

Abstract: This paper describes the concept of a decision support system aimed to solve the machine learning (ML) model selection problem in the data mining process. The problem with traditional methods of selecting the best ML model for a task is that they rely on resource-intensive testing and comparison. The described system solves it by ranking machine-learning models given the task specifics, metric of interest and dataset statistics. The system internally utilizes an ML model to learn dependencies between dataset statistics and task specifics with the best ML model for the task. This paper covers the main problems of the data mining process, the necessity of the proposed system, the requirements for such system, chosen architecture behind the proposed system, as well as its advantages and disadvantages, and the potential future improvements for it.

Keywords: Decision support system, data mining, machine learning.

ITHEA Keywords: object-oriented programming, design, software architectures.

Introduction

At the current time, the data mining process (DM) has gained a widespread use, and it is utilized in business, finances, medicine and others. It is based on inspecting and analyzing large volumes of data, and as a result finding dependencies, trends and relationships amongst the data. The DM process itself is composed several stages, such as data selection, data cleaning, transformation, building a model of the data (frequently using ML methods), model evaluation and knowledge extraction [1]. Since there is a choice between several ML algorithms during the model building stage, quality of which is dependent on the specific task [2], then to select the most desirable algorithm an experimental comparison should be conducted. However, the larger the amount of data is, and the more models are being compared, the more time and memory consuming the testing process is.

To solve the stated problem, authors propose a decision support system, which will allow to obtain a comparative rating of the ML methods for the given task. The list is generated by considering the task details, extracted dataset statistics per column and for target variables, and the desired metric for the model to excel at. Compared to the usual experimental comparison of the models, the system will provide a rating in a comparatively short amount of time and using much less space.

The key component of the system is the ML model, trained on predicting the most likely best-suited model out of several for the given task. The described system works with the supervised learning tasks, however it is possible to further its functionality to with unsupervised learning.

The main point of such system is that it will allow for quick evaluation of several ML models for a given task and dataset, allowing experts to either select the possibly best performing one for the task, or select a few of the highly rated models and decide on the best one through the experimental comparison.

Requirements for the proposed system

Since the data in the DM process can be stored in different formats, it is essential for the system to be able to work with datasets stored in different file formats, such as .CSV or .XLS, as well as databases of SQL and NoSQL structures.

In addition, when working with large amounts of data, the size of the dataset can be too large to extract statistics from it quickly and without overusing the available memory. To circumvent this issue, the system should support selecting only a sample of the original dataset.

Upon importing the dataset, user will be displayed the dataset overview and the extracted statistics. In addition, user must have the ability to specify which of the dataset columns to consider during analysis, as well as which of them are target variables. Depending on the selected target variables, the task specifics must be determined (classification or regression, amount of labels, etc.).

Architecture of the proposed system

It was decided to split the system into several modules: user interface, dataset loading and selection, statistics extraction module and rating generation module. General process of working with the system is displayed on the figure 1.

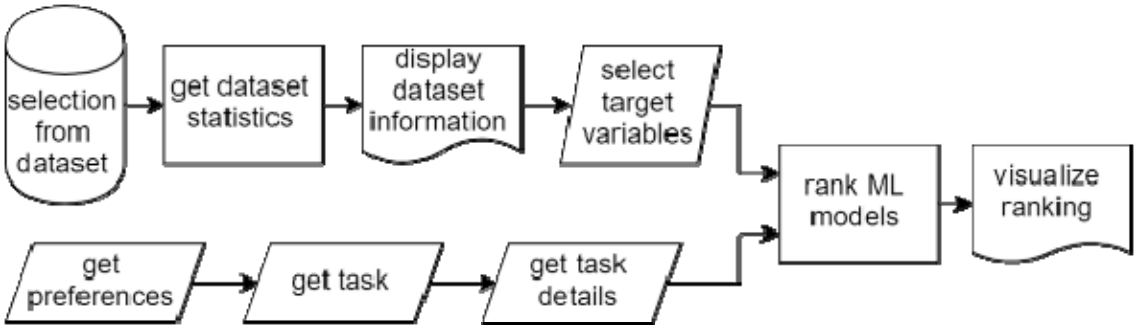

Figure 1. Process of working with the system

In the rating generation module, it was decided to use a ML model based on a deep neural network. As the input data, the model uses meta-statistics (like dataset features amount to size proportion), task details (like the desired metric), and statistics for each feature columns (each column's statistics are encoded as a vector). To allow for consideration of all features and their effect on the best model for the task, it was decided to use an attention mechanism in the neural network. As the output data, system generates a value between 0 and 1, symbolizing relative quality of the ML model for the task, which allows to for their ranking. This approach allows the system to learn complicated relations between which models are good for the task and the dataset statistics. It also allows for training using stochastic gradient descent. During training, the target data are the experimentally obtained model ratings. The biggest disadvantage of this method is the long process of collecting the training data, as it is necessary to experimentally obtain ML model ratings for each dataset.

Conclusion

The developed decision support system concept allows to reduce the time and computational resources required to compare different ML models, by means of creating a rating of models based on the dataset and the task. The advantage of the described system is the ability to learn complicated dependencies, as well as support for different file formats and databases.

Bibliography

[Chudán, D., Svátek, V., Kuchař, J., & Vojíř, S., Data Mining from Open Fiscal Data. Journal Acta Informatica Pragensia, Vol.7, Issue 1. Springer-Verlag, 2018. pp. 58-73. doi: 10.18267/j.aip.114.

[Kovalerchuk B., Vityaev E., Data Mining for Financial Applications., Journal Data Mining and Knowledge Discovery Handbook. Boston: Springe, 2005. pp. 1153-1169. doi: 10.1007/978-0-387-09823-4_60

Authors' Information

Nickolay Rudnichenko – Odessa National Polytechnic University; assistant professor IT department: 65001, Odesa; e-mail: nickolay.rud@gmail.com

Major Fields of Scientific Research: data mining, software programming and design, machine learning

Gezha Nikita – Odessa National Polytechnic University; student of IT department: 65001, Odesa; e-mail: znongz1@gmail.com

Major Fields of Scientific Research: data mining, software programming and design, machine learning

Tishchenko Sophia – Odessa National Polytechnic University; student of IT department: 65001, Odesa; e-mail: s5187565@gmail.com

Major Fields of Scientific Research: data mining, software programming and design, machine learning

THE MAKING DECISION IN THE ALARM SYSTEM OF THE PASSIVE MONITORING OF SPECIAL OBJECTS.

Sergiy Mostovyi, Vasyl Mostovyy

Abstracts: A flow-like mathematical model is proposed that represents a seismic field in the form of a convolution of a point process of the moments of occurrence of seismic events with signals that describe each such event. Signals are complex compositions of elastic waves of various types that arise both at the beginning of this process and during its propagation up to the moment of being recorded by a sensor. The concept of a generalized seismic signal is introduced. In this concept, it is taken into account that due to the natural background, which always complicates the analysis of the moment of occurrence of an event, the concept of a conditionally physically realizable signal is introduced. The prehistory of such a signal, before its appearance above the background level, is represented by a stochastic process with an indefinite time of signal appearance, but satisfying certain conditions. These conditions consist of the assumptions that its energy in the prehistory, before its excess of the background level, nowhere exceeded the energy of such a background, and that its derivative is a smooth function. The generalized seismic signal, depending on the values of the matrix of free parameters, describes various types of seismic waves. The matrix of free parameters of this model determines physically meaningful, both kinematic and dynamic, parameters of each event. In the model, it is taken into account that a specific additive micro-seismic noise accompanies the registration. At each registration point, in aggregate, we obtain a three-component vector stochastic process. For the data obtained in the experiment, on the set of alternative models, by mean of a minimization of a criterion, the optimal one is chosen [Mostovyi et al, 2008].

Keywords: conditionally physically realizable signal, point stochastic process, generalized elastic seismic signal.

Introduction

The modern development of the research base in seismology requires the creation of new mathematical models, both a seismic wavefield in general and the mathematical models of seismic signals. As well and timely noted in a recently published article, which shows the inconsistency of many wave models, a long time, and traditionally used till

now a day. We believe that in this row and an outdated view of the seismic event, as such. In our work, we want to emphasize that a seismic event is rather long in time process that can only be described as a flow containing many local partial events. The seismic event formation process is one, which is the amplification of the tension and the attenuation of energy accompanied by the generation of weak events, as a result of a weakening of its energy. The sequence of such weak seismic events is modeling by the flow of many signals, the energy of which is commensurate with the background energy.

This work is aiming at investigating the flow of weak events; such events are quite frequent and carry information about rare intense seismic events. The proposing mathematical models in this contribution are using in the monitoring of the state of both a natural object and human-made one. Each potential source of stress and their unloading itself migrates in space, and these processes of overstrain and unloading occur spontaneously in time. These sources generate a seismic event. All events are associating with a particular source are drawing by a chain in time. It is unknown when events are beginning, and it is unknown when they will end. The very nature of these events is random. The discrete occurrence of these events in time is a stochastic flow. From this point of view, the flow model seems entirely natural. Investigation of the seismic signals flow, comparable in energy with background noise, requires from the mathematical models describing the process under study certain conditions. Due to the high level of interference, we can’t talk about the specific signal arrival time, but only about the probability confidence interval of the signal beginning. Considering signal flow models, mathematical models of individual signals should take into account the prehistory containing the previous signals of the investigated flow.

The classical mathematical and empirical models of seismic signals do not allow these conditions to be taking into account. In this paper, a new approach is proposing in the mathematical modeling of the seismic wavefield and seismic signals that take into account the above circumstances. The authors' many years of experience in mathematical modeling of active and passive monitoring of seismic and acoustic fields, both natural and human-made objects, allowed to formulate the problems are setting.

The classical mathematical and empirical models of seismic signals do not allow these conditions to be taking into account. In this paper, a new approach is proposing in the mathematical modeling of the seismic wavefield and seismic signals that take into account the above circumstances. Within this conceptual framework, the notion of a generalized seismic signal is introducing, the mathematical model of which takes into account the conditions considered above. It is under consideration and to offer a new concept of mathematical modeling of the seismic wavefield.

Mathematical model of the generalized seismic signal.

The mathematical models of signals used in seismological practice can be called one family, generated by a vector of free parameters. The signal's free parameters vector variety must produce the point’s migration into the n-dimensional space. It gives all the conceivable models used in practice. These points are vectors of free parameters of mathematical models of seismic signals. Each signal is generating by different point locations of the hyperplanes of this space. Choosing a specific point in space, we can obtain a class of signals widely used in seismology practice or other areas of signal processing. We settled on the dimension of this space equal to eight.

In the analysis of the seismic wavefield, it is necessary to introduce an additional condition for the generalized seismic signal, namely, the concept of conditional physical feasibility of this signal. This circumstance appears because of a physically realizable of the signal must satisfy the causality condition. This condition is that the signal has an arrival point of origin and then develops in time. If the signal flow is recording against a background of microseismic noise, then we cannot determine the beginning of this signal. This is due to the microseismic noise. It is possible to eliminate this disagreement of our model with the real seismic background by replacing the model of the physically realized signal with a conditionally physically feasible signal, which is not much different in energy from it. The condition is that we consider the moment when the signal appears undefined before it appears above the level of the microseismic background. The second condition for physical feasibility is that the signal energy must be finite. This condition for

the conditionally physically realized signal in our model is preserving. Because of the interval of its indefinite development, we demand from its model that its energy does not exceed the energy of the natural background.

Because on the interval of its indefinite development, we demand from its model that its energy does not exceed the energy of the natural background noise. Thus, the energy of the physically feasible signal and the conditionally physically feasible signal differ by not more than the value of the background energy level from the beginning of the observation to the appearance of the signal above the background level.

Conditionally physically feasible signal or a seismic wave $F_G(t, \mathrm{P}_G)$, which will be used, as a generalized seismic signal or seismic wave.

$$F_G(t, \mathrm{P}_G)$$

Where:

$$P_G \in \mathbb{R} \times [-A, A] \times [0,1) \times [\Omega_1, \Omega_2] \times \mathbb{N} \cup \{0\} \times \mathbb{N} \cup \{0\} \times \mathbb{R} \times \mathbb{R}_+ \setminus \{0\}.$$

$$F_G(t, P_G) \in \mathbb{R} \times \mathbb{R} \times [-A, A] \times [0,1) \times [\Omega_1, \Omega_2] \times \mathbb{N} \cup \{0\} \times \mathbb{N} \cup \{0\} \times \mathbb{R} \times \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}.$$

Below we can see the transposed vector P_G of free parameters of the model, which contains physically meaningful parameters.

$$P_G^T(\tau, \alpha, \beta, \omega_0, \gamma, \gamma_2, \psi, \sigma).$$

Where:

τ is the mathematical expectation in the cumulative probability function of a normal distribution. Simultaneously, this is the moment when the signal appears above the noise level.

α is the amplitude of the seismic signal.

β is the exponent in the decaying exponent that enters the envelope function of our signal.

ω_0 is the carrier frequency of our signal.

ψ is the phase of the carrier frequency of our signal.

$\beta, \omega_0, \gamma, \gamma_2, \psi$ are the parameters that determine the shape of the seismic waves.

Let us afford a generalized signal in the form of a product of two functions. The first function is responsible for the waveform. The second function is responsible for the parameters of the carrier frequency of the simulated signal.

$$F_G(t,\tau,\alpha,\beta,\omega_0,\gamma,\gamma_2,\psi,\sigma) = Envel(t,\tau,\beta,\gamma,\gamma_2,\sigma)\cdot Carrier(t,\tau,\alpha,\omega_0,\psi).$$

Where

$$Carrier(t,\tau,\alpha,\omega_0,\psi) = \alpha\cdot\sin\left[\omega_0\cdot(t-\tau)+\psi\right].$$

And

$$Envel(t,\tau,\beta,\gamma,\gamma_2,\sigma) = Envel_1(t,\tau,\beta,\gamma,\gamma_2,\sigma)\cdot Fprob(t,\tau,\sigma).$$

Here

$$Envel_1(t,\tau,\beta,\gamma,\gamma_2,\sigma) = (t-\tau)^{\gamma}\cdot\exp\left[-\left[\beta\cdot(t-\tau)^{\gamma_2}\right]\right].$$

For the model of a physical realizable signal, the $Fprob(t,\tau,\sigma)$ is a step function.

For the model of a generalized signal, the $Fprob(t,\tau,\sigma)$ is a cumulative normal distribution function.

We have to remember that the closer the value of σ to zero, the closer the function $Fprob(t,\tau,\sigma)$ to the step function.

Ergo:

$$Envel(t,\tau,\beta,\gamma,\gamma_2,\sigma) = \exp\left(-\beta\cdot(t-\tau)^{\gamma_2}\right)\cdot\left(\frac{\sqrt{2}\cdot erf\left[\tau\cdot\sqrt{\frac{1}{2(\sigma)^2}}-t\cdot\sqrt{\frac{1}{2(\sigma)^2}}\right]}{4\cdot\sigma\cdot\sqrt{\frac{1}{2(\sigma)^2}}}-\frac{1}{2}\right)\cdot(t-\tau)^{\gamma}$$

Here $erf(\cdot)$ is well-known error function.

Conclusions: The space of free parameters of the model characterizes the state of the investigated object [Mostovyi and Mostovyy, 1996]. We consider the subset of free

parameters included in the model discretely, as a set of alternative hypotheses. Each of the tested hypotheses is evaluated for the quality of agreement with the observed data within the framework of the selected criterion, in the subspace of the continuous model parameters. Of all the hypotheses, the one that has the highest quality is selected. If the vector of free parameters is in the stationary region, then this does not cause alarm. If informative parameters of the object go beyond the stationary area, a decision is making on the alarm [Mostovyi and Mostovyy, 2005]. Choosing a point in space, we can obtain a class of signals widely used in practice, in seismology, in the exploration seismology and in the other areas of signal processing. The flow of this type of signal is a set of vectors of free parameters. Many of these signals represent a superposition of the streams of each of these signals. The result of assessing the state of an object is a matrix, which composed of estimates of the parameters of each of the vectors in this stream. The matrix is a stack of all vectors of free parameters that participated in this stream. The alarm decision is taken immediately as soon as the point, which is defined by the stacked matrix in n-dimensional space, goes beyond the bounds of the allowable parallelepiped in this space. The dimension of this space is equal to the length of the stacked matrix of free parameters. In the space of the informative parameters of the model is represented by a matrix of the state object evaluation criterion based decision about alarm [Mostovyi et al, 2001], [Mostovyi and Mostovyy, 2016], [Mostovyi and Mostovyy, 2014].

References

[Mostovyi et al, 2008] S. Mostovyi, V. Mostovyy, M. Panchenko. Seismic Signal and Microseismic Background (Mathematical Models and Methods of Estimations), Geoinformatica, 2008, Vol. 1, pp. 28-33.

[Mostovyi and Mostovyy, 1996] S. Mostovyi, V. Mostovyy. Seismic Signal Analysis and Prediction the Seismic Events //Consel de l'Europe. Journees Luxembourgeoises de Jeodynnamique. 79th and 80th session, 1996. p.40-45.

[Mostovyi and Mostovyy, 2005] S. Mostovyi, V. Mostovyy. Monitoring problem in sphere of dynamic object identification // KDS 2005, Proceedings, V. 1, FOI-COMMERCE, Sofia, 2005, p 256-262.

[Mostovyi et al, 2001] S. Mostovyi, et al. Model and Experimental Studies of the Identification of Oil/Gas Deposits Using Dynamic Parameters of Active Seismic Monitoring, Geophys. J., 2001, Vol. 20, pp. 895-909.
[Mostovyi and Mostovyy, 2016] S. Mostovyi, V. Mostovyy. Mathematical model of seismic signal, as a flow of physically non realizable single seismic waves, Geophys. J., 2016, Vol. 38(5), pp. 164-169.
[Mostovyi and Mostovyy, 2016] S. Mostovyi, V. Mostovyy. Estimation of parameters of seismic waves, Dokl. Acad. Nauk Ukr., 2014, Vol. 2, pp. 118-123.

Authors' Information

Sergiy Mostovyi – Doctor of Physical and Mathematical Sciences, Professor, National Aviation University; e-mail: sergey.mostovoy2110@gmail.com
Major Fields of Scientific Research: Mathematical models of monitoring, signal processing.
Vasyl Mostovyy – Doctor of Physical and Mathematical Sciences, Senior Researcher, S. Subbotin Institute of Geophysics National Academy of Science of Ukraine; e-mail: vasyl.mostovyy@gmail.com
Major Fields of Scientific Research: Mathematical models of monitoring, signal processing.

Тези доповідей

Міжнародної науково-практичної конференції

Інженерія програмного забезпечення

Технічні редактори – Markov K., Чебанюк О.В.

Матеріали друкуються за редакцією авторів.

Підписано до друку.