SUMMARY

“INFORMATION SYSTEMS AND NETWORKS” BULLETIN 2017/I

INFORMATION SYSTEMS, NETWORKS AND TECHNOLOGY

1. Пастернак І. І. Діагностика та налагодження вузлів корпоративної мережі кіберфізичних систем

DIAGNOSIS AND ADJUSTMENT OF UNITS OF BUSINESS NETWORK OF CYBER PHYSICAL SYSTEMS

Iryna Pasternak

Department of electronic computer machines, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail: pas_irusj@ukr.net

In malfunction of network devices it is recorded the fact of a failure, determined by its location and type. Message was transmitted about a fault, device or node disconnected and replaced by a reserve or adjusted. Continuous monitoring of the work in local network, which is basis of any corporate computer network is required to support in working condition. Given the importance of this function it is often separated from other functions and implemented by special means. This separation of control functions is useful for small and medium sized networks. The use of independent controls helps network administrators identify problem areas and network devices, and their disconnection or reconfiguration they can perform manually in this case. Also, in some cases, standalone tool can diagnose the problem and fix it.

This article describes the process of monitoring network which is usually divided into two stages – monitoring and diagnosis. On the stage of monitoring it is analyzed, that a procedure for the collection of primary data on the network CFS is performed: statistics on the number of circulating frames and packets of various protocols in the network, state of port hubs, switches and routers, and so on. As well the diagnosis stage is executed which means more complex and intellectual process of understanding gained information during monitoring, comparing it with data obtained before and making assumptions about possible causes of slow or unreliable network work. The task of monitoring the network is solved using software and hardware meters, testers, network analyzers, built-in monitoring and communication devices, system management agents. Also, diagnosis of CFS task requires more active participation of a person and use of sophisticated tools such as expert systems, accumulating experience of many network professionals.

Obviously it is much easier to prevent network problems CFS than correct problems that have arisen. Diagnosis of servers and network devices CFS will advance to learn about potential problems and prevent their occurrence. Tracking functioning of network and maintaining its background work, besides administrator can provide users with accurate information, which is sometimes the wrong idea about a frequency of various malfunctions.

It is analyzed tools for monitoring and diagnosing networks CFS and they are grouped into several large classes:

– Network management system – centralized software systems that collect data on components and communication devices of network and traffic data in the network. These systems do not only monitor and analyze, but also perform automatic or semi-automatic mode of network management: enabling and disabling port devices, change of settings of table bridges, switches and routers, etc.

– System management. System management means often perform functions similar to the management systems, but in relation to other objects.

– Embedded systems of diagnostics and control. These systems are carried out the form of hardware and software modules that are installed in communication equipment as well as software modules built into operating system. They serve as diagnosis and management of only one device, and this is a main difference from centralized control systems.

– Protocol analyzers. They are software or hardware and software systems that are limited in contrast to the management systems only by functions for monitoring and analyzing network traffic.

– Protocol analyzers provide some logical conditions to capture individual packets and perform full decoding of captured packets shown in user-friendly form nesting packet protocols at different levels in each one of deciphering the contents of individual fields of each packet.

– Equipment for the diagnosis and certification of cable systems. Conventionally, this equipment can be divided into four main groups: network monitors, devices for certification of cable systems, cable scanners and testers.

– Expert systems. This type accumulates human knowledge about identifying the causes of abnormal networks and possible ways to bring the network for functioning.

– Multifunctional devices of analysis and diagnostics. Due to proliferation of local networks it was necessary to develop low-cost handheld devices that combine functions of several devices: protocol analyzers, cable scanners and even some opportunities to network management software.

The article presents an algorithm testing computer network in a specialized environment of unit diagnostics of CFS corporate network. Also, algorithm of diagnosing of CFS corporate computer network using modules inspection services can be used as basis for more complex CFS corporate computer networks with many nodes.

Key words: corporate network, client, server, cyber physics system.

1. Chris Giametta. Pro Flex on Spring, 2009. – p.445. 2. Оберг Роберт Дж. Технология COM + Основы и программирование = Understanding and Programming COM+: A Practical Guide to Windows 2000 First Edition. – М.: Вильямс, 2000. – С. 480. 3. Липаев В. В. Обеспечение качества программных средств. Методы и стандарты. – М.: Синтег, 2001. – С. 246. 4. Макгрегор Дж., Сайкс Д. Тестирование объектно-ориентированного программного обеспечения. – К.: Диасофт, 2002. – С. 432. 5. Тамре Л. Введение в тестирование программного обеспечения. – М.: Вильямс, 2003. – С. 368. 6. Татарчук М. І. Корпоративні інформаційні системи: навч. посіб. – 2005. – С. 245. 7. Мухамедзянов Н. Java. Server applications. – М.: Солон–Р, 2003. – С. 267. 8. Орфалі Роберт, Ден Харкі JAVA and CORBA in client server applications.

2. Lupenko S. A., Orobchuk O. R., Vakulenko D. V., Sverstyuk A. S., Horkunenko A. B. Integrated onto-based information analytical environment of scientific research, professional healing and e-learning of Chinese image medicine

INTEGRATED ONTO-BASED INFORMATION ANALYTICAL ENVIRONMENT OF SCIENTIFIC RESEARCH, PROFESSIONAL HEALING AND E-LEARNING OF CHINESE IMAGE MEDICINE

S. A. Lupenko, O. R. Orobchuk
Ternopil Ivan Pul'uj National Technical University, 56, Ruska str., Ternopil
D. V. Vakulenko, A. S. Sverstyuk, A. B. Horkunenko
I. Horbachevsky Ternopil State Medical University, 1, m.Voli, Ternopil

The timeliness of development and formulation of general requirements and architecture of integrated onto-based information analytical environment of scientific research, professional healing and e-learning of Chinese image medicine as a promising component of integrative medicine is substantiated in the article. Scientific evidence-based integrative medicine is typical for conventional medicine; but unlike the conventional, the integrative medicine synthesizes experience and engages all the best achievements of ancient medicine and contemporary Western one. The integrative medicine is not a new field of medicine; it is its new paradigm that facilitates the new quality of healthcare services.

Traditional Chinese medicine experienced a number of comprehensive clinical researches, theoretical scientific studies and relevant information analysis means were developed (ontologies, expert systems, grid systems), there is no such research and significant information analysis means for Chinese image medicine. The development of integrated onto-based information analytical environment of scientific research, professional healing and e-learning of Chinese image medicine is aimed to ensure the effective organization and coordination of existing professionals of Chinese image medicine, its scientific researchers, people who study Chinese image medicine and the establishing of modern intellectualized information means and resources in traditional, complementary and integrative medicine on a national and worldwide basis. The developing information environment will enable on a high scientific, technological and infrastructure levels data collection and automated statistical and intellectualized analysis of treatment results by means of Chinese imagine medicine; will facilitate the creation of a unified database of theoretical, experimental and clinical research in integrative medicine.

Onto-basis of the developed integrated information environment will help to unify, standardize the technologies of information submission (data and knowledge) in traditional Chinese medicine and Chinese image medicine that will make it possible to solve the problem of semantic heterogeneity of poorly structured and hard formalized knowledge of Chinese image medicine because the use of ontologies eliminates subjective factors, polysemantics, fuzziness of images and concepts that are explicitly or implicitly operated by complementary medicine specialists in diagnostic and therapeutic decision-making. In addition, the developed onto-based environment allows maintaining the necessary level of integration and sustainability of knowledge and data in Chinese medicine for different information technology and systems, and also the possibility of multiple reuse of knowledge for various information systems and applications.

Also the requirements and general architectural components of integrated onto-based information analytical environment of scientific research, professional healing and e-learning of Chinese image medicine were developed in the study, in particular for the information system of professional healing Image Therapist, a knowledge-base of Chinese image medicine, expert system for diagnostic and therapeutic decision-making support, information system of e-learning of Chinese image medicine, information system of research in Chinese image medicine.

Key words: ontology, integrated information analytical environment, Chinese image medicine.

1. Barnes PM, Bloom B, Nahin R. CDC National Health Statistics Report #12. The Use of Complementary and Alternative Medicine in the United States. Findings from the 2007 National Health Interview Survey (NHIS) conducted by the National Center for Complementary and Alternative Medicine (NCCAM) and the National Center for Health Statistics. December 2008. http://nccam.nih.gov/ news/camstats/2007/cam-survey_fs1.htm Accessed November 5, 2011, вільний (дата звернення: 23.11.2016) 2. Ananth S. 2010 Complementary and Alternative Medicine Survey of Hospitals: Summary of Results.Health Forum (American Hospital Association) and the Samueli Institute. September2011.http://www.siib.org/news/2468-SIIB/version/default/part/AttachmentData/data/ CAM%20 Survey%20FINAL.pdf Accessed December 11, 2011. 3. Guarneri E., Horrigan B., Pechura C. The Efficacy and Cost Effectiveness of Integrative Medicine: A Review of the Medical and Corporate Literature. Explore: The Journal of Science and Healing. 2010; 5:308–312. 4. Maizes V., Rakel D., Niemiec C. Integrative medicine and patient-centered care. Explore: The Journal of Science and Healing. 2009;5(5):277–89. 5. Horrigan B. What is Integrative Medicine? Published by The Bravewell Collaborative; 2010.http://www.bravewell.org/ integrative_ medicine/what_is_IM Accessed November 6, 2011, вільний (дата звернення: 23.11.2016). 6. Wang Y., Zhonghua Y., Jiang Y., Liu Y., Chen L., Liu Y. A Framework and Its Empirical Study of Automatic Diagnosis of Traditional Chinese Medicine Utilizing Raw Free-text Clinical Records.Journal of Biomedical Informatics. 2012;45(2):210–223. doi: 10.1016/j.jbi.2011.10.010. 7. Wang H. A computerized diagnostic model based on naive bayesian classifier in traditional chinese medicine. Proceedings of the 1st International Conference on BioMedical Engineering and Informatics (BMEI '08); May 2008; IEEE; pp. 474–477. 8. Wang X, Qu H, Liu P, Cheng Y. A self-learning expert system for diagnosis in traditional Chinese medicine. Expert Systems with Applications. 2004;26(4):557–566. 9. Huang M-J, Chen M-Y. Integrated design of the intelligent web-based Chinese Medical Diagnostic System (CMDS)—systematic development for digestive health. Expert Systems with Applications.2007;32(2):658–673. 10. Mao Y, Yin A. Ontology modeling and development for Traditional Chinese Medicine. Proceedings of the 2nd International Conference on Biomedical Engineering and Informatics (BMEI '09); October 2009; IEEE; pp. 1–5. 11. Lukman S, He Y, Hui SC. Computational methods for traditional Chinese medicine: a survey. Computer Methods and Programs in Biomedicine. 2007;88:283–294. 12. Chen H., Wang Y., Wang H. et al. Towards a semantic web of relational databases: a practical semantic toolkit and an in-use case from traditional Chinese medicine. Proceeding of the 5th international conference on The Semantic Web (ISWC '06); 2006; pp. 750–763. 13. International program of scientific research in Chinese image medicine and Zhong Yuan Qigong for 2017-2023 [Electronic source]. On-line mode: https://kundawell.com/ru/mezhdunarodnaya-programma-nauchnykh-issledovanij-kitajskoj-imidzh-meditsiny-i-chzhun-yuan-tsigun-na-2017-2023-god, free access (date of access: 22.11.2016). 14. WHO strategy for traditional medicine for 2014-2023 [Electronic source] – 2013. – 72 p. – On-line mode: http://www.who. int/medicines/publications/traditional/ trm_strategy14 _23/ru/, free access (date of access: 20.11.2016). 15. State enterprise "The Committee on traditional and alternative Medicine of the Ministry of Health of Ukraine" [Electronic source]: website of MHU. – On-line mode: http://moz.gov.ua/ua/portal/ oth_narmed.html, free access (date of access: 21.11.2016).

3. Бодянський Є. В., Дейнеко А. О., Жернова П. Є., Золотухін О. В., Хаустова Я. В. Послідовне ядерне нечітке кластерування великих масивів даних на основі гібридної системи обчислювального інтелекту

SEQUENTIAL KERNEL FUZZY CLUSTERING OF BIG DATA BASED ON COMPUTATIONAL INTELLIGENCE HYBRID SYSTEM

Yevgeniy Bodyanskiy[1], Anastasiia Deineko[2], Polina Zhernova[3], Oleh Zolotukhin[4], Yana Khaustova[5]
[1, 2, 3, 4, 5] Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, 14, Nauky av. Kharkiv, 61166, UKRAINE, E-mail: [1]yevgeniy.bodyanskiy@nure.ua, [2]anastasiya.deineko@gmail.com, [3]polina.zhernova@gmail.com, [4] oleg.zolotuxin@gmail.com, [5]yana.kutsenko@nure.ua

Nowadays systems and methods of computational intelligence are widely used in solving various problems of Data Mining, such as image processing, classification, clustering, etc. Clustering of multidimensional observations in large databases often occurs in many real practical tasks and for its solution many algorithms were proposed. Today in the frame of Big Data concept special attention is given to the processing of information, that is stored either in the

very large data bases, or is located sequentially in on-line mode from the cloud in the form of data stream. For solving these tasks mathematical apparatus of computational intelligence can be successfully used, especially artificial neural networks and soft computing based on fuzzy logic. It is clear that known systems of computational intelligence must be essentially modified for processing large data volumes that are sequentionally integrating into the processing. Obviously, most of the available methods for solving clustering tasks do not construct a single hypothetical algorithm that would potentially be effective for all of the occurring situations. One of such hypothetical and complex situations is connected with assumption, that every vector-observation can be associated with different levels of possibilities, probabilities or their involvement into several or all formed clusters at the same time. The most popular systems designed for data processing in batch mode are BSB– and ART– neural networks; T. Kohonen's self-organizing maps (SOM) were designed for sequential data clustering due to their implementation simplicity and computational possibilities of sequential data processing in on-line mode that allow their appliance to tasks of Dynamic Data Mining and Data Stream Mining. It is as well supposed that the recovered classes aren't mutually overlaping and have a convex shape, i.e. in the self-adjustment recovering process separating hyper-planes that clearly separate different clusters.

In situations when clusters have arbitrary shape the so-called kernel self-organizing maps (KSOM) may be used to solve the problems of clustering. These systems are built using the J. Mercer's kernels and are based on minimization of empirical risk criterion that takes place in so-called support vector machines (SVM) introduced by V. Vapnik. SVM-neural networks are really effective means for solving many problems of Data Mining including clustering. However, because the number of neurons in this network is determined by the volume of the processed dataset, it's clearly that it is not appropriate for the problems associated with on-line analysis of the information that is integrated into the system. In this regard, instead of the conventional approach to SVM-kernel systems in clustering tasks it is possible to use the ideas associated with the E. Parzen's windows, D. Specht's generalized regression neural networks (GRNN) and T. Cover's theorem of linear classes separability in spaces of higher dimensions.

The architecture and self-learning method for hybrid neuro-fuzzy system for big data fuzzy clustering in on-line mode are proposed in this paper. The architecture of proposed system consists of four information processing layers and represents the hybrid of the fuzzy generalized regression neural network and T. Kohonen's clustering self-organizing network. During a learning procedure in on-line mode the proposed system tunes both its parameters, and its architecture. For shaping the belonging functions of neuro-fuzzy system the method based on competitive learning is introduced. The hybrid neuro-fuzzy system shapes its synaptic weights, centers and width parameters of membership functions. Experimental results have proved the fact that the proposed system could be used to solve a sequential data stream clustering task. The proposed system is characterized by computational simplicity. A distinguishing feature of the proposed system is that this system combines both supervised learning, and self-learning procedures.

Key words: self-learning method for hybrid neuro-fuzzy system, fuzzy clustering, clustering self-organizing network, fuzzy generalized regression neural network

1. Kohonen T. Self-Organizing Maps / T. Kohonen // Berlin: Springer-Verlag. – 1995. – 362 p. 2. Bezdek, J.-C. Pattern Recognition with Fuzzy Objective Function Algorithms [Text] / J. C. Bezdek. – N.Y.: Plenum Press, 1981. – 272 p. 3. Tsao E.C.-K. Fuzzy Kohonen clustering networks [Text] / E.C.-K. Tsao, J. C. Bezdek, J. C. Tsao, N. R. Pal // Pattern Recognition. – 1994. – No. 27. – P. 757–764. 4. Pascual – Marqui R. D. Smoothly distributed fuzzy C-means: a new self-organizing map / R. D. Pascual – Marqui, A.D. Pascual – Montano, K. Kochi, J.M. Caroso // Pattern Recognition. – 2001. – No. 34. – P. 2395–2402. 5. MacDonald D., Fyfe C. Clustering in data space and feature space [Text] : ESANN'2002 Proc. European Symp. on Artificial Neural Networks. Bruges (24-26 April 2002). – Belgium. – 2002. – P. 137–142. 6. Girolami, M. Mercer kernel-based clustering in feature space [Text] / M. Girolami // IEEE Trans. on Neural Networks. – 2002. – Vol. 13. – No. 3. – P. 780–784. 7. Camastra F. A novel kernel method for clustering [Text] / F. Camastra, A. Verri // IEEE Trans. on Pattern Analysis and Machine Intelligence. – 2005. – No. 5. – P. 801–805. 8. Schölkopf, B. Learning with Kernels [Text] / B. Schölkopf, A. Smola // Cambridge M. A.: MIT Press. – 2002. – 648 p. 9. Kacprzyk J. Springer Handbook of Computational Intelligence [Text] / J. Kacprzyk, W. Pedrycz. – Berlin Heidelberg: Springer – Verlag, 2015. – 1634 p. 10. Haykin, S. Neural Networks and Learning Machines [Text] / S. Haykin. – N.Y. :Prentice Hall, 2009. – 1634 p. 11. Cortes C. Support Vector Networks [Texy] / C. Cortes, V. Vapnik // Machine Learning. – 1995. – No. 20. – P. 273–297. 12. Parzen E. On the estimation of a probability density function and the mode / E. Parzen // Ann. Math. Statist. – 1962. – No. 38. – P. 1065–1076. 13. Specht, D.F. A general regression neural network [Text] / D.F. Specht // IEEE Trans. on Neural Networks. – 1991. – Vol. 2. – P. 568–576. 14. Zahirniak D. Pattern recognition using radial basis function network. [Text] / D. Zahirniak, R. Chapman, S. Rogers, B. Suter, M. Kabrisky, V. Piati // Proc 6th Ann. Aerospace Application of Artificial Intelligence Conf. – Dayton, OH. – 1990. – P. 249–260. 15. Cover T. M. Geometrical and statistical properties of systems of linear inequali-ties with applications in pattern recognition [Text] / T.M. Cover // IEEE Trans. on Electronic Computers. – 1965. – No. 14. – P. 326–334. 16. Angelov, P. Evolving Rule-based Models: A Tool for Design of Flexible Adaptive Systems [Text] / P. Angelov // Heidelberg-New York: Springer-Verlag. – 2002. –

211 p. 17. Kasabov N. Evolving Connectionist Systems [Text] / N. Kasabov – London: Springer-Verlag. – 2003. – 307 p. 18. Angelov P. Evolving computational intelligence systems [Text] / P. Angelov, N. Kasabov // Proc. 1st Int. Workshop on Genetic Fuzzy Systems. – Granada, Spain. – 2005. – P. 76–82. 19. Lughofer E. Evolving Fuzzy Systems – Methodologies and Applications [Text] / E. Lughofer. – Studies in Fuzziness and Soft Computing. – Springer-Berlin. – 2011. – 410 p.

4. Буров Є. В., Микіч Х. І. Формальна модель опрацювання знань у системах із ситуаційною обізнаністю

FORMAL MODEL OF KNOWLEDGE PROCESSING IN SITUATIONAL AWARENESS SYSTEMS

Yevhen Burov[1], Khrystyna Mykich[2]

[1, 2] Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail: [1] Yevhen.V.Burov@lpnu.ua, [2] *Khrystyna.I.Mykich@lpnu.ua*

Models and methods of formalizing and modeling process of decision making in systems with situation awareness (SAW systems) were considered and analyzed. The important feature of autonomous intellectual systems is situation awareness – the ability to obtain information about the state of environment and using this data in combination with knowledge about domain to make decisions on necessary actions.

Acquisition and representation of data in the form of algebraic structures are widely used by the experts in the field of information technology. Nowadays different methods are used for domain modeling and the resolution of problems in SAW systems. Existing methods of formalizing situations allows to display only certain subtasks SA. But decision-making and identification of situations in the real world require a combination of various methods.

Thus, there is a need to develop a holistic approach that allows using within a system of different techniques of situation awareness modeling. This can be achieved by developing a central formal model and building mappings to other methods and models from it.

For research situation awareness use models in which detailed parts of SAW and their interdependence.

Formal models were considered, because building a formal model is a prerequisite for better understanding of the SAW system requirements and is the basis for further development frameworks and SAW systems architectures.

The aim of the paper is to develop formal models and frameworks for knowledge representation and processing in SAW systems (particularly algebraic framework) and as well as building the mapping of algebraic model into model of description logic.

The process of knowledge processing in SAW systems were viewed and analyzed. The situation as a mathematical unit (infon) was presented.

The unifying algebraic model, allowing the usage of different tools for situation awareness modeling, based of algebra of systems was proposed. The representation of situations and mapping between algebraic model and description logic model was developed. Also, the modeling process of complex systems using this framework was analyzed.

The process of evaluation and identification of situations is based on ontological modeling of subject area using the framework of Algebra of Systems.

The proposed algebraic approach is based on Algebra of System and meets the general JDL model for SAW systems. It has sufficient flexibility and allows using for modeling and solving problems other mathematical methods by building mappings between them.

The mapping of algebraic model to description logic model is described allowing to perform necessary logical reasoning in SAW system.

Description logics (DL) are used for reasoning in ontology based systems. They are based on first order predicate logic.

The usage of logical reasoning is an important part of any knowledge-based system because it helps to maintain the logical consistency of domain model and the correctness of data. However, the usage of DL as a primary model for domain representation has also substantial drawbacks. The simplified model of DL based on first order logic does not allow to make other types of reasoning (inductive, statistical, based on analogies) which are helpful in human-like cognitive activities.

The mappings between algebraic model, description logic and interpreted systems are proposed. It allows using logical reasoning that are developed for description logic and interpreted systems for research ontology based systems.

Key words: situation awareness, formal model, framework, Algebra of Systems, description logic, interpreted systems.

1. Endsley M. Theoretical underpinnings of situation awareness: a critical review Process More Data ≠ More Information / M. Endsley, R. Mica // Edited by Array. Most, Vol. 301, 2000. – P. 3–32. 2. BeAware!-Situation

awareness, the ontology-driven way / N. Baumgartner [and other] // Data and Knowledge Engineering, Vol. 69, 2010. – P. 1181–1193. 3. Jousselme A. Interpreted Systems for Situation Analysis / Jousselme Anne Laure, Patrick Maupin // 10th International Conference on Information Fusion. – IEEE, 2007. – P. 1–11. 4. *Steinberg A. N. Revisions to the JDL Model / Steinber A. N., Bowman C. L., White F. E // Joint NATO/IRIS Conference Proceedings. – Quebec, 1998 and in Sensor Fusion: Architectures, Algorithms, and Applications, Proceedings of the SPIE, Vol. 3719, 1999. – P. 430–441. 5. Codd E. Relational Completeness of Data Base Sublanguages / E. F. Codd. – IBM corporation, Vol. 54, 1972. – P. 65–98. 6. Koo B. Algebra of Systems: A metalanguage for Model Synthesis and Evaluation / B. Koo, W.L. Simmons, E.F. Crawley // IEEE Transactions On 39.3, Systems, Man and Cybernetics, Part A: Systems and Humans, 2009. – P. 501–513. 7. Koo B. Meta-language for Systems Architecting / B. A. Koo // In Engineering Systems and Management Massachusetts Institute of Technology, 2005. – P. 168. 8. Koo B. A valuation technology for product development option using an executable meta-modeling language in Complex Systems Concurrent Engineering / B. Koo, W. L. Simmons, E. F. Crawley // Collaboration, Technology Innovation and Sustainability. New York: Springer-Verlag, 2007. – P. 107–115. 9. Burris S. A Course in Universal Algebra / S. Burris, H. P. Sankappanavar. – New York: Springer-Verlag, 1981. – P. 129–154. 10. Barwise J. Situation Theory and Its Applications / Jon Barwise // Center for the Study of Language (CSLI), Vol. 26, 1991. – P. 631. 11. Devlin K. Situation theory and situation semantics / K. Delvin. – Handbook of the History of Logic, vol. 7, 2006. – P. 601–664. 12. White F. E. A Model for Data Fusion / F. E. White // Proc. 1st National Symposium on Sensor Fusion. – 1988. – P. 153–158.*

5. Романишин Ю. М., Єлманова О. С. Узагальнений опис контрасту елементів зображення

GENERALIZED DESCRIPTION OF CONTRAST OF IMAGE ELEMENTS

Yuriy Romanyshyn[1, 2], Elena Yelmanova[3]

[1, 3] Department of Electronics and Computer Technologies, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, [2] University of Warmia and Mazury in Olsztyn, Olsztyn, POLAND, E-mail: [1]yuriy.romanyshyn1@gmail.com, [3]yelmanova@lp.edu.ua

The problem of image quality assessment in the transmission and processing of video data in information and telecommunication systems is considered. The image quality is characterized by several basic parameters [1–4]. Generalized contrast of image is the most important quantitative characteristic which determines the objective quality of the image [3, 5]. The generalized contrast of complex images is generally defined basing on the quantitative assessments of contrast for individual pairs of image elements (objects and background) [5–7].

One of the most important problems when measuring the generalized contrast of complex images is the choice of definition of contrast for two elements of image [7]. Definitions of the contrast of image elements (of the contrast kernels) will meet the basic requirements to the contrast definition and should ensure reasonably accurate quantitative assessment (measurement) of contrast of image elements for the real complex images and should allow evaluate (predict) the perceived values of image contrast at carrying out of subjective (qualitative) expert assessments [8].

The problem of contrast measuring of image elements (objects and background) on complex (multi-element) images is considered in this paper. The paper deals with the basic requirements to the contrast definition [7, 8, 10]. Known definitions of a weighted and relative contrast of image elements are considered [5–7, 9, 11–13].

The main disadvantages of known contrast definitions are the uncertainty and the multiplicity of the conditions under which the extreme values of contrast are achieved. To address these shortcomings we propose a new method of contrast measurement of image elements on the basis of assessments of contrast for appropriate elements on the primary (initial) and inverted (negative) images.

The generalized description of image elements contrast for different definition of contrast kernels was suggested. The new definitions of weighted and relative contrast of image elements were proposed.

For contrast definition of elements of complex images using the chosen (the specified) definition of contrast kernel we propose the generalized contrast description on the basis of analytical definitions of contrast for the initial (primary) and for inverted (negative) images.

The research of known and proposed definitions of a weighted and relative contrast to evaluate the efficiency of measuring (of quantitative assessment) of contrast of image elements was carried out. Experimental researches were carried out through the comparative analysis of known and proposed definitions of a weighted and relative contrast for compliance with the basic requirements to the contrast.

The proposed new descriptions of weighted and relative contrast of image elements on the pre-normalized images satisfy to the all known basic requirements to the contrast definition and ensure reasonably accurate quantitative assessment (measurement) of contrast of image elements for the real complex images [14] and also allow evaluating (predict) the perceived values of image contrast at carrying out of subjective (qualitative) expert assessments.

The proposed definitions of contrast can be recommended for image quality assessment in transmission and processing of video data in telecommunication systems of various destinations.

Key words: image quality assessment, image contrast, contrast of image elements, weighted contrast, relative contrast, contrast kernel.

1. Wang, Z. and Bovik, A. C. (2006), Modern Image Quality Assessment, Morgan & Claypool Publishers, San Rafael, California, USA. 2. Gonzalez, R.C. and Woods, R. E. (2002), Digital Image Processing, 2nd ed., Prentice Hall, New Jersey, USA. 3. Wang, Z. and Bovik, A. (2002). A universal image quality index // IEEE Signal Processing Letters, vol. 9, no. 3, pp. 81–84. 4. Pratt, W.K. (2001), Digital Image Processing: PIKS Inside, 3rd ed., John Wiley & Sons, Inc., New York, USA. 5. Peli, E. (1990). Contrast in Complex Images // Journal of the Optical Society of America, A, vol. 7, no. 10, pp. 2032–2040. 6. Nesteruk, V.F. and Porfyryeva, N.N. (1970). Contrasting law of light perception // Optics and Spectroscopy, vol. XXIX, no. 6, pp. 1138–1143 (In Russian). 7. Vorobel, R.A. (1999), Digital image processing based on the theory of contrast, D. Sc. Thesis, Lviv, Ukraine. 8. Vorobel, R. A. (2012). Logarithmic Image Processing, Naukova Dumka, Kyiv, Ukraine (In Ukrainian). 9. Vorobel, R. A. (2011). The generalized description of weighted and relative contrasts of elements of monochrome images // Information Extraction and Processing, issue 34 (110), pp. 120–128. 10. Yelmanova, E. (2016). Quantitative Evaluation of Contrast for a Complex Image by its Histogram // Proc. of 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science TCSET'2016, Lviv-Slavske, Ukraine, pp. 688-692. 11. Pratt, W.K. (1978), Psychophysical properties of vision // Digital Image Processing, John Wiley & Sons, Inc., New York, USA, pp. 23–49. 12. Stockham, T.G. (1972). Image processing in the context of a visual model // Proc. of the IEEE, vol. 60, no. 7, pp. 828–442. 13. Jourlin, M., Pinoli, J-C. and Zeboudj R. (1989). Contrast definition and contour detection for logarithmic images // Journal of Microscopy, vol. 156, issue 1, pp. 33–40. 14. Public-Domain Test Images for Homework's and Projects, available at: https://homepages.cae.wisc.edu/~ece533/images/ lena.bmp (Accessed 1 December 2016).

6. Василюк А. С., Басюк Т. М. Інтелектуальний аналіз математичного забезпечення процесів згортання та розгортання формул алгоритмів

INTELLIGENT ANALYSIS OF MATHEMATICAL SOFTWARE COLLAPSING AND EXPANDING OF FORMULAE ALGORITHMS

Andrii Vasyliuk, Taras Basyuk
Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, Ukraine

Algebra algorithms – with algorithms as its objects – are described as formulae that can be converted with the intention to minimize, substitute, reduce, and expand formulae algorithms. This theory has the specific features of operations, such as sequencing, elimination, and cyclic paralleling operations that appear as special characters being unknown mathematical symbols. In order to simplify the process of algebra algorithms formulae typing and editing, it is needed to implement formulae reduction and expancion.

This article discusses the processes of reduction and expancion of formulae of algebra algorithms. The process of reducing of algorithm formula presupposes replacement of the part of the formula for trivial algorithm uniterm with the sequential number of a reduced formula. In its turn, the process of formula expansion presupposes the replacement of trivial unitherms of the sequential number of a reduced formula with the responding expanded formula with the preservation of the corresponding hierarchical relationships. These processes are further accompanied by adapting and adjusting the formulae of algebra algorithms.

Mathematical software should describe the possibility of reducing and expanding of formulae of abstract algorithms that visually reduce abstract algorithm formula.

Reducing of formulae of abstract algorithm is implemented as follows:

– select a formula abstract algorithm;

– select the tool on the toolbar;

"reduce or expand an abstract algorithm formula".

The system will automatically minimize abstract algorithm formula.

Expanding of formula of abstract algorithm is implemented in the following way:

– select the required formula abstract algorithm (in the form of the uniterm);

– select the tool on the toolbar "reduce or expand an abstract algorithm formula."

The synthesising of algorithm reduction and expansion can be considered accomplished. They are described by means of algebra algorithms. The minimization of abstract algorithms and formulae deployment of algorithms by the sequential unitherms can be considered accomplished as well.

The synthesising process of software deployment and minimization of formulae algorithms is described in the relevant seqence field variables, allowing us to perform research of software deployment processes coagulation and formulae algorithms.

The advantage of the synthesized and minimized Mathematical providing of reduction and expansion of formulae of algebra algorithms is that it ensures the correct implementation of adaptation with consideration of geometrical parameters and orientation of the unitherms.

Key words: uniterm, algorithm, minimization, reduction, expansion, mathematical model.

1. Ovsyak V., Brytkovs'kyy V., Ovsyak O., Ovsyak Yu.. Syntez i doslidzhennya algorytmiv komp"yuternykh system. – L'viv, 2004. – 276 s. 2. V. Ovsyak. ALGORYTMY: metody pobudovy, optymizatsiyi, doslidzhennya virohidnosti. – L'viv: Svit, 2001. – 160 s. 3. Vasyliuk A. S. Intelektual'nyy analiz struktury danykh matematychnoho zabezpechennya redaktora formul alhorytmiv / A. S. Vasyliuk, // Visnyk Natsional'noho universytetu "L'vivs'ka politekhnika". – 2015. – # 832 : Informatsiyni systemy ta merezhi. – S. 34–48. 4. Katrenko A. V.Analiz matematychnykh modeley planuvannya v mul'typroektnomu seredovyshchi / A. V. Katrenko, A. S. Mahats // Visnyk Natsional'noho universytetu "L'vivs'ka politekhnika". – 2014. – # 783 : Informatsiyni systemy ta merezhi. – S. 443–450, 5. Ovsyak V., Kozelko M. Zahal'na model' redaktora hrafichnykh unitermiv / V. Ovsyak, M. Kozelko // Visnyk TNTU. – 2013. – Tom 69. – #1. – S. 183–192. 6. Vasyliuk A. Abstraktnyy alhorytm redaktora formul abstraktnykh alhorytmiv "AbstraktAl" / A. Vasyliuk // Komp'yuterni tekhnolohiyi drukarstva : Zbirnyk naukovykh prats'. – L'viv: UAD, 2006. – #16. – S. 99 –108.

7. Грабовська Н. Р., Русин Б. П., Іванюк В. Г. Метод тривимірної реконструкції поверхні за тріадою зображень та оцінка його точності

METHOD OF THREE-DIMENSIONAL RECONSTRUCTION OF SURFACE AFTER TRIAD OF IMAGES AND HIS EXACTNESS ESTIMATION

Natalya Hrabovska[1], Bohdan Rusyn[2], Vitaliy Ivaniuk[3]

[1, 2, 3] Department of "Methods and systems of the remote sensing ", Karpenko Physico– Mechanical Institute of NAS of Ukraine, 5, street Scientific, Lviv, 79601, UKRAINE, E-mail: [1]dep32@ipm.lviv.ua, [2]dep32@ipm.lviv.ua, [3]vivan@imp.lviv.ua

The article considers the method of three-dimensional surface reconstruction from the triad of two-dimensional images. This method is based on Lambertian reflection model. A method of 3D reconstruction of the prospected surface is described step by step. It is proposed to use three directions of illumination for the recording process of the triad of images. For the recording process of every image it is suggested to use the personal direction of illumination.

During the recording process of first image it is applied vertical direction of illumination.

During the recording process of second image it is applied lateral direction of illumination. This lateral direction of illumination belongs to the ortogonal plane of XZ.

During the recording process of third image it is applied other lateral direction of illumination. This lateral direction of illumination belongs to the ortogonal plane of YZ.

The technologically pre-arranged value of every direction of illumination is a parameter of proposed reconstruction. Other parameters of proposed reconstruction are values of pixels of the triad of images.

After video recording three directions of illumination together with the triad of images being reconstructed to derivatives on the surface are in every point.

Errors that can appear when the algorithm is used are shown.

These errors are conditioned by deviation of lateral direction of illumination from the technologically pre-arranged value in plane XZ and YZ. A method of the estimation of errors of the reconstructed derivatives is described step by step.

First step. For the analytical show of errors of every direction of illumination it is suggested to use function of the two angles. First angle of direction of illumination belongs to the ortogonal plane of XZ. Second angle of direction of illumination belongs to the ortogonal plane of YZ. Second step. Functions of the two angles that can appear when the reconstructed derivatives are used in the real work are shown. Third step. The row of Teilor when the reconstructed derivative used the every angle, is shown. Fourth step. The error of reconstructed derivative when the reconstructed derivative used the every angle is extracted from the approximated row of Teilor.

These errors that can appear when the algorithm is used are estimated. The reconstruction and the estimation of errors of the reconstructed derivatives is described in the three stages.

First stage. Before video recording it is measurement of direction of illumination. That measurement is done with device of lateral ray. The device has a source of light and two screens. These screens are located between the

source of light and the prospected surface. Every screen has small opening. Opening is located so, that providse the light ray through them if direction of illumination has the technologically pre-arranged value. For that the source of light has the adjusting device of direction of illumination. So, pre-arranged mode of illumination is set by the structural adjusting of direction of illumination. It is known the diameter of opening and known distance between screens. In this case it is shown a value of the pre-arranged absolute error of direction of illumination. During realization of quantitative analysis of exactness the used absolute error of establishment of lateral direction of illumination does not exceed 0,01 advices.

Second stage. The opening is large. Lighting on the surface of the prospected object exists. After that the video recording is started. Then it happens reconstruction of derivatives.. During realization of quantitative analysis of exactness it is used technical descriptions of system of three-dimensional reconstruction. Its relative error of reconstruction of horizontal and vertical derivative does not exceed ten percents.

Third stage. When known absolute error of lateral direction of illumination and when known relative error of reconstruction of horizontal and vertical derivative, these parameters apply in the proposed formula of the error of reconstructed derivative.

In this case during realization of quantitative analysis of this formula, we got recommendations for establishment of ranges of the reconstructed derivatives.

Key words: 2D image, 3D reconstructio, direction of illumination.

1. Sharland S. M. A review of the theoretical modeling of crevice and pitting corrosion Corrosion Science 1987. Vol. 27. N 3. – P. 289–323. 2. Marcus P. Corrosion *mechanisms in theory and practice 2nd Edition – 2002 – New York, Technology & Engineering – 742 p. 3. Броек. Д. Основы механики разрушения. / Д. Броек. – М.: Высшая школа, 1980. – 368 с. 4. Liu Y., Su X. and Zhang Q. A novel encoded-phase technique for phase measuring profilometry // Opt. Express 19(15), 2011. – P. 14137–14144 . 5. Zhang Q., Su X., Xiang L. and Sun X. 3-D shape measurement based on complementary gray-code light, Opt. Lasers Eng., vol. 50, 2012. – P. 574–579. 6. Zhang S., Van Der Weide D. and Oliver J. Superfast phase-shifting method for 3-D shape measurement, Opt. Express vol. 18, No. 9, 2010. P. 9684–9689. 7. Woodham R. J. Photometric method for determining surface orientation from multiple images Optical Engineering. 1980. – 19(1). – P. 139–144. 8. Maria E. Angelopoulou, Maria Petrou Evaluating the effect of diffuse light on photometric stereo reconstruction Machine Vision and Applications. 2014, Vol. 25, Issue 1. – P. 199–210. 9. Yoshizawa, T. Handbook of Optical Metrology -principle and applications– CRC Press, New York 2009. – 744 p. 10. Zhang S. High-resolution 3-D profilometry with binary phase-shifting methods, Appl. Opt.50(12), 2011. – P. 1753–1757. 11. Sudipta N. Sinha, D. Scharstein, R. Szeliski. Efficient high-resolution stereo matching using local plane sweeps // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2014), 2014. – P. 1219–1222. 12. Richard Szeliski. Computer Vision: Algorithms and Applications. Springer, New York, 2010. – 655 P. 13. Неруйнівний контроль стану поверхонь, уражених корозійними пітингами / Б. П. Русин, Н. П. Ануфрієва, Н. Р. Грабовська, В. Г. Іванюк // Фіз.-хім. механіка матеріалів. – 2013. – № 4. –С. 90–96. 14. Оцінка глибини тріщини за тріадою зображень. / Б. П. Русин, В. Г. Іванюк, О. В. Капшій, Н. П. Ануфрієва // Радіоелектроніка і інформатика. – 2010. – № 2. – С. 70–78. 15. J. H. Lambert. Photometria, sive de Mensura et gradibus luminis, colorum et umbrae / sumptibus viduae E. Klett, 1760. 16. Shintaro Watanabe, Koji Miyajima Detecting Building Changes Using Epipolar Constraint From Aerial Images Taken At Different Positions // ICIP 2001. – 2001. – P. 201–204. 17. Смирнов В. И. Курс высшей математики. Т. 1. – М.: Наука, 1967. – 480 с. 18. tool@microtech-ua.com. 19. Wood W. A. Recent observations on fatigue fracture in metals, / W. A.Wood // ASTM STP 237, 1958. – P. 110–121. 20. Tetelman A. S. Fracture of structural materials / A. S. Tetelman, A. J. McEvily // John Wiley. – 1967. – 697 p.*

8. Левус Є. В., Нечипір Н. І., Полиняк Ю. В. Аналіз алгоритму Apriori для структурованих та неструктурованих даних

THE ANALYSIS OF APRIORI ALGORITHM FOR STRUCTURED AND UNSTRUCTURED DATA

Yevheniya Levus, Nadiya Nechypir, Yurii Polyniak

Software Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE,
E-mail: [1]Yevheniia.V.Levus@lpnu.ua, [2]nadia.nechypir@gmail.com, [3]yurii.polyniak@gmail.com

Modern information society is characterized by not only sheer growth of data amounts, but also by its unlimited variety. Practically 75% of all the information is both unordered and unstructured. Structured data is saved in data bases, which are easily processed, hence, it's easy to get all kinds of regularities and dependencies from them. The term "Big Data" is closely related to unstructured data. According to experts, unstructured data forms up to 90% of big data [1–7].

There's a range of hardware-software complexes, which offer an already configured solution to process large volumes of data: Aster MapReduce appliance, Oracle Big Data appliance, Greenplum appliance. However, all those highly productive systems have major drawbacks: 1) high prices; 2) expensive support of this class of solutions. Those factors play a huge role when it comes to small and medium companies [4, 9, 10].

Association rule learning algorithms have become one of the most popular existing methods for identifying data patterns. Considering all pros and cons of this kind of algorithms, as well as their implementation particulars, we have chosen to analyze the Apriori algorithm on big amounts of unstructured and structured data [11–15].

The article contains the comparison of results and performance of the Apriori algorithm for big amounts of unstructured and structured data. Performance quality of the algorithm is defined by the number of generated rules for different values of support and confidence coefficients. Productivity is defined by the processing time with the required computing capacities of the system.

For the computational experiments, it was required to implement a prototype of a recommendatory program system for a sales outlet, based on the Apriori algorithm implementation. The Apache Hadoop platform has been used for distributed processing of big data [19–24].

The software system provides us with the following functions: data pre-processing; data analysis; parallel data processing; association rules search or creation. The project is made up of two modules: unstructured data processing and structured data processing. Association rule learning algorithm has been implemented using Java programming language in the NetBeans IDE.

For this goal we used a database of items, which contains information on all transactions with a certain set of products. The unstructured data downloaded from the * .csv files is stored using HDFS file system and processed with a computing framework for distributed tasks named MapReduce. The structured data is located in the "Products" database in MS SQL Server 2012. The algorithm processes the same sets of data.

The system for the unstructured data processing requires the minimum of 16 GB RAM and an **8-core** processor. On the other hand, in order to process structured data we only need 2 GB RAM and a dual-core processor. To estimate the program system's work results we executed the following processes: 450MB, 1GB, 2.5Gb of the structured and unstructured data respectively. The data contains information on a retail outlets' products and corresponding transactions.

When we set support and confidence coefficients to more than 0.5 (for example, 0.6 and 0.7 respectively), we found only 3 rules for the structured data and none for the unstructured data. This points out the limitations of the Apriori algorithm when dealing with unstructured data. For the average support and confidence coefficients the algorithm finds approximately 20% more association rules for the structured data than for the unstructured data.

Apriori algorithm works 3 times faster with small sets of structured data compared to the same kind of sets of unstructured data. As the data volume grows, speed advantages of the unstructured data processing drastically decreases. For the experiment using 1Gb of data, the processing time is 70% bigger when dealing with unstructured data. But the experiment with 2.5Gb of data only shows 12% difference.

Aforementioned results can be used to reason an optimal choice of implementation technologies.

Key words: apriori algorithm, structured data, unstructured data, associative rule, data volume.

1. Montgomery Karen. Big Data Now: 2014 Edition. O'Reilly Media. – Junuary, 2015. – 165 p. 2. Майер-Шенбергер Виктор, Кукьер Кеннет. Большие данные. Революция, которая изменит то, как мы живём, работаем и мыслим = Big Data. A Revolution That Will Transform How We Live, Work, and Think / пер. с англ. И. Гайдюк. – М.: Манн, Иванов, Фербер, 2014. – 240 с. 3. Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. John Wiley & Sons. 2014-12-19. 300p. 4. Big Data Applience [Електронний ресурс] // Oracle Big Data: сайт. – Режим доступу https://www.oracle.com/engineered-systems/big-data-appliance/index.html. 5. Almasi, G.S. and A. Gottlieb (2009). Highly Parallel Computing. Benjamin // Cummings publishers, Redwood City, CA. – 235 с. 6. Шаховська Н. Б. Організація великих даних у розподіленому середовищі / Н. Б. Шаховська, Ю. Я. Болюбаш, О. М. Верес// Наукові праці ДонНТУ. Серія: обчислювальна техніка та автоматизація. – № 2(27). – 2014. – С. 147–155. 7. Павич Н. Я. Оцінювання ефективності опрацювання даних великих обсягів технологіями Spark та Hive / Н. Я. Павич, О. П. Крохмальна // Вісник Нац. ун-ту "Львів. політехніка" "Комп'ютерні системи та мережі". – 2015. – № 830. – С. 128–135. 8. Сєдушев О. Ю. Методи видобування даних з баз нечітких знань / О. Ю. Сєдушев, Є. В. Буров // Вісник Нац. ун-ту "Львів. політехніка" "Інформаційні системи та мережі". – 2014. – № 783. – С. 193–203. 9. Mapreduce Appliance [Електронний ресурс] // MapReduce: сайт. – Режим доступу http://www.teradata.com/products/Aster_MapReduce_Appliance. 10. GreenPlum. [Електронний ресурс] // сайт. – Режим доступу http://www.emc.com/campaign/global/ greenplumdca/index.htm. 11. Zhu Yixia, Yao Liwen, Huang Shuiyuan, Huang Longjun. A association rules mining algorithm based on matrix and trees[J]. Computer science.

2006, 33(7):196-198. 12. Tong Qiang, Zhou Yuanchun, Wu Kaichao, Yan Baoping. A quantitative association rules mining algorithm[J]. Computer engineering. 2007. 13. Agrawal R., Imielinski T., Swami A. Mining association rules between sets of items in large database, In Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data, 1993: 207–216. 14. Agrawal R. and Srikant, R. Fast algorithms for mining association rules. In Proc.20th Int. Conf. Very Large Data Bases, Santiago, Chile, 1994. 487–499. 15. Purdom P. W., Guch D. V., Groth D. P. Avarage case performance of the apriori algorithm – SIAM Journal on Computing, 33(5): 1223–1260, 2004. 16. Mohammed J. Zaki. Scalable algorithms for association mining – IEEE Transactions on Knowledge and Data Engineering, 12(3):373–390, 2000. 17. Brin S., Rajeev Motwani, Ullman J., Tsur S. Dynamic itemset couting and implication rules for market basket data// Proc. ACM SIGMOD Intern. Conference on Management of Data, 255–264 p., USA, 1997. 18. Apache Hadoop. [Електронний ресурс]// Big Data: сайт. – Режим доступу https://hadoop.apache. org/docs/r1.2.1/hdfs_design.html 19. Harris, Dereck Intel jettisons its Hadoop distro and puts millions behind Cloudera (27 March 2014). 20. Уайт, Том Hadoop. Подробное руководство= Hadoop: The Definitive Guide. – СПб., 2013. – 672 с. 21. Hadoop File System. [Електронний ресурс]// hadoop-distributed-file-system: сайт. – Режим доступу https://www. safaribooksonline. com/blog/2013/02/13/the-hadoop-distributed-file-system. 22. White T. Hadoop: The Definitive Guide, 4th Edition. O'Reilly Media. – March, 2015 – 756 p. 23. Нечипір Н. І. Опрацювання великих обсягів неструктурованих та структурованих даних алгоритмом Apriori / Н. І. Нечипір, Є. В. Левус // Математичне та програмне забезпечення інтелектуальних систем: матер. XIII Міжнар. наук.-практ. конф. – Вид-во Дніпропетр. Нац. у-ту ім. Олеся Гончара, 2015. – С. 34–36.

9. Корж Р. О., Пелещишин А. М., Мастикаш О. В. Інформаційна система взаємодії ВНЗ з соціальними середовищами інтернету

INFORMATION SYSTEM INTERACTION OF UNIVERSITIES WITH SOCIAL ENVIRONMENT INTERNET

Roman Korzh[1], Peleshchyshyn Andriy[2], Mastykash Oleg[3]

[1] Vice-Rector for Educational and Pedagogical Work, [2] Professor and Head of Department of Social Communications and Information Activities, [3]PhD of Department of Social Communications and Information Activities, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail:[1]korzh@lp.edu.ua, [2]apele@ridne.net, [3]mastykash@itstep.org

Information activities of higher educational institutions in social media provide a wide range of accounting and analytical works as well as a complex of work on accounting, analysis and management of the departments of higher education institution in the communities to form the image of the university. The formalization of the process of creating an information image allows to significantly reducing the costs for information activities and management. In addition, accounting the results of universities information activities gives an opportunity to develop and implement an integrated system of remuneration and promotion for departments achieved high results.

A formal accounting structure of the higher educational institution is implemented in the present article. The core software system for managing universities information activities consists of a multi-component database that stores reference data and current university activities in social media is documented. Information from the database is used and updated within all workplaces of the complex. The article presents the implementation of the information model of each database component. Chart models presented in the article are somewhat simplified in comparison with a real database, so the physical aspects of the structure are not shown, attributes that are of a technical ancillary nature are not presented, the system of data integrity and codification of data are not described.

The main purpose of writing the article is to develop the structure components of database, which stores information necessary for the execution of the automated search items of information image of higher educational institutions in communities, as well as to develop the structure of the following databases: the organizational structure of the university, indicators of communities, specific indicators of communities, generators, communication characteristics of generators, areas of responsibilities, etc.

The architecture of the information model of each database components in the form of "entity-relationship" diagrams is developed. In particular, the structure of the database "Organizational structure of the higher educational institution" is developed, a multi-level model of the generator for information activity of universities is developed. The account of generators is constructed in such a way that it actually realizes and accounts all available image information of higher educational institution in social media. The structure of the database for account of administratively approved areas of responsibility of departments for information activity of higher education institutions is developed.

An important component of the overall database of the complex and part of its core is the database component, which is to take into account the social media (communities) that are involved or can be involved in the university

information activity. The article defines data arrays that should be stored in the community database: the registration of communities in which active information activities are carried out, the accounting of communities for which content is monitored and activity analysis, the registration of communities that are previously detected and for which identification is carried out basic characteristics, determining the relevance and appropriateness of conducting university information activity and the similar. Despite the different purpose of each array, their structure is of the same type, so the same entities are used to account them, although for different arrays a set of attributes with certain and indeterminate data may differ substantially.

The component operates with basic entity "Communities" and the entity "Community technology characteristics" that corresponds to the group of characteristics VT. The entity "Search characteristics of the Community" executes the "many-to-many" relation between the "Communities" and the "Search engine", as well as stores typical specialized queries for searching within the community in a particular GPS.

The structure of the DB "Generators" is based on the formal definition of the generator and its relations with communities and individual discussions. The entity "Elements of the generator" simultaneously provides possibility of storing information about individual discussions in social media without referring to a particular generator.

Key words: university, information activities, information image, Internet marketing.

1. Пелещишин А. М. Визначення комплексу показників віртуальної спільноти для вищих навчальних закладів / А. М. Пелещишин, Р. О. Корж, О. Р. Трач // Східно-Європейський журнал передових технологій. – Харків, 2014. – № 2/2 (68). – С. 16–23. Бібліогр.: 12 назв. 2. Жежнич П. І. Часові бази даних (моделі та методи реалізації): монографія / П. І. Жежнич. – Львів: Вид-во Львівської політехніки, 2007. – 260 с. 2. Пасічник В. В. Глобальні інформаційні системи та технології (моделі ефективного аналізу, опрацювання та захисту даних) / В. В. Пасічник, П. І. Жежнич, Р. Б. Кравець, А. М. Пелещишин, Д. М. Тарасов. – Львів: Вид-во Львівської політехніки, 2006. – 350 с. 3. Жежнич П. І. Консолідовані інформаційні ресурси баз даних та знань / П. І. Жежнич. – Львів: Видавництво Львівської політехніки, 2010. – 212 с. 4. Жежнич П .І. Технології інформаційного менеджменту: навч. посіб. / П. І. Жежнич. – Львів: Вид-во Львівської політехніки, 2010. – 260 с. 5. Співаковський О. В. Побудова ІКТ інфраструктури ВНЗ: проблеми та шляхи вирішення / О. В. Співаковський et al. – 2014. 6. Лисенко В. С.; Єгоров С. О.; Рудницький Є. А. Інформаційно-комунікаційні інструменти мережі трансферу технологій вищих навчальних закладів України. Математические машины и системы. – 2014. – 4. 7. Стрюк А. М.; Рассовицька М. В. Система хмаро орієнтованих засобів навчання як елемент інформаційного освітньо-наукового середовища ВНЗ. Інформаційні технології і засоби навчання. – 2014. – 4.42. – 150–158. 8. Спірін О. М.; Новицька Т. Л., Лупаренко Л. А. Науково-методичний та координаційний супровід розвитку інформаційного освітньо-наукового простору України // Комп'ютер у школі та сім'ї. – 2015. – 125.5. – 11–17. 9. Корж Р. О. Аналіз сучасних видів та методів інформаційної діяльності ВНЗ в Інтернеті / Р. О. Корж, А. М. Пелещишин // Східно=Європейський журнал передових технологій. – Харків, 2013. – № 2/2 (62). – С. 62–65. 10. Пелещишин А. М. Базові характеристики та модель підрозділів ВНЗ як суб'єкта інформаційної діяльності / А. М. Пелещишин, Р. О. Корж // Східно-Європейський журнал передових технологій. – Харків, 2015. – № 2/2 (74). – С. 27–34. 11. Korzh R. Protection of University Information Image from Focused Aggressive Actions / R. Korzh, A. Peleshchyshyn, S. Fedushko , Y. Syerov // Advances in Intelligent Systems and Computing: Recent Advances in Systems, Control and Information Technology, Proceedings of the International Conference SCIT 2016, May 20–21, 2016, Warsaw, Poland // Springer International Publishing. – 2017. – Vol. 543. – P. 104–110.

10. Верес О. М., Оливко Р. М. Класифікація методів аналізу великих даних

CLASSIFICATION OF METHODS FOR THE BIG DATA ANALYTICS

Oleh Veres, Roman Olyvko
Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail: Oleh.M.Veres@lpnu.ua

The rapid spread of intelligent and interconnected devices and systems stimulates the amount of collected data to grow at an alarming rate. To maintain competitiveness, innovation and the rapid withdrawal of products and services to the market it is of paramount importance to be able to quickly and economically analyze data and get through this analytical information. The current solutions do not provide the speed of system response needed to work on tasks of analysis of the big data and that is why it reduces user productivity and delaying decisions.

Based on the definition of Big Data it is possible to formulate basic principles of such data: horizontal scalability, resistant to failures, locality data. All modern means of working with Big Data meet these three principles to certain extend. In order to observe them we have to select methods, ways and means of development paradigms for

data processing. Today a different set of available techniques for analyzing data sets is available and is primarily based on tools borrowed from the Statistics and Informatics.

Nowadays there is no difference in the use of the term Big Data and Big Data Analytics. These terms describe the data, management technologies, and methods of analysis. Big Data Analytics is the development of the concept of Data Mining. The same tasks, scope, data sources, methods, and technologies are incorporated here. If to supplement the approach Data Mining with MapReduce technology as well as with the requirement 4V (Volume, Velocity, Variety, Veracity), it will display the functional connection Big Data Analytic.

Groups of methods and technologies, analytics Big Data that are graded according to the functional relationships and formal model of information technology are divided into: methods for Data Mining, Technology Text Mining, MapReduce technology, data visualization, and other technologies and methods of analysis.

Data Mining is a set of techniques that allows you to determine the most suitable product, promoted or service category consumers identify the characteristics of the most successful workers predict behavioral patterns of consumers. Data Mining Methods are divided into two groups: Supervised Learning and Unsupervised Learning. Another classification divides the diversity of Data Mining methods into two subgroups: statistical and cybernetic methods. Statistical methods for Data Mining are divided into four groups of methods: descriptive analysis; the initial data analysis; analysis of the links; Multivariate statistical analysis; time series analysis. The group of cybernetic methods is such: evolutionary programming; associative memory; fuzzy logic; decision trees; processing systems of expert knowledge; Artificial neural networks; genetic algorithms.

Visual presentation of the analysis of Big Data is essential for their interpretation. As their technology, Big Data uses visualization methods, such as Tag Cloud, Clustergram, History Flow, Spatial Stream.

The basis of the technology Text Mining is the statistical and linguistic analysis, methods of artificial intelligence. This technology is used for analysis, providing navigation and search unstructured text. The main methods of Text Mining Technology are classification; clustering; Relationship, Event and Fact Extraction; feature extraction; summarization; question answering; thematic indexing; keyword searching; of taxonomies and thesauri.

Other technologies and research methodologies applied to Big Data are: A/B testing, Splittesting; Natural Language Processing (NLP); Sentiment Analysis; Network Analysis; Optimization; Pattern Recognition; Predictive Modeling; Signal Processing; Spatial Analysis; Statistics; Simulation; Crowdsourcing; Data Fusion and Data Integration.

Creating and maintaining data warehouses in terms of terabytes, petabytes or more is made possible through the technology of distributed file systems. In terms of implementation, analytic platform for Big Data the most reasonable is the usage of new technologies MapReduce.

The abovementioned description of the methods and technologies enabled us to build tools Protégé-OWL ontology analysis of the Big data.

Key words: analysis, Big data, visualization data, model, Data Mining, Text Mining, MapReduce.

1. Mayer-Shenberher V. Bol'shye dannye. Revolyutsyya, kotoraya yzmenyt to, kak my zhyvem, rabotaem y myslym / Vyktor Mayer-Shenberher, Kennet Kuker ; per. s anhl. Ynny Haydyuk. – M. : Mann, Yvanov y Ferber, 2014. – 240 p. Майер-Шенбергер В. 2. Bol'shye dannye i analytyka [Electronic resource]. – Mode of access: http://www-03.ibm.com/systems/ru/technicalcomputing/bigdata.html 3. Ageev Anna. Analitiki predupredili ob opasnosti bolschyh danych [Electronic resource] . – Mode of access: http://bigdata.cnews.ru/news/top/2015-10-23_eksperty_predosteregayut_ot_nepravilnogo_obrash cheniya. 4. Nazvany prichiny tormosheniya rynka bolschyh danych [Electronic resource]. — Mode of access : http://bigdata.cnews.ru/news/top/2015-11-20_analitiki _otsenili_tempy_rosta_mirovogo_rynka. 5. Koén Dzh. MOHuchye sposobnosty: novye pryemy analyza bol'shykh dannykh [Electronic resource] / Dzheffry Koén, Brayen Dolén, Mark Danlép, Dzhozef Khellersteyn, Keyléb Velton; per. s anhl. Serhey Kuznetsov. – Mode of access: http://citforum.ru/database/articles/mad_skills/ 6. History and evolution of big data analytics [Electronic resource]. – Mode of access: https://www.sas.com/en_us/ insights/analytics/big-data-analytics.html 7. Mitchell R. 8 big trends in big data analytics [Electronic resource] / Robert L. Mitchell // Computerworld, OCT 23, 2014. – Mode of access : http://www.computerworld.com/article/ 2690856/big-data/8-big-trends-in-big-data-analytics.html 8. Bol'shye danye (Big Data) [Electronic resource]. – Mode of access: http://tadviser.ru/a/125096. 9. Inmon W. H. Big Data – getting it right: A checklist to evaluate your environment / [Electronic resource] / W. H. Inmon. // DSSResources.COM, – 2014. – Mode of access: http://dssresources.com/papers/features/inmon/inmon01162014.htm. 10. Shakhovska N. B. Organizacya velykych danych u rozpodilenomu seredovyshi / N. B. Shakhovska, Yu .Ja. Bolubash, O. M. Veres // Obchyslyuvalna technika ta avtomatyzaciya: [zb. nauk. pr. DonNTU]. – Donezk, 2014. – S. 147–155. – (Visnyk / DonNTU) ; No. 2 (27). 11. Shakhovska N. B. Big Data Federated Repository Model / N. B. Shakhovska, Yu. Ja. Bolubash, O. M. Veres // The Experience of Designing and Application of CAD Systems in Microelectronics (CADMS'2015) Proc. ofthe XIII-thInt. Conf., (Polyana–Svalyava (Zakarpattya), Ukraine, 24–27 February, 2015). – Lviv: Publishing Lviv Polytechnic, 2015. – P. 382–384. 12. Veres O. Elements of the Formal Model Big Data / Oleh Veres, Natalya Shakhovska // Perspective Technologies and Methods in MEMS Design, MEMSTECH 2015 – Proceedings of 11th International Conference,

2–6 veresnya 2015, Lviv / Naz. un-t "Lviv. politechnika". – Lviv : Vyd-vo Lviv. politechniky, 2015. – S. 81–83. 13. Shakhovska N. Data space architecture for Big Data managering / N. Shakhovska, O. Veres, Y. Bolubash, L. Bychkovska-Lipinska // Xth International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT'2015), pp. 184-187, Lviv, 2015. DOI: 10.1109/STC-CSIT.2015.7325461 14. Shakhovska N. Generalized formal model of Big Data / N. Shakhovska, O. Veres and M. Hirnyak, // ECONTECHMOD: an international quarterly journal on economics of technology and modelling processes, vol. 5, no. 2, 2016, – pp. 33–38. 15. Shakhovska N. Big Data Information Technology and Data Space Architecture / N. Shakhovska, O. Veres, Y. Bolubash // Sensors & Transducers, vol. 195, no. 12, pp. 69–76, 2015. 16. Barsehyan A. A. Analyz dannykh y protsessov / A. A. Barsehyan, M. S. Kupryyanov, Y. Y. Kholod, M. D. Tess, S. Y. Elyzarov. Z-e yzd. pererab. y dop. – SPb.: BKHV-Peterburh, 2009. 512 p. 17. Paklyn N. B. Byznes-analytyka: ot dannykh k znanyyam (+ SD) / N. B. Paklyn, V. Y. Oreshkov. – SPb.: Yzd. Pyter, 2009. – 624 p. 18. Dyuk V. Data Mining: uchebnyy kurs (+CD) / V. Dyuk, A. Samoylenko. – SPb.: Yzd. Pyter, 2001. – 368 p. 19. Manyika J. Big data: The next frontier for innovation, competition, and productivity / Manyika James. Mc Kinsey Global Institute, June, 2011. – 156 p. 20. Zhuravlëv YU. Y. Raspoznavanye. Matematycheskye metody. Prohrammnaya systema. Praktycheskye prymenenyya. / YU. Y. Zhuravlëv , V. V. Ryazanov, O. V. Sen'ko. – M. : Yzd. "Fazys", 2006. – 176 p. 21. Zynov'ev A. YU. Vyzualyzatsyya mnohomernykh dannykh / A. YU. Zynov'ev. – Krasnoyarsk: Yzd. Krasnoyarskoho hosudarstvennoho tekhnycheskoho unyversyteta, 2000. – 180 p. 22. Chubukova Y. A. Data Mining: uchebnoe posobye / Y. A. Chubukova. – M. : Ynternet-unyversytet ynformatsyonnykh tekhnolohyy: BYNOM: Laboratoryya znanyy, 2006. – 382 p. 23. Sytnyk V. F. Intelektual'nyy analiz danykh (deytamayninh): Navch. posibnyk. / V. F. Sytnyk, M. T. Krasnyuk. – K.: KNEU, 2007. – 376 p. 24. Ian H. Witten. Data Mining: Practical Machine Learning Tools and Techniques. / Ian H. Witten, Eibe Frank, Mark A. Hall . – 3rd Edition. – Morgan Kaufmann, 2011. – 664 p. 25. Marr B. Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance / Bernard Marr. – John Wiley&Sons Ltd, 2015. – 256 p. 26. Einav L. The Data Revolution and Economic Analysis [Electronic resource] / Liran Einav, Jonathan Levin // NBER Working PaperNo. 19035, 2013. – Mode of access : http://www.nber.org/chapters/c12942.pdf 27. Vanyashyn A. Za bol'shymy dannymy sledyt PANDA / A. Vanyashyn, A. Klymentov, V. Koren'kov // "Superkompyutery", No. 3 (11), 2013. С. 56–61 28. Serov D. Analytyka "bol'shykh dannykh" – novye perspektyvy [Electronic resource] / Denys Serov // "StorageNews", No. 1 (49), 2012. – Mode of access : http://www.storagenews.ru/49/EMC_BigData_49.pdf 29. Ronen Sh. Links that speak: The global language network and its association with global fame [Electronic resource] / Shahar Ronen, Bruno Gonçalves, Kevin Z. Hu, Alessandro Vespignani, Steven Pinker, César A. Hidalgo // PNAS, Vol. 111, No.52, 2014. – Mode of access : http://stevenpinker.com/files/pinker/files /pnas_hildago_et_al_global_language_network_2014.pdf 30. Aflalo Y. Spectral multidimensional scaling [Electronic resource] / Yonathan Aflalo, Ron Kimmel // PNAS, vol. 110, no. 45, November 5, 2013. – Mode of access : http://www.cs.technion.ac.il/~ron/PAPERS/Journal/ AflaloKimmelPNAS2013.pdf 31. Gadepally V. Big Data Dimensional Analysis [Electronic resource] / Vijay Gadepally, Jeremy Kepner. arXiv:1408.0517v1. – Mode of access : https://arxiv.org/pdf/1408.0517v1.pdf 32. Weinstein M. Analyzing Big Data with Dynamic Quantum Clustering [Electronic resource] / M. Weinstein, F. Meirer, A. Hume, Ph. Sciau, G. Shaked, R. Hofstetter, E. Persi, A. Mehta, D. Horn. arXiv:1310.2700.– Mode of access : https://arxiv.org/ftp/arxiv/papers/1310/1310.2700.pdf 33. Paklyn, N. B.Byznes-analytyka: ot dannykh k znanyyam [Tekst] : ucheb. posobye / N. B. Paklyn, V. Y. Oreshkov. – 2-e yzd., yspr. — SPb. : Pyter, 2013. – 702 p. 34. Zhelyazny D. Hovory na yazyke dyahramm : posobye po vyzual'nym kommunykatsyyam dlya rukovodyteley / D. Zhelyazny. – M. : Ynstytut kompleksnykh stratehycheskykh yssledovanyy, 2004. – 220 p. 35. Roém D. Praktyka vyzual'noho myshlenyya. Oryhynal'nyy metod reshenyya slozhnykh problem / D. Roém. – M . : Mann, Yvanov y Ferber, 2014. – 396 p. 36. Tafty É. Predstavlenye ynformatsyy [Electronic resource] / É. Tafty. – Mode of access : http://envisioninginformation.daiquiri.ru/15 37. Yau N. Yskusstvo vyzualyzatsyy v byznese. Kak predstavyt slozhnuyu ynformatsyyu prostymy obrazamy / N. Yau. – M. : Mann, Yvanov y Ferber, 2013. – 352 p. 38. Iliinsky N. Designing Data Visualizations / N. Iliinsky, J. Steele. – Sebastopol :O'Reilly, 2011. – 110 p. 39. Krum R. Cool infographics: effective communication with datavisualization and design / R. Krum. – Indianapolis: Wiley, 2014. – 348 p. 40. T'yuky Dzh. Analyz rezultatov nablyudenyy: razvedochnyy analyz / Dzh. T'yuky / Pod red. V. F. Pysarenko. – M.: Myr, 1981. – 693 p. 41. Alper C. New Software for Visualizing the Past, Presentand Future [Electronic resource] / C. Alper, K. Brown, G. R. Wagner // DSSResources.COM, 09/23/2006. – Mode of access :http://dssresources.com/ papers/ features/alperbrown&wagner/alperbrown&wagner09212006.html 42. Barsehyan A. A. Analyz dannykh y protsessov: ucheb. posobye / A. A. Barsehyan, M. S. Kupryyanov, Y. Y. Kholod, M. D. Tess, S.Y. Elyzarov. – 3-e yzdanye pererab. y dop. – SPb.: BKHV-Peterburh, 2009. – 512 c. 43. Text Mining [Electronic resource] . – Mode of access: http://statsoft.ru/home/textbook/modules/sttextmin.html#index 44. Landé D. Hlubynnyy analyz tekstov tekhnolohyya éffektyvnoho analyza tekstovykh dannykh [Electronic resource] / Dmytryy Landé. – Mode of access : http://visti.net/~dwl/art/dz/ 45. Barsehyan A. A. Tekhnolohyy analyza dannykh. Data Mining, Visual Mining, Text Mining, OLAP / A. A. Barsehyan, M. S. Kupryyanov, V. V. Stepanenko, Y. Y. Kholod. – 2-e yzdanye pererab. y dop. – SPb.: BKHV-Peterburh, 2007. – 384 p. 46. Lynyuchev P. Text Mining: sovremennye tekhnolohyy na ynformatsyonnykh rudnykakh [Electronic resource] / Pavel Lynyuchev // PC Week/RE No. 6 (564), 27 fevralya –

5 marta 2007. – Mode of access: https://www.pcweek.ru/idea/article/detail.php?ID=82081 47. Pleskach V. L. Informatsiyni systemy i tekhnolohiyi na pidpryyemstvakh : pidruchnyk / V. L. Pleskach, T. H. Zatonats'ka. – K. : Znannya, 2011. – 718 p. 48. Stounbreyker M. MapReduce y parallelnye SUBD: druzya yly vrahy? [Electronic resource] / Maykl Stounbreyker, Dényél Abady, Dévyt Devytt, Sém Médden, Éryk Paulson, Éndryu Pavlo, Aleksandr Razyn ; per. s anhl. Serhey Kuznetsov // Communications of the ACM, vol. 53, no. 1, January 2010.– Mode of access: http://citforum.ru/database/articles/mr_vs_dbms-2/ 49. Berezyn A. Map-Reduce na prymere MongoDB [Electronic resource] / Anton Berezyn. – Mode of access : https://habrahabr.ru/post/184130/ 50. Lebedenko E. Tekhnolohyya Google MapReduce: razdelyay y vlastvuy [Electronic resource] / Evhenyy Lebedenko // – Mode of access : http://www.computerra.ru/82659/mapreduce/ 51. Pavlo É. Sravnenye podkhodov k krupnomasshtabnomu analyzu dannykh [Electronic resource] / Éndryu Pavlo, Éryk Paulson, Aleksandr Razyn, Dényél Abady, Dévyd Devytt, Sémyuél Médden, Maykl Stounbreyker; per. s anhl. Serhey Kuznetsov. – Mode of access: http://citforum.ru/database/articles/mr_vs_dbms/2.shtml 52. BigData ot A do YA. Chast 1: Pryntsypy raboty s bolshymy dannymy, paradyhma MapReduce [Electronic resource]. – Mode of access: https://habrahabr.ru/company/dca/blog/267361/ 53. Big data ot A do YA. Chast 3: Pryemy y stratehyy razrabotky MapReduce-prylozhenyy [Electronic resource] . – Mode of access : https://habrahabr.ru/company/dca/blog/270453/ 54. Havrylova T. A. Bazy znanyy yntellektualnykh system / T. A. Havrylova, V. F. Khoroshevskyy. – SPb. : Pyter, 2000. – 384 p. 55. Havrylova T. A. Ontolohyya dlya yzuchenyya ynzheneryy znanyy // Trudy Mezhdunarodnoy nauchno – praktycheskoy konferentsyy KDS-2001. – 2001. 56. Havrylova T. A. Ontolohycheskyy podkhod k upravlenyyu znanyyamy pry razrabotke korporatyvnykh ynformatsyonnykh system // "Novosty yskusstvennoho yntellekta". – 2003. No. 2. – P.24-30. 57. Lytvyn V. V. Bazy znan intelektualnykh system pidtrymky pryynyattya rishen: monohrafiya / V. V. Lytvyn; Ministerstvo osvity i nauky, molodi ta sportu Ukrayiny, Natsionalnyy universytet "Lvivska politekhnika". – Lviv : Vydavnytstvo Lvivskoyi politekhniky, 2011. – 240 p.

11. Заяць В. М., Заяць М. М. Образний підхід до обчислення кількості інформації та оцінки її цінності

THE FIGURATIVE APPROACH TO CALCULATE THE AMOUNT OF INFORMATION AND ESTIMATES ITS VALUES

Vasyl Zaiats[1], Mariia Zaiats[2]

[1] Department of Telecommunication, Computer Science and Electrical Engineering, 7, al. Akad. Kaliskiego, Bydgoszcz, 85-785, Poland., [2] Information Systems and Networks Department, Lviv Polytechnic National University, S. Bandery Str., 12, Lviv, 79013, UKRAINE, E-mail: [1]zvm01@ramblerl.ru , [2]mariyazayats@gmail.com

The term information has multiple definitions nowadays [1-10] (from the Latin word informatio – clarification, presentation). The most general definition has it as "a reflection of the real world" while the narrowest holds it as "a set of data, messages, information, symbols, images subject to processing, organizing and learning and using them". Since this concept is used in almost all spheres of human activity, a particular definition is used depending on the objectives of the study. In practical terms, the most appropriate definition of information seems the one made by the French scholar Buaze who describes it as anything that reduces the uncertainty of our knowledge of the object, phenomenon or process. Thus, information is the raw material that is connected with the data collection. Knowledge formation and sahping presupposes certain considerations that arrange the data on the basis of their comparison and classification. To form a scientific theory, clear and unambiguous definitions that would exclude any ambiguity within that branch of science application should be given.

Precise definitions of terms for the scientific language are based on two approaches. The first approach (theoretical) is often used in mathematical disciplines where the definition begins with the formulation of the basic postulates or hypotheses. All other displays of complex nature may be initially introduced in case they are not contradictory. The second approach to the formulation of definitions is typical for experimental sciences and is often called operational (practical). This approach is considered appropriate to be introduce in the scientific language of these values that can be determined experimentally or evaluated so indirectly through their impact on the behavior of the system. The terms of innovations that defy the definition of operational and have no practical value are excluded from the scientific vocabulary. In particular, it was the case with notion of ether, the theory of which had no practical content, thus was eventually excluded from the use in scientific circles, in spite of numerous publications on this topic.

The purpose of this study is to compare the existing methods and approaches to determining the information database and assessing its value, to indicate the limits of their applicability and offer hypothetical new approach to quantitative assessment of the value of information.

In classical information theory, developed in the fundamental works of R. Hartley [2] and C. Shannon [3], the basic concepts are based on probability theory. According to the theory R. Hartley, if the system can be in N states, the total amount of information can be defined as a binary logarithm of the number of states:

$$I = Log_2 N \qquad (1)$$

In this case, the unit of measurement of information is one bit. Thus, the bit of information can be defined as the amount of information contained in messages with two possible states. Note that the probability of these messages may be different.

If the probability of characters of the message is different, the formula Shannon [3] can be used to calculate the amount of information

$$I = -k \cdot \sum_i^n p_i \cdot Log_2 p_i , \quad (2)$$

where p_i is the probability of i-quality characteristics. In the case, if the probability of each of the signs are equal, then

$$I = -k \cdot Log_2 \frac{1}{n} = k \cdot Log_2 n , \quad (3)$$

that corresponding to the formula (1), because $N = n^k$.

Thus, the formula (1), (2) and (3) allow calculating the amount of information as with the equal and non-equal probability of qualitative characteristics regardless of the quality received of information (messages, sound, images or any nature). Such approach, however, does not take into account the value of the information that is largely subjective and depends on the characteristics of an object and purposes of the user. In fact, it is the charge for the calculation of the amount of information that is based on statistical considerations, regardless of its quality and value.

A fundamentally different approach to the definition of information is proposed by a Russian scientist AM Kolmogorov [4]. His algorithmic information theory is based on the concept of algorithm complexity transformation of one object to another. This approach aims at establishing the principle of balance and interconnection between the objects studied and the length of the program that is conducting the study. The amount of information in the theory of algorithms of transforming one object into another is determined by the length of the program, which enables transformation of the object A to B:

$$I = f[G(A,B)], \quad (4)$$

where G – the conversion program object A object B; f – function that determines the length of the conversion program in bits.

Note that the amount of information on this approach greatly depends on the choice of structural element selected for transformation. The amount of information is at its maximum, if the selected element conversion is pixel and the minimum is while choosing one letter as an element; it will be considered medial in case of choosing the letter segment as an element for processing. Measuring the same amount of data for a single letter is quite impossible, because of the algorithm (4) which needs two objects. Comparison with the same letter to itself is senseless as well, since the length of the conversion program in this case is zero.

Thus, the obvious disadvantages of algorithmic approach are:

a) while measuring the amount of information, the length of the conversion program is used as a parameter; the length of the conversion program essentially depends on the structure of the elements that enable it to convert one object into another. Consequently, the smaller structure elements are selected, the longer the program of conversion for the same objects becomes;

b) one and the same conversion program can be applied to a set of objects, analysis of which requires different amounts computer performance commands, while the length of the program remains unchanged;

c) non-defined method for measuring the amount of information when considering an individual object; to specify, a measurement algorithm (8) requires two;

d) loss of additively as a property of information when considering interacting systems, which significantly complicate the analysis.

These disadvantages are eliminated in the approach, which is based on the use of methods of theory of recognition and operates with the concept of information and amount of information that is different from the classical definitions derived from the theory of probability.

The amount of information I*o* contained in a certain image, which is successfully recognized by a cybernetic machine, is defined by the formula:

$$I_o = q^{-1} F[G(O)], \quad (5)$$

where q – probability of correct recognition image; F – expanded function of length (including cycles) image recognition software; G – length of image recognition program, expressed in bits.

Obviously, according to approach (5) there appears the deductibility of information of sophisticatedly structured images, such as individual words, sentences, texts or drawings. The amount of information substantially depends on the length of the expanded program of recognition, the probability of correct recognition and functional opportunities of cybernetic machine that implements the procedure of recognition.

Regarding the value of information, it is probable in cases of a particular purpose of obtaining information for the user, i.e. the attainment of some objective function. In [5] A.A. Harkevich proposes to calculate the value of information as

$$F = Log_2 \frac{p_0}{p_i}, \tag{6}$$

where p_0 – the probability of a correct solution to the problem prior to obtaining the information, and p_1 – the probability of a correct solution to the problem after receiving the information. This approach has a right to exist, although it raises some doubt about the effectiveness and appropriateness of use. Firstly, the unit of measurement of the value of information in this approach is bit; the same is the case with calculating the amount of information. Obviously, with introduction of new parameter there should also appear a new dimension; otherwise deduction should be made in dimensionless units. Secondly, the formula (6) cannot be considered objective as the assessment p_0 so p_1 will be made by the user. Thirdly, the method of estimating these probabilities is not obvious. Fourthly, the value of information is dynamic [7], and should be variable while receiving the information. Obviously, the calculation value of information as an objective function of achieving the set goal should presuppose meeting certain needs of the user (physical, spiritual, aesthetic, cognitive, etc.) or performing certain actions.

Given the variety of needs of the user, the most appropriate approach appears to be the one in which the value of information is calculated in the percentages: 100% – if the information is valuable; 0 % – provided the target function is not reached. Thus, the value of information F for the current value of the objective function Z and efficiency E, that achieved after receiving the message at a given time can be embedded in a rule:

$$F = \begin{matrix} 100\%, & якщо & Z - E = 0 \\ 0\%, & якщо & Z - E > 0 \end{matrix}. \tag{7}$$

To improve the accuracy of deducting value of information it is reasonable to have a formed objective function for the perspective; and to gradually expand the scale of calculating the value of information. The first condition significantly increases the value of the information after receiving the entire message, and the second can be realized using the formula:

$$F = \frac{E}{Z} \cdot 100\%. \tag{8}$$

Equation (8) is a specification (7), as confirmed by the examination of specific examples of consideration of queuing systems and game theory.

Obviously, for the implementation of this approach for calculating the value of information, the declarative programming languages (Lisp, Prolog, or derivatives thereof depending on the specific subject of the problem) should be used [26–34], which is most successfully adapted to implementing logic functions of the form (4)–(8) and can solve problems related to qualitative analysis and recognition of objects with complex structures (handwriting recognition, handwriting, psycho physiological state of a person, receiving, treatment, analysis of storage, processing and transformation of information) [8-25, 35-37]; and implement relevant definite target functions into a particular applied or scientific problem.

Key words: information theory, amount of information, probability, algorithm, image, value of information.

1. Bryllyuen L. Science and information theory / L. Bryllyuen. – M .: State Publishing House, 1960. – 392 p. 2. Hartly R. V. Information theory and applications / R. V. Hartly. – M .: Fyzmathyz, 1959. – 356 p. 3. Shannon K. Work on information theory and cybernetics / K. Shannon. – M .: Publishing For. Lit., 1963. – 286 p 4. Kolmogorov A.N. Approach to definitions. for concepts "Quantity of information" / A. N. Kolmogorov // Problems transfer information. – Vol. 1. – 1965. – P. 63–67. 5. Kharkevich A. A. On Values of information / A. A. Kharkevich // Problems of Cybernetics. – Vol. 4.– Fyzmathyz. – 1960. – P. 53–57. 6. Partyko Z. V. Imaginative concept of information theory / Z. V. Partyko.– Lions: The publishing house LNU. Franko, 2001. – 98 p. 7. Bongard M. M. Oh concept "value info" / M. M. Bongard // Problems of Cybernetics. Vol. 8. – Fyzmathyz. – 1963. – S. 71–102. 8. Tesler G. S. New Science / G.S. Tesler. – K .: Logos, 2004. – 404 p. 9. Syavavko M. S., M. I. Rozhankivska. Entropy as a measure of fuzziness fuzzy set / M. S. Syavavko, M. I. Rozhankivska // Coll. scientific papers LDINTU by Chornovil. – No. 2. – 2009. – S. 3–19. 10. De Luca A. On the convergence of entropy measures of fussy sets / A. De Luca, S. Termini // Kybernetes. – Vol. 6. – 1971. – P. 219–227. 11. G. Schuster Determation Chaos: Introduction. / H. Shuster // Trans. c anhl.– M.: Mir, 1988. – S.240. 12. Zaiats V. The mathematical description of the computer user recognition / V. M. Zaiats, M.M. Zaiats // Coll. "Physical modeling and information technology." – Lviv. – 2005 – Vol. 1.– S. 146–152. 13. Kharkevich A.A. Recognition Images / A. A. Kharkevich // Radyotehnyka.– 14.– 1959.– Tom S. 15–19. 14. Fukunaha C. Introduction to Theory Statistic Recognition / C Fukunaha., M .Kar. Nauka, 1979. – 512 p. 15. Gorelik A. L. Methods recognition / A. L. Horelyk, V. A. Skrypnyk. – M.: Higher School, 1989.– S.232. 2009. – S. 3–19. 16. Duda R. Recognition images and analysis stages. / R. Duda, P. Hart.– M .: Piece, 1976.– 512 p. 17. Zaiats V. M. Identifying priority deterministic attributes when building the system of recognition of objects /

V. M. Zayats, O. Shokyra // Coll. works scientific-practical conference. LDINTU of V. Chornovil "Mathematical modeling of complex systems." – 2007. – S. 135–137. 18. Zayats V. M. Architecture event-oriented systems on the example of handwriting recognition / V. M. Zayats, D. O. Ivanov // Herald "Lviv Polytechnic" "Computer Engineering and Information Technology". – Lviv. – 2004. – No. 530. – S. 78–83. 19. Tomaszewskyy O. M. Methods and algorithms for information security system based on neural network technology. Author. ... dis. tech. Sciences: 05.13.23. / O. M.Tomashevskyy – Lviv, 2002.– p. 20. 20. Chala L.E. Comparative Analysis Authentication Methods. Computer journal of Man-Machine Studies.– 1988 – V. 28, No. 1.– P. 67–76. 21. Cohell O. Biometric Identification System Based in Keyboard Filtering / O. Cohell, J.Badia, G. Torres // Proc. Of XXXIII Annual IEEE International Carnahan Conference of Security Technology / – 1999.– P. 203–209. 22. Wolf O. B. Mat. Problems as specific fonts copyright / O. B. Wolf // Herald "Lviv Polytechnic", "Information Systems and Networks". – 2008. – No. 610. – P. 85–83. 23. Platonov A. V. The use of expert situational models in the field of national security. / A. V.Platonov, I. V.Baklan, K. V.Kramer // Coll. Intern jobs. scientific conference. ISDMCI '2008. – Vol. 1, Yevpatoriya. 2008. – P. 39–43. 24. Suzdal A. I. Definition Psycho Physical State operational staff by writing to keyboard nefte presiding mini-factories / A. I. Suzdal, V. A.Lobanov, V. G. Abashin // Nefte gazes work.– 2006.– P. 1—6. 25. Zaiats V. Structural method of hand-written text recognition / V. Zaiats, D. Ivanov // Pros. International Conf. "The experience of designing and application of CAD systems in microelectronics". – Lviv–Polyana. – 2005. – P. 493–494. 26. Mccarthy J. Recursive functions of symbolic expressions and their computation by machine // Comm. ACM.:1960. – Vol. 3. – P. 184–195. 27. Badaev Y. The theory of functional programming. Languages Common Lisp and Auto Lisp. – Kyiv, 1999. – 150 p. 28. Zaiats V. Functional programming: Training. Textbook. – Lviv: Publishing House "Beskids Bit", 2003.– 160 p. 29. P. Henderson. Functional programming. Application and Implementation. – M.: Peace. – 1983. – 349 p. 30. Bratko I. Programming in Prolog language for Artificial Intelligence. – M: Williams, 2004. – 640 p. 31. In C. Solomon D. Using Turbo Prolog: Tran. From Eng. – Moscow, Peace, 1993, 608 p. 32. Zaiats V. Logic Programming: Part 1: Synopsis of lectures on "Logic Programming" for students basic direction 06.08.04 "Software Engineering". – Lviv: Publishing House of the National University "Lviv Polytechnic", 2002. – 48 p. 33. Macalester J. Artificial Intelligence and Prolog for micro computers. – M.: Engineering, 1990. – 240 p. 34. Zaiats V., Zaiats M. Logical and functional programming: Training. Manual. – Lviv: Publishing House "Beskids Bit", 2006.– 352 p. 35. Erich Gamma, Richard Helm, Ralih Johnson, John Vlissides. Wzorce projektowe, 2012. 36. Nilsson N. J. Principles of Artificial Intelligence. – Tioga. – Springer-Verlag. – 1980.– 164 p. 37. Stanislaw Wrycha i inni. UML 2.1. Cwiczenia. ISBN: 978-83-246-0612-12.

12. Гомон К. О., Дияк І. І., Копитко М. Ф. Паралельний алгоритм розв'язування задач теорії пружності

THE PARALLEL ALGORITHM FOR SOLVING PROBLEMS OF ELASTICITY

Katya Gomon, Ivan Dyyak, Maria Kopytko
Applied Mathematics and Informatics Department, Ivan Franko National University of Lviv, 1, Universitetska Str., Lviv, 79000, UKRAINE

Adequately reflect the actual behavior of structures requires new methods and approaches for the construction and study of new computer models and the reliability of the results. This problem requires a significant increase of the amount of information that is processed. This requires the use of high performance computing that causes increasing demand for computing resources. At the present stage of computing the transition from single-core processor computers, the possibility of increased productivity are already exhausted, computers with multicore processors, which are also a lot. Creating high performance computing technologies closely related to the use of multi-core and multiprocessor computer systems. Correspondence between algorithmic software and technical means of multi-core and multiprocessor computers determine the effectiveness of such technologies. The highest performance of such systems is achieved using algorithms and programs with parallel organization of calculations. Therefore creating a brand of new software that uses parallel algorithms for solving problems gives us that opportunity.

In this article the algorithm based on parallel computing for solving elasticity problem was created. The originality of this approach lies in the fact that parallelization occurs at all stages of solving the problem: setting input, building the mesh of finite element method (FEM), and solving the systems of linear equations, analysis of the results. Domain decomposition method with isoperimetric quadrilateral approximations of finite element method in each of the subregions is used. Global matrix of system of equations for the entire region is not formed and is represented through local matrices and vectors for subdomains using Boolean matrices of connectivity. For solving the systems of linear equations, the parallel algorithm based on the method of conjugate gradients is developed.

The algorithm is implemented by means of program language C ++ using library of parallel execution Message Passing Interface (MPI). Calculation is done on a cluster of National Lviv Ivan Franko University which includes 14 computer nodes and servers.

A proposed approach is tested on a model example. Calculations were carried out on one (sequential algorithm), five and ten knots. The sphere of task solving was divided into the corresponding number of subdomains

so that dimension SLE was similar in all cases, that is for sequential and parallel algorithms. The results showed that the larger dimension of SLE, the more efficient use of parallel algorithm as time solving of SLE is more than 99 % of the time solving of the whole problem.

Key words: the problem of the theory of elasticity, domain decomposition method, parallel computing, finite element method, parallel conjugate gradient method.

1. Grigorenko A. Ya. Domain Decomposition method with geterogeneouse approximation of solution for the problems of theory of elasticity / A. Ya. Grigorenko, I. I. Dyyak, I. I. Prokopyshyn // Applied mechanics. – 2008. – Vol. 44, No. 11. – P. 18–29. 2. Dyyak I. I. Numerical investigation of 2D problems of elasticity with boundary element method / I. I. Dyyak // Math. methods and phys. and mech. fields. – 1997. – T. 40. – P. 60–64. 3. Dyyak I. Numerical efficiency hybrid finite-boundary approximation of elasticity problems based on domain decomposition method/ Dyyak I., J. Makar., J. Prokopyshyn // Bulletin of Lviv University. – 2007. – No. 12: Series Applied Mathematics and Informatyks. – P. 93–100. 4. Khimich A. M. Hybrid algorithm for solving linear systems with direct methods ribbon matrix / A. M. Khimich, A. J. Baranov // Computer mathematics: Coll. Science. works. – 2013 – Vol. 2. – 87. 5. Khimich A. M. Grid computing technology for the mathematical modeling of ductile fracture / A. M. Khimich, V. V. Meadow, O. V. Popov, V. A. Sidoruk, O. V. Chistyakov // Artificial intelligence. – 2014. – No. 4. – S. 101–110. 6. A High Performance Message Passing Library [electronic resource]. – Access: http://www.open-mpi.org – 10.03.2017.

13. Кравець П. О. Ігрова модель прийняття рішень в ієрархічних системах

GAME MODEL OF DECISION-MAKING IN HIERARCHICAL SYSTEMS

Petro Kravets

Computer Science Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, AN E-mail: krpo@i.ua

The majority of the modern distributed systems are constructed by a hierarchical principle. Such systems consist of the active elements functionally connected among them under the scheme “head-executor”. It is natural that in such systems there is a necessity for acceptance and performance of supervising decisions.

The structure of hierarchical system can be represented in the form of a tree which knots are represented by subjects of decision-making, and communications between them show rank subordination. The subject with the highest rank (a tree root) develops the decision and leads up them to data to subjects of an average level that in turn develops the decision for subjects of the lowest level. Subjects of local level (correspond to tree leaves) have no subordinates and carry out the decision of a direct high level.

Decision-making process by subjects of system is influenced by uncontrollable factors, i.e. decisions are accepted in the conditions of uncertainty.

It is supposed that all decisions of the highest level are recommendatory and do not limit a freedom in choosing of subjects of the lowest level. For a deviation of own decision of the subject from the recommendatory the penalty is imposed. Development of the co-ordinated collective decision which minimises total expenses of system is the purpose of hierarchical system.

Process of collective decision-making in a hierarchical system leads to occurrence of competitive states which are studied by the theory of games, and in the conditions of uncertainty are studied by the theory of stochastic games.

In references to this problem basically fully connected stochastic games are investigated. Games with local communications between players (games on not fully connected graphs, trees), actual for modelling of decision-making processes in organizational and ergative systems require versatile and profound studying.

Processes of game decision-making in the conditions of uncertainty are object of this research.

Subject of this research is a game model of decision-making in hierarchically organised systems.

The work purpose is decisions of stochastic game for decision-making in hierarchical system and definition of factors which influence on convergence of a game method in the conditions of uncertainty.

Practical value of work consists in possibility of application of the received results for construction of effective organizational and cybernetic systems with hierarchical structure.

For the decision of a problem of hierarchical decision-making in the conditions of uncertainty it is used an adaptive game method in which uncertainty of parameters is compensated by their stochastic identification reached at the expense of self-learning. The game method is constructed on the basis of stochastic approximation of a complementary slackness condition.

The model developed in this work, the method and the algorithm of the decision of stochastic game provide the co-ordinated decision-making in hierarchical systems at restrictions on parameters which satisfy fundamental conditions of stochastic approximation. Coordination of strategies of players is reached during the decision of stochastic game on the basis of gathering of the current information and its adaptive processing.

Efficiency of decision-making in hierarchical system is supervised by means of characteristic functions of average losses, factor of the co-ordinated decisions and Euclidean norms of a deviation of the dynamic mixed strategies from optimum values. Falling off the function of average losses, growth of factor of the co-ordinated decisions and the reduction of norm of a deviation of the mixed strategies testifies to convergence of a game method according to the formulated purpose.

Dimension of a problem and parameters of a method of its solving define convergence rate of stochastic game. Optimisation of parameters of a game method taking into account restrictive conditions of stochastic approximation provides close to 1 exponential order of convergence rate.

Reliability of experimental results proves to be true repeatability of values of the calculated characteristics of stochastic game for different sequences of random variables.

The received results can be used for support of acceptance of the co-ordinated collective decisions in systems with the hierarchical organisation.

Key words: hierarchical system, decision-making, stochastic game, uncertainty condition.

1. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Моко, Я. Такахара. – М.: Мир, 1973. – 334 с. 2. Воронин А. А. Оптимальные иерархические структуры / А. А. Воронин, С. П. Мишин. – М.: ИПУ РАН, 2003. – 210 с. 3. Шарапов О. Д. Економічна кібернетика: навч. посіб. / О. Д. Шарапов, В. Д. Дербенцев, Д. Є. Семьонов. – К.: КНЕУ, 2004. – 231 с. 4. Саати Т. Принятие решений. Метод анализа иерархий / Т. Саати. – М.: Радио и связь, 1993. – 320 с. 5. Теорія і практика прийняття управлінських рішень / А. С. Крупник, К. О. Линьов, Є. М. Нужний, О. М. Рудик. – К.: Видавничий дім „Простір", 2007. – 119 с. 6. Катренко А. В. Теорія прийняття рішень : підручник з грифом МОН / А. В. Катренко, В. В. Пасічник, В. П. Пасько. – К. : Видавнича група BHV, 2009. – 448 с. 7. Кононенко А. Ф. Принятие решений в условиях неопределенности / А. Ф. Кононенко, А. Д. Халезов, В. В. Чумаков. – М.: ВЦ АН СССР, 1991. – 196 с. 8. Бурков В. Н. Теория активных систем: состояние и перспективы / В. Н. Бурков, Д. А. Новиков. – М. Синтег, 1999. – 128 с. 9. Айзерман М. А. Выбор вариантов: основы теории / М. А. Айзерман, В. Ф. Алескеров. – М.: Наука, 1990. – 240 с. 10. Данилов В. И. Механизмы группового выбора / В. И Данилов, А. И. Сотсков. – М.: Наука, 1991. – 172 с. 11. Нейман Дж. Теория игр и экономическое поведение / Дж. Нейман, О. Моргенштерн. – М.: Наука, 1970. – 708 с. 12. Гермейер Ю. Б. Игры с непротивоположными интересами. – М.: Наука, 1976. – 328 с. 13. Горелик В. А., Кононенко А. Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. М.: Радио и связь, 1982. – 144 с. 14. Кукушкин Н. С. Теория неантагонистических игр / Н. С. Кукушкин, В. В. Морозов. – М.: МГУ, 1984. – 104 с. 15. Воробьев Н. Н. Основы теории игр. Бескоалиционные игры / Н. Н. Воробьев. – М.: Наука, 1984. – с. – 496 с. 16. Мулен Э. Теория игр с примерами из математической экономики / Э. Мулен. – М.: Мир, 1985. – 200 с. 17. Губко М. В. Теория игр в управлении организационными системами / М. В. Губко, Д. А. Новиков. – М.: Синтег, 2002. – 148 с. 18. Новиков Д. А. Игры и сети / Д. А. Новиков // Математическая теория игр и ее приложения. – Т. 2, в.1. – 2010. – С. 107–124. 19. Эпштейн Г. Л. Теория игр: учеб. пособ. – М.: М ГУ ПС (МИИТ). – 2014. – 114 с. 20. Доманский В. К. Стохастические игры / В. К. Доманский // Математические вопросы кибернетики. – 1988. – № 1. – С. 26–49. 21. Fudenberg D. The Theory of Learning in Games / D. Fudenberg, D. K. Levine. – Cambridge, MA: MIT Press, 1998. – 292 pp. 22. Назин А. В. Адаптивный выбор вариантов / А. В. Назин, А. С. Позняк. – М.: Наука, 1986. – 288 с. 23. Weiss G. Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence / G. Weiss, editor. – Springer Verlag, Berlin, 1996. – 643 p. 24. Wooldridge M. An Introduction to Multiagent Systems / M. Wooldridge. – John Wiley & Sons, 2002. – 366 p. 25. Граничин О. Н. Введение в методы стохастической аппроксимации и оценивания: учеб. пособ. / О. Н. Граничин. – СПб.: Изд-во СПб. ун-та, 2003. – 131 с.

THE METHODS OF CONTEXT PROCESSING IN INTELLIGENT SYSTEMS

I. Zavuschak, E. Burov
[2]Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail: iryska2009@ukr.net, eugeneburov01@gmail.com

When problems are solved or tasks executed in the intelligent knowledge-based system, not all data are defined explicitly. A large part of them is assumed or taken from context. Thus, the intelligent system should correctly elucidate implicit data and knowledge and process it. The lack of explicit context definition and processing is the common reason why knowledge-based system fails.

Up to now, context processing is performed in linguistics, robotics, ubiquitous computing and mobile systems. However, there is no common understanding of context's definition between those different research areas. This creates inability to treat context from a unified point of view, develop methods for context formalization and processing viable across different application areas.

The dictionary sources define context as "interrelated conditions in which something exists or occurs" (Merriam-Webster dictionary) or "the situation within which something exists or happens, and that can help explain it" (Cambridge dictionary). Therefore, the concept of context implies the existence of some central agent (which can also be a multi-agent system) and the world in which this agent operates. The part of the world used for understanding of agent's behaviour or for making decisions by agent is the context of this agent. The scope of context is dynamically defined by agent's state and by agent's goals and limitations.

The purpose of this article is provide analysis of current context representation and processing methods and delineate yet unresolved problems and promising areas for research.

In practice context aware system takes information from numerous sources, integrates it with knowledge. On the other hand, knowledge itself is dependent from context and is applicable only within some predefined context. This is the reason why in some powerful expert systems, such as CYC, logical rules are true or false only within some context. Therefore, the important task for intelligent system is selection of relevant information and knowledge.

Current knowledge about context make distinction between contextual knowledge about how to execute a task (know how, surface knowledge) and knowledge about application domain (know that, deep knowledge). Also, it defines static and dynamic contexts. Existing definitions of context consider it as a product of interaction between objects. However, this approach does not provide the reason for selection of those interacting objects, nor does it explain why and how context changes. In our approach based on central intelligent agent, other interacting objects become the part of context. They are selected depending of goals pursued by agent, its state and the state of environment.

In this article, the formal methods for context representation and processing are analyzed. The "key-value" models provide information for interpreting variables in form of environment variables. Markup languages use hierarchical tag structures with attributes for context representation. Those context models are often specified in profile documents such as *Composite Capabilities / Preferences Profile (CC/PP), Comprehensive Structured Context Profiles (CSCP), Pervasive Profile Description Language (PPDL)* and others. Context logic is used when context is represented as series of axioms. This logic is an extension of first order logic where axioms are true only within some predefined context. Rule-based formalisms describe context as part of knowledgebase. The structure of knowledge is represented by rule packages. The flaw of rule-base knowledge representation is the lack of unified domain model, fragmentation of rule packages and complexity of keeping them relevant and conflict-free.

The implementation of context-aware intelligent systems implies the usage of some form of formal domain model and reasoning for creation of new facts and interpreting existing facts in domain. The deepness of reasoning is an important reasoning property. So, procedural context reasoning is usually shallow, because only a limited set of objects needed for a specific task at hand is taken into consideration. On the other hand, in deep reasoning, the knowledge about domain is used and number of objects and relations considered can be very large.

Different formal models used for representation of context have also different capabilities for supporting reasoning. Thus, simple models such as "key-value" models or markup language models are not suitable for extensive reasoning because of limitation introduced by syntax and semantics language. Rule-based models are used in complex reasoning systems (such as medical diagnostics). However, taking in consideration context makes related knowledge substantially more complex. Moreover, rule-based systems do not work with a domain model and therefore cannot use domain knowledge consistently.

Ontology-based models are mainstream now for context representation and reasoning, because ontology encapsulates domain knowledge within a single consistent knowledge structure. Additionally, there are a large number of reasoners available working with ontological models and using description logic as base for their operation. Context itself is often represented as small ontology or part of some larger ontology. The drawbacks of

using ontology in context-aware system source from limitations of reasoning tasks, which can be performed using description logic. Typically, this logic is applied for checking consistency, implementation of inheritance, building new concepts from existing using constructors. Other forms of reasoning such as causative and probabilistic reasoning while common in everyday practice are not supported by descriptive logic. Additionally, deep reasoning requires substantial computational resources which can become prohibitive taking in consideration time and processing power limitations in practice. Therefore, an important area of research is the reduction of context size prior to reasoning by including only relevant objects.

The notion of context has much in common with the notion of situation. In some classical works, situation and context are synonyms. The creation of situation-aware agent requires also an implementation of context awareness. Methods used for situation identification also require an identification of current context.

Several problems in the area of context-aware computing are still unresolved. Among most promising problems of context representation is the combined use of procedural context and domain knowledge depending on specific situation. For dynamic systems, it is important to track static and dynamic parts of context and minimize the size of dynamic part. In order to be able to implement deep reasoning in context it is important to resolve the problem of selecting only relevant information. One of possible ways to resolve this problem is usage of multi-layer context models.

In situation aware systems, it is important to research the influence of contextual data in all levels of JDL model. Also, the research of contextual dependencies for fuzzy, unreliable, or probabilistic data has still many unresolved problems. Finally, the study of context in different domains and building corresponding context models can contribute new knowledge which can be shared between domains.

Keywords: data, knowledge, context, database, knowledge base, the situation

1. Bazire M., July. Understanding context before using it / Bazire, M., Brézillon P // Proceedings of International and Interdisciplinary Conference on Modeling and Using Context. – Springer Berlin Heidelberg, 2005. – P. 29–40. 2. Brezillon P. Context in Artificial Intelligence: I. A Survey of the Literature / Brezillon P. // Computer & Artificial Intelligence. – 1999. – No. 4 – P. 321–340. 3. Chen, Guanling, and David Kotz. A survey of context-aware mobile computing research / Chen, Guanling, and David Kotz // Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College. Vol. 1. No. 2.1. – 2000. 4. Context. Retrieved 1 Dec. 2016 from http://www.merriam-webster.com/dictionary/context. 5. Meaning of "context" in the English Dictionary. Retrieved 1 Dec. 2016 from http://dictionary.cambridge.org /dictionary/english/context 6. Context. Retrieved 1 Dec. 2016 from http://www.dictionary.com/browse/context. 7. Context. Retrieved 1 Dec. 2016 from https://en.oxforddictionaries.com/definition/context_8. Dey A. K. Understanding and Using Context / Dey A. K. // Personal and Ubiquitous Computing. – 2001. – No. 1. – P. 4–7. 6. 9. Буров Є. В. Опрацювання контексту у когнітивній інформаційній системі, керованій моделями / Буров Є. В. // Східно-Європейський журнал передових технологій. – Харків: Технологічний центр. – 2010. – No. 1/7(43). – С. 40–47. 10. Brezillon P. Context in Artificial Intelligence: II. Key Elements of Contexts / Brezillon P. // Computer & Artificial Intelligence. – 1999. – No. 5. – P. 425–446. 11. Pomerol J.-Ch. About Some Relationships between Knowledge and Context / Pomerol J.-Ch., Brézillon P. // Modeling and Using Context (CONTEXT-01), Dundee, Scotland. Lecture Notes in Computer Science, Springer Verlag. – 2001.– P. 461–464. 12. Steinberg N. Adaptive Context Discovery and Exploitation / Alan N. Steinberg and Christopher L. Bowman // Information Fusion (FUSION), 16th International Conference. – 2013. – P. 2004–2011. 13. Смирнов А. В. Модели контекстно-управляемых систем поддержки принятия решений в динамических структурированных областях / А. В. Смирнов, Т. В. Левашова, М. П. Пашкин // Труды СПИИРАН. – 2009. – No. 9. – С. 116–147. 14. Steinberg N. Adaptive Context Assessment and Context Management / Steinberg N, Christopher L. Bowman, Gary Haith, and Erik Blasch // Information Fusion (FUSION), 17th International Conference. – 2014. – P. 1–8. 15. Krummenacher R., Ontology-Based Context Modeling / Krummenacher and Strang // Ieice Transactions On Information And Systems. – 2007. 16. Strang, T.A Context modeling survey / Strang, Thomas, and Claudia Linnhoff-Popien // Workshop Proceedings. – 2004. 17. What is Cyc? /Cycorp, Inc. Retrieved 11 Oct. 2016 from:http://www.cyc.com/cyc /technology/whatiscyc. 18. Winograd T. Architectures for Context // Human-Computer Interaction. 2001. Vol. 16. P. 2–3. 19. Wiederhold G. An Algebra for Ontology Composition / Wiederhold G // Proc. 1994 Monterey Workshop on Formal Methods. U. S. Naval Postgraduate School, Monterey CA, 1994. – P. 56–61. 20. Bettini, Claudio, et al. A survey of context modelling and reasoning techniques/ Bettini, Claudio// Pervasive and Mobile Computing , 6.2, – 2010. – P. 161–180. 21. Sanchez-Pi N. An Information Fusion Framework for Context-based Accidents Prevention / Nayat Sanchez-Pi, Luis Martí, José Manuel Molina, and Ana Cristina Bicharra Garcia // Information Fusion (FUSION), 17th International Conference. – 2014.

COMMUNICATIONS SYSTEM FOR IT-DEVELOPERS

Natalia Dluhunovych[1], Yurii Forkun[2]
[1, 2] Software Engineering Department, Khmelnytsky National University,
11, Instytutska Str., Khmelnytskiy, 29016, UKRAINE E-mail: [1] dlugunovich@gmail.com,
[2] forkun@ridne.net

Remote working and cooperation in territorially separated project teams during the software development process is the most frequent working pattern in IT companies. Article is dedicated to solve difficulties concerning establishment of the communication system providing compatible work conditions for territorially separated project team members and problems, which occur during the cooperation process between associates of development team.

During the process of problem solving following solutions were proposed. Specified type of communication, remote communication, which might be used in territorially separated project teams was extracted. The classification system of remote communication, which further was used as the basic model of communication system for IT companies' usage.

The authors proposed to identify the following types of tasks for the communication system of cooperation in IT-companies. The first task is to establish communication in the development team (communication). The second task is to provide a support for the common organization of the group work on the processing of general information and information linked to the project (cooperation). The third task is to create conditions for a collaborative software development (collaboration). The fourth task is to promote the formation and development of corporate culture in IT companies (corporate culture).

Each of the tasks was divided to subtasks in order to determine methods of interaction between developers to settle efficient process of software development.

Establishment of communication in team management processes relies on following subtasks: formation of feedback channels; upkeep of permanent and operative connections with team members, conduction of meetings and conferences connected to work and monitoring plans; arrangement of e-learning and strengthening relations between team associates.

Providing a support for the common organization of the group work on the processing of general information and information linked to the project focuses on involving team members and leaders in information proceeding and forming process. This process includes following subtasks: research, information forming and saving, controlling proper access rights, distribution and general information access.

Support of the collaborative software development relies on solving problems concerning collective formation of common strategies of work, shared documentation development and, particularly, cooperation in project development.

Working on previous tasks leads to the next one, creating and developming corporate culture. It should contain rules of communication initiating, standards and rules for corporate communication channels usage, forms of information introducing by different channels, rules of proper responses to information, etc.

For each type of tasks authors proposed range of methods, which ensure the implementation of remote communication during software development process. Moreover, certain elements that make up the system of communication IT company were identified. Conditionally, they might be divided to three groups. First group contains projects, tasks and subtasks. Second group includes offices, teams, and employees, roles in project and access rights. Third group consists of communication types and purposes, channels and methods. Consideration of all these elements allows creating communication system in IT company, engaging territorially separated project team.

Keywords - team projects developing software, remote communications, communications, cooperation, collaboration, corporate culture.

1. Amble S.r Agile Modeling: Effective Practices for eXtreme Programming and the Unified Process [Text] / Scott Ambler. – John Wiley & Sons, Inc., New York, 2002. – 402 p. 2. Herbsleb James D. Geographically Distributed Software Development. Bell Labs, Lucent Technologies [Electronic resource] / James D. Herbsleb Rebecca E. Grinter, and Lawrence Votta Jr. – Mode of access: http://kluedo.ub.uni-kl. de/volltexte/2000/217/pdf/no_series_210.pdf. – Title from the screen. 3. Markovets O. Modeling of citizen claims processing by means of queuing system / O. Markovets, A. Peleschyshyn // International Journal of Computer Science and Business Informatics (IJCSBI). – Vol. 15, No. 1. – India : IJCSBI.ORG, 2015. – P. 36–46. 4. Sutherland Jeff. Distributed Scrum: Agile Project Management with Outsourced Development Teams [Electronic resource] / Sutherland Jeff, Ph. D. Anton Viktorov. – Mode of access: http://jeffsutherland.com/SutherlandDistributedScrumHIC CS2007.pdf. – Title from the screen. 5. Fedushko S. Design of registration and validation algorithm of member's personal data / S. Fedushko, Yu. Syerov // International Journal of Informatics and Communication Technology. – Indonesia: Institute of Advanced Engineering and Science, 2013. – Vol. 2. – No. 2. – P. 93–98.

16. Кульчицький І. М. Дослідження довжини речення та слова у творах Романа Іваничука

THE EXAMINATION OF SENTENCE AND WORD LENGTH IN THE WRITING OF ROMAN IVANYCHUK

Kulchytskyi I. M.
Departmant of Applied Linguistics, Lviv Polytechnic National University, 12, S. Bandery Str., Building 12, 79013, Ukraine, email: bis.kim@gmail.com

Linguistics can boast of its various methods and approaches to the research, but recently there has been a tendency to the use of quantitative methods for the needs of linguistics.

Statistical methods have significantly changed and increased our knowledge of what is language as a system and how it works. They are widely used for quantitative calculations, construction of stochastic linguistic models with the use of previously obtained data, testing of different hypotheses about certain lingual phenomena.

In order to investigate the length of words and sentences all R. Ivanychuk's literary works were divided into two groups – novels (6) and short stories (43). Among writer's works short stories prevail as there are 51 of them. Novels in their turn, which are in a minority, better reflect the writer's style in different time periods. Each investigating word or sentence was given the absolute frequency of values. This data was the basis for variation range. For received variational series the average statistical estimates of the average frequency, namely standard deviation, a measure of the average fluctuation, frequency and standard deviations were found. The calculations were carried out according to standard formulas. To determine the limits of the sentences, the pre— made text structure tags were used. This complex and compound sentences were counted as one, and the words written with a hyphen were counted as one word as well. The length of sentences was measured in word forms. The length of words – in phonemes. The results were compared with those of others in the works of Ukrainian authors.

The article presents the following results which have been obtained in the course of the research:

According to the frequency descending, the length of words was distributed as follows: the most common words with length 4 – 12001 (13.41 %), 5 – 11 199 (12.51 %) and 2 – 11058 (12.35 %) phonemes. Central group consists of the word length equal to 6 – 10 015 (11.19 %) 3 — 9,449 (10.56 %) 1 – 9047 (10,11 %), 7 – 8667 (9.68 %) 8 – 6670 (7.45 %) and 9 – 4895 (5.47 %) phonemes. Words of more than 10 phonemes constitute only 3.8 % (10 – 3072 (3.43 %), 11 – 1861 (2.08 %) 12 – 906 (1.01 %) 13 – 369 (0, 41 %) 14 – 155 (0.17 %) 15 – 71 (0.08 %) 16 – 28 (0.03 %) 17 – 21 (0.02 %) 18 – 12 (0, 01 %), 19 – 5 (0.01 %), 20 – 3 (0.0 %), 22 – 2 (0.0 %), 25 – 1 (0.0 %)) of all words. Words that are longer than 19 phonemes are compounds or ones written with a hyphen. They were counted as one word.

The length of words in the novels by R. Ivanychuk according to its descending were ranked as follows: 4 – 61 041 (12,63 %) 5 – 60,338 (12.48 %) 6 – 54 422 (11.26 %) 2 – 53 196 (11, 01 %) 1 – 49,892 (10.32 %), 3 – 47655 (9,86 %), 7 – 46 810 (9,68 %) 8 – 38,261 (7.92 %), 9 – 29110 (6, 02 %), 10 – 19 841 (4.1 %) 11 – 11 949 (2.47 %), 12 – 6119 (1.27 %) 13 – 2,780 (0.58 %), 14 – 1028 (0, 21 %) 15 – 468 (0.1 %) 16 – 201 (0.04 %) 17 – 114 (0.02 %) 18 – 53 (0.01 %) 19 – 32 (0, 01 %) 20 – 30 (0.01 %) 21 – 11 (0.0 %), 22 – 7 (0.0 %), 24 – 3 (0.0 %), 23 – 2 (0, 0 %) 25 – — 2 (0.0 %), 26 – 1 (0.0 %).

The comparison of the length of words in the works by R. Ivanychuk with the works by other writers has revealed that the average length of words in the short stories by R. Ivanychuk is similar to the works of J. Yanovskyi. As to the Ivanychuk's novels, their average length of words is close to the novels by M. Stelmakh and P. Pancha.

The average length of sentences in novels by Roman Ivanychuk according to the number of words is the following: 4– 6 – 823 (16.54 %) 7– 9 – 774 (15.55 %), 10– 12 – 523 (10.51 %) 13– 15 – 403 (8.1 %) 1– 3 – 389 (7.82 %), 16– 18 – 308 (6.19 %), 19– 21 – 271 (5.45 %) ,> 60 – 239 (4.8 %), 22– 24 – 228 (4.58 %), 25– 27 — 210 (4.22 %), 28– 30 – 148 (2.97 %), 31– 33 – 124 (2.49 %), 34– 36 – 112 (2.25 %), 37– 39 – 83 (1.67 %), 40– 42 – 76 (1.53 %), 46– 48 – 61 (1.23 %), 43– 45 – 58 (1.17 %), 49– 51 – 45 (0.9 %), 52– 54 – 38 (0.76 %), 55– 57 — 35 (0.7 %), 58– 60 – 28 (0.56 %).

The comparison of the data with the similar data of the works of different styles in the Ukrainian language shows that the results correspond with the established data for Ukrainian prose. The only significant difference is a high percentage of long sentences in the works by R. Ivanychuk. This allows concluding that the presence of long sentences is a specific feature of Roman Ivanychuk's style.

Key words: quantative research, sentence length, word length, linguostatistics, style.

1. Arnold Y. V. Osnovi nauchnikh issledovanyi v lynhvystyke: Ucheb. posobye. – M.: Vissh. shk., 1991. – 140 s. 2. Baranov A. N. Vvedenye v prykladnuiu lynhvystyku: ucheb. posobye / A. N. Baranov. – M.: Edy-toryal URSS, 2001. — 360 s. 3. BSE — Bolshaia sovetskaia entsyklopedyia [Elektronnii resurs]. – 1978. – Rezhym dostupa : dic.academic.nJoments.nsflbse/ 4. Buk S. Suchasni metody doslidzhennia movy pysmennyka u slovianoznavstvi / S. Buk // Problemy slovianoznavstva . – 2012. – Vyp. 61. – S. 86–95 5. Buk Solomiia Linhvostatystychnyi opys "Ne

spytavshy brodu" Ivana Franka / Solomiia Buk [Elektronnyi resurs]. – Rezhym dostupu : http://www.lnu.edu.ua/faculty/ Philol/www/visnyk/55_2011/55_2011_Buk.pdf 6. Buniakovskyi V. Ya. O vozmozhnosty vvedenyia opredelytelnikh mer doveryia k rezultatam nekotorikh nauk nabliudatelnikh y preymushchestvenno statystyky / V. Ya. Buniakovskyi // Sovremennyk. – 1847. – T. 3, razdel 2. – S. 48. 7. Vasylev L. M. Metodi sovremennoi lynhvystyky / L. M. Vasylev. – Ufa : Yzd-vo Bash-kyrst. un-ta, 1997. – 182 s. 8. Verkhozyn S. S. O statuse kolychestvennikh metodov v lynhvystyke / S. S. Verkhozyn [Elektronnyi resurs]. – Rezhym dostupu : http://cyberleninka.ru/article/n/o-statuse-kolichestvennyh-metodov-v-lingvistike 9. Hladkyi A. V. Matematycheskye metodi yzuchenyia estestvennikh yazikov / A. V. Hladkyi // Matematycheskaia lohyka, teoryia alhorytmov y teoryia mnozhestv : sb. rabot, 1973. – S. 95–108. 10. Holovyn B. N. Yazik y statystyka / B. N. Holovyn. – M.: Prosveshchenye, 1971. – 189 s. 11. Zvehyntsev V.A. Ocherky po obshchemu yazikoznanyiu / V. A. Zvehkhntsev. - M.: MHU, 1962. – 382 s. 12. Yvanov V. V. Nekotorie problemi sovremennoi lynhvystyky / V. V. Yvanov // Narodi Azyy y Afryky. – 1963. – # 4. 13. Kryhin M. Yu. , Shyrokov V. A. Doslidzhennia informatsiino-statystychnykh vlastyvostei ukrainskoho tekstu / M. Yu. Kryhin, V. A. Shyrokov // Matematycheskye mashyni y systemi. – IPMMS NANU, 2000. – N 1. – S. 120–127. 14. Neliubyn L. L. Tolkovii perevodovedcheskyi slovar : ucheb. posobye / L. L. Neliubyn. – 3-e yzd., pererab. – M: Flynta, 2003. – 320 s. 15. Perebyinis V. S. Kilkisni ta yakisni kharakterystyky systemy fonem suchasnoi ukrainskoi literaturnoi movy / V. S. Perebyinis. – K.: Naukova dumka, 1970. 16. Pyotrovskyi R. H. Ynzhenernaia lynhvystyka y teoryia yazika [Tekst] / R. H. Pyotrovskyi. – L.: Nauka, S 1979. – 111 s. 17. Pyotrovskyi R. H. Matematycheskaia lynhvystyka: ucheb. posobye / R. H. Pyotrovskyi. – M. : Vissh. shk., 1977. – 383 s. 18. Sosiur, Ferdinan de. Kurs zahalnoi lingvistyky / Ferdinan de Sosiur: Per. z fr. A. Korniichuk, K. Tyshchenko. – K: Osnovy, 1998. – 324 s. 19. Statystychni parametry styliv / za red. V. S. Perebyinis. – Kyiv: Naukova dumka, 1967. – 260 s. 20. Frumkyna R. M. Rol statystycheskykh metodov v sovremennikh lynhvystycheskykh yssledovanyiakh // Shaumian, S. K. Matematycheskaia lynhvystyka / otvetstvennii redaktor S. K. Shaumian. – M. : Nauka, 1973. – S. 156–183. 21. Shyrokov V. A. Informatsiina teoriia leksykohrafichnykh system / V. A. Shyrokov. – Kyiv: Dovira, 1998. – 331 s. 22. Fucks W Mathematische Analyse von Sprachele-menten, Sprachstil und Sprachen / W. Fucks. – Koln und Oplade, 1955. – 110 p. 23. Herdan G. Language as Choice and Chance / G. Herdan. – Noordhoff : Groningen, 1956. – 356 p. 24. Ross A. Philological Probability Problems / I. A. Ross. – "Statist. Soc.". – Vol. XII. – 1950. – P. 19–59. 25. Shannon C. Weaver W. The Mathematical Theory of Communication / C. Shannon, W. Weaver. – Univ of Illinois Press, 1949. — 117 p. 26. Zipf G. K. Human Behavior and the Principle of Least Effort / G. K. Zipf. – Addison-Wesley, 1949. —573 p.

17. Висоцька В. А., Наум О. М. Порівняння складності автоматичного опрацювання англійських та українських текстів з урахуванням семантики та синтаксису природних мов

AUTOMATIC PROCESSING COMPLEXITY COMPARISON FOR UKRAINIAN AND ENGLISH TEXTS TAKING INTO ACCOUNT SEMANTICS AND SYNTAX OF NATURAL LANGUAGES

Victoria Vysotska[1], Oleh Naum[2]

[1]Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandera Str., Lviv, 79013, UKRAINE, E-mail: Victoria.A. Vysotska@lpnu.ua

[2]Information Systems and Technologies Department, Drohobych Ivan Franko State Pedagogical University, 24, I. Franko Str., Drohobych, 82100, UKRAINE, E-mail: oleh.naum@gmail.com

Now is important and topical of the English / Ukrainian text content flow dynamics analysis and the information resources processing stages construction for this content.

Is it necessary quality processing of these textual data sets with efficient syntax analysis (keywords determination and text categorization, authorship definition, effective translation, etc.) and semantics (for selection for the text content and the appointment). It is necessary for an operational obtaining of modern data in the world that overloaded as valuable information and information noise excess. Today, effective development and implementation of thematic information resources is not possible without the correct definition of the keywords set for effective positioning and ranking of these resources in search engines. They also must use in a text content array is not chaotic, but on a particular distribution that search engines not to index and not classify it as spam. Developing method of Ukrainian and English textual information processing for the automatic identification of meaningful keywords and content categorization is one of the strategic directions of domestic e-business and IT development.

Automation capabilities providing of information resources processing for meaningful keywords identification and content categorization, the text syntax and semantics analysis helps to increase content, a product or services sales volumes for a constant user, active involvement of potential users and boundaries expanding of the target audience. In particular, these principles and technologies in e-commerce are actively used at sales on-line / off-line systems creation, content analysis and / exchange / store, online store developing, cloud storage / computing. The

common standardized approach absence for Ukrainian textual information processing to automatically identification of meaningful keywords and content categorization, the text syntax and semantics analysis of and so the e-commerce support systems functional design leads to appearance some problems in the typical structure implementation of such systems.

The aim of this article is optimal method determination to automatically processing of Ukrainian textual content set for meaningful keywords identification and automatically categorization of content, syntax and semantics analysis of the text.

This paper presents the generative grammar application in linguistic modelling. Description of syntax sentence modelling is applied to automate the processes of analysis and the synthesis of texts in the natural language. The complexity of automatic processing of English and Ukrainian texts with regard to syntax and semantics of natural languages is analyzed. The comparative linguistic analysis of syntax and semantics of English and Ukrainian languages synthesis buildings natural texts or automatic text processing relevant data sets is analyzed. Comparison of English and Ukrainian languages facilitates the development of linguistic processing algorithms natural texts Germanic and Slavic type.

Content set processing for meaningful keywords identification is built on the principle of keywords finding on the content (terms). This is based on the Zipf law and consists on a words choice from an average frequency of occurrence. Because this is a direct and most simple way, it is proposed to use more complex and appropriate experimental research is conducted. Experimental basis for this research 100 scientific publications Lviv Polytechnic National University Bulletin of the Information systems and networks series (http://science.lp.edu.ua/sisn) is selected, two numbers 783 (http://science.lp.edu.ua/SISN/SISN-2014) and 805 (http://science.lp.edu.ua/sisn/vol-cur-805-2014-2). So another algorithm of keywords set finding taking into account the thematic words basics is developed and in the public domain at http://victana.lviv.ua/index.php/kliuchovi-slova is placed.

Key words: generative grammar, structured scheme sentences, information linguistic system.

1. Bekh P. Anhliys'ka mova. Samovchytel' / P. Bekh, L. Byrkun. - Kyyiv: Lybid'. 1993. – 232 p. 2. Methods based on ontologies for information resources processing / [V. Lytvyn, V. Vysotska, L. Chyrun, D. Dosyn] // LAP Lambert Academic Publishing, Germany, 2016. – 324 c. 3. Berko A. Systemy elektronnoyi kontent-komertsiyi: monohrafiya / A. Berko, V. Vysotska, V. Pasichnyk. – Lviv: LPNU, 2009. – 612 p. 4. Matematychna lingvistyka / [V. Pasichnyk, Y. Shcherbyna, V. Vysotska, T. Shestakevych]. – Lviv : Novyy svit-2000, 2012. – 359 p. 5. Vysotska V. Metody i zasoby opratsyuvannya informatsiynykh resursiv v systemakh elektronnoyi kontent-komertsiyi : avtoreferat dysertatsiyi na zdobuttya naukovoho stupenya kandydata tekhnichnykh nauk : 05.13.06 – informatsiyni tekhnolohiyi / V. Vysotska. – Lviv: LPNU, 2014. – 27 pp. 6. Berko A. Features of information resources processing in electronic content commerce / A. Berko, V. Vysotska, L. Chyrun // ACS journal. – 10(2). – Poland, 2014. – P. 5–19. 7. Vysotska V. Web Content Processing Method for Electronic Business Systems / V. Vysotska, L. Chyrun // International Journal of Computers & Technology. – 12(2). – 2013. – P. 3211–3220. 8. Vysotska V. Modelyuvannya etapiv zhyttyevoho tsyklu komertsiynoho web-kontentu / V. Vysotska, L. Chyrun, L. Chyrun // LPNU Visnyk. – Lviv: LPNU, 2011. – No. 715. – P. 69-87. 9. Vysotska V. Osoblyvosti proektuvannya ta vprovadzhennya system elektronnoyi komertsiyi. / V. Vysotska // LPNU Visnyk. – L'viv 2008. – No. 631. – P. 55-84. 10. Vysotska V. Analysis and evaluation of risks in electronic commerce / V. Vysotska, I. Rishnyak, L. Chyrun // CADSM'07. – Lviv, 2007. – P. 332–333. 11. Vysotska V. Linguistic Analysis of Textual Commercial Content for Information Resources Processing / V. Vysotska // TCSET'2016. –2016. Lviv-Slavske, Ukraine. – P. 709–713. 12. Bisikalo O. Identifying keywords on the basis of content monitoring Ukrainian texts / O. Bisikalo, V. Vysotska // Scientific journal "Electronics. Computer Science. Management.". – 1 (36). – Zaporozhye: ZNTU. – 2016/1. – P. 74–83. 13. Chyrun L. Informational resources processing intellectual systems with textual commercial content linguistic analysis usage constructional means and tools development / L. Chyrun, V. Vysotska, I. Kozak // Econtechmod. – Lublin ; Rzeszow, 2016. – 5(2). – P. 85–94. 14. Kondratyev Y. Kontent-analiz tekstovykh masyviv danykh / Y. Kondratyev, V. Vysotska // IKS-2015. – Ukrayina, Lviv-Slavske. – P. 170-171. 15. Vysotska V. Features of the content-analysis method for text categorization of commercial content in processing online newspaper articles / V. Vysotska, L. Chyrun // ACS journal. – 11(1). – Poland, 2015. – P. 5–19. 16. Vysotska V. Linguistic analysis and modelling semantics of textual content for digest formation / V. Vysotska, L. Chyrun // MEST Journal. – 3(1). – P. 127–148. 17. Vysotska V. Osoblyvosti modelyuvannya syntaksysu rechennya slov'yans'kykh ta hermans'kykh mov za dopomohoyu porodzhuval'nykh kontekstno-vil'nykh hrammatyk / V. Vysotska // LPNU Visnyk. – No. 814. – Lviv: LPNU, 2015. – P. 246-276. 18. Vysotska V. Osoblyvosti rubrykatsiyi tekstovoho komertsiynoho kontentu / V. Vysotska // LPNU Visnyk. – No. 826. – Lviv: LPNU, 2015. – P.359–367. 19. Zastosuvannya kontent-analizu dlya opratsyuvannya tekstovykh masyviv danykh / Y. Kis, V. Vysotska, L. Chyrun, V. Foltovych // LPNU Visnyk. – Lviv: LPNU, 2015. – No. 814. – P. 282–292. 20. Chyrun L. Osoblyvosti metodiv kontent-analizu tekstovykh masyviv danykh web-resursiv v mezhakh rehionu / L. Chyrun, V. Kuchkovskyy, V. Vysotska // LPNU Visnyk. – No. 829. – Lviv: LPNU,, 2015. – P. 296–320. 21. Shestakevych T. Modelyuvannya semantyky rechennya pryrodnoyu movoyu za dopomohoyu

porodzhuval'nykh hrammatyk / T. Shestakevych, V. Vysotska, L. Chyrun, L. Chyrun // LPNU Visnyk, No. 814. – Lviv: LPNU, 2015. – P. 335–352. 22. Bisikalo O. V. Eksperymental'ne doslidzhennya poshuku znachushchykh klyuchovykh sliv ukrayinomovnoho kontentu / O. Bisikalo, V. Vysotska // LPNU Visnyk. – No. 829. – Lviv: LPNU, 2015. – P. 255–272. 23. Berko A. Linhvistychnyy analiz tekstovoho komertsiynoho kontentu / A. Berko, V. Vysotska, L. Chyrun // LPNU Visnyk, No. 814. – Lviv 2015 – P. 203–227. 24. Vysotska V. Generative regular grammars application to modeling the semantics of sentences in natural language / Victoria Vysotska // LPNU Visnyk. – No. 808. – Lviv: LPNU, 2014. – P. 43–56. 25. Vysotska V. Osoblyvosti heneruvannya semantyky rechennya pryrodnoyu movoyu za dapomohoyu porodzhuval'nykh neobmezhenykh ta kontekstno-zalezhnykh hramatyk / V. Vysotska // LPNU Visnyk. – No. 783. – Lviv: LPNU, 2014. – P. 271-292. 26. Vysotska V. Zastosuvannya porodzhuval'nykh hramatyk dlya modelyuvannya syntaksysu rechennya / V. Vysotska, T. Shestakevych, Y. Shcherbyna // LPNU Visnyk. – No. 743. – Lviv: LPNU, 2012. – P. 175–190. 27. Berko A. Zastosuvannya metodu kontent-analizu dlya formuvannya informatsiynykh resursiv v systemakh elektronnoyi kontent-komertsiyi / A.. Berko, V. Vysotska, M. Sorokovs'kyy // LPNU Visnyk. – No. 743. – Lviv: LPNU, 2012. – P. 3–15. 28. Vysotska V. Utvorennya rechen' anhliys'koyu ta nimets'koyu za dopomohoyu porodzhuval'nykh hramatyk / V. Vysotska, T. Shestakevych, Y. Shcherbyna // LPNU Visnyk. – No. 744. – Lviv: LPNU, 2012. – P.142–152. 29. Vysotska V. Intelektual'na systema rozpodilu daydzhestiv mizh pratsivnykamy elektronnykh zasobiv masovoyi informatsiyi / V. Vysotska, O. Okrushko // LPNU Visnyk. – No. 744. – Lviv: LPNU, 2012. – P. 41–53. 30. Shcherbyna Y. Utvorennya ukrayins'kykh diyeprykmetnykiv za dopomohoyu porodzhuval'nykh hramatyk / Y. Shcherbyna, Y. Nikolskyy, V. Vysotska, T. Shestakevych // LPNU Visnyk. – Lviv: LPNU, 2011. – No. 715. – P. 354–369. 31. Proekt stvorennya vidkrytykh slovnykiv dlya perevirky orfohrafiyi ukrayins'koyi movy [Electronic recourse]. – Access mode : https://code.google.com/archive/p/spell-uk/wikis/aspell_uk.wiki. 32. Hunspell/Aspell for Ukrainian [Electronic recourse]. – Access mode: https://sourceforge.net/projects/ispell-uk/.

18. Кушнір Олег Степанович, Альфавіцький Микола Андрійович, Дзіковський Віктор Євгенович, Іваніцький Любомир Богданович, Рихлюк Сергій Вікторович, Сокульський Володимир Іванович. Статистика появи слів у природних і рандомних текстах

STATISTICS OF WORDS OCCURRENCES IN NATURAL AND RANDOM TEXTS

Oleg Kushnir, Mykola Alfavitskyi, Viktor Dzikovskyi [*], Lyubomyr Ivanitskyi, Sergiy Rykhlyuk, Volodymyr Sokulskyi

Department of Optoelectronics and Information Technologies, Ivan Franko National University of Lviv
[*] Natural Sciences College of Ivan Franko National University of Lviv
E-mail: o_kushnir@franko.lviv.ua, alfnickolas@gmail.com, dzup4uk@dzup4uk.org.ua, liubomyr.ivanitskyi@globallogic.com, rykhlyuk@gmail.com, volodya199514@gmail.com

In this paper we study experimentally the statistical distributions describing the occurrence of words in several natural texts, as well as random derivatives of these texts. To study the impact of randomness, a program is designed for both global randomization of texts and local randomization with a running replacement window that operates on the linguistic levels of characters, words and sentences. A program for statistical analysis of the occurrence of different kinds of words provides calculations of the relevant probability density (or mass probability) functions and cumulative probability distributions.

Interpretation of our experimental data is based on the graphical methods of linear approximation. It is shown that the tails of the statistical distributions for different types of texts are either described by exponential function or deviate from this function, thus revealing a presence of a so-called heavy tail. The latter phenomenon means deviation from the purely stochastic nature of ordering of the linguistic elements in text, including abnormally high probability for occurrence of extremely large distances between repeatedly used linguistic elements. A critical analysis of methodological principles used in some researches on the subject known from the literature is done and methodological weaknesses of these studies are pointed to.

We also study the dependence of power-law parameter for the distributions with heavy tails on the total lengths of the texts and show that this dependence is decreasing and linear. The latter phenomenon originates from partial violation of the power Herdan's law known for many complex systems and, in particular, for texts, rather than from variety of text genres as assumed earlier in the literature.

It is proved that a properly averaged probability mass function for the time intervals between words is almost the same for both the natural and random texts. This coincides with the result known from the literature for the English texts described well by the frequency dependences of Zipfian type. The probability density for the word

occurrence is calculated for different text types. It is either the uniform distribution, or is approximately described by the inverse power function. Analytical relationship between the mentioned statistical distribution and the function determining the vocabulary richness of text is revealed. Namely, the uniform distribution corresponds to a linear increase in the number of different words with increasing total number of words in the text, and the power-law distribution to a slightly slower growth, with the power-law exponent less than one.

It is proved that the empirical dependence of the vocabulary richness on the running length of natural and random texts shows some nonlinearity in the double logarithmic scale, which serves as another proof of the deviation from the Herdan's law for these texts. However, the above nonlinearity disappears for the texts in which the density of the word occurrence is constant and independent of the position in the text. Significant deviations of the dynamics of vocabulary growth for the natural and random texts from the dynamics predicted by the simplest power law, along with a break found by us in the vocabulary dependence for one of the natural texts, confirm a need in generalization of the above law.

We discuss in detail relationships of the non-exponential statistical distributions found by us and the phenomena of memory and temporal correlation between the linguistic elements in text, and find links of those statistical parameters with the cross-correlation function. The analytical links found in this work among the quantitative parameters of different statistical distributions and the fact of qualitatively similar behaviors of these distributions for different types of texts indicate that the laws of statistical linguistics can be best explained as a consequence of general properties of random symbolic sequences, rather than evidence of text semantics or some linguistic patterns. The empirical results obtained indicate that, in order to understand correctly the laws of statistical linguistics, it is fundamentally important to take into account the deviations of statistical distributions for the linguistic elements in texts from the exponential function.

Key words: natural texts, random texts, statistical laws of linguistics, vocabulary, fat-tailed distributions.

1. Baek S. K. Zipf's law unzipped / S. K. Baek, S. Bernhardsson, P. Minnhagen // New J. Phys. – 2011. – Vol. 13. – 043004 (21 pp.). 2. Adamic L. Unzipping Zipf's law / L. Adamic // Nature. – 2011. – Vol. 474. – P. 164–165. 3. Kornai A. How many words are there? / A. Kornai // Glottometrics. – 2002. – Vol. 4. – P. 60–85. 4. van Leijenhorst D. C. A formal derivation of Heaps' law / D. C. van Leijenhorst, Th. P. van der Weide // Inf. Sci. – 2005. – Vol. 170. – P. 263–272. 5. Gerlach M. Stochastic model for the vocabulary growth in natural languages / M. Gerlach, E. G. Altmann // Phys. Rev. X. – 2013. – Vol. 3. – 021006 (10 pp.). 6. Bernhardsson S. The meta book and size-dependent properties of written language / S. Bernhardsson, L. E. Correa da Rocha, P. Minnhagen // New J. Phys. – 2009. – Vol. 11. – 203015 (15 pp.). 7. Bernhardsson S. Size-dependent word frequencies and translational invariance of books / S. Bernhardsson, L. E. Correa da Rocha, P. Minnhagen // Physica A. – 2010. – Vol. 389. – P. 330–341. 8. Lü L. Zipf's law leads to Heaps' law: Analyzing their relation in finite-size systems / L. Lü, Z.-K. Zhang, T. Zhou // PLOS ONE. – 2010. – Vol. 5. – e14139 (11 pp.). 9. Yan X.-Y. Comment on 'A scaling law beyond Zipf's law and its relation to Heaps' law' [Electronic resource] / X.-Y. Yan, P. Minnhagen. – 2014. – Access mode: http://arxiv.org/abs/1404.1461. – Title from the screen. 10. Lü L. Deviation of Zipf's and Heaps' laws in human languages with limited dictionary sizes / L. Lü, Z.-K. Zhang, T. Zhou // Sci. Rep. – 2013. – Vol. 3. – 1082 (7 pp.). 11. Font-Clos F. A scaling law beyond Zipf's law and its relation to Heaps' law / F. Font-Clos, G. Boleda, A. Corral // New J. Phys. – 2013. – Vol. 15. – 093033 (16 pp.). 12. Bochkarev V. V. Deviations in the Zipf and Heaps laws in natural languages / V. V. Bochkarev, E. Yu. Lerner, A. V. Shevlyakova // J. Phys.: Conf. Ser. – 2014. – Vol. 490. – 012009 (4 pp.). 13. Font-Clos F. Log-log convexity of type-token growth in Zipf's systems / F. Font-Clos, A. Corral // Phys. Rev. Lett. – 2015. – Vol. 114. – 238701 (5 pp.). 14. Egghe L. Untangling Herdan's law and Heaps' law: Mathematical and informetric arguments / L. Egghe // J. Amer. Soc. Inf. Sci. Technol. – 2007. – Vol. 58. – P. 702–709. 15. Ebeling W. Long-range correlations between letters and sentences in texts / W. Ebeling, A. Neiman // Physica A. – 1995. – Vol. 215. – P. 233–241. 16. Hierarchical structures induce long-range dynamical correlations in written texts / E. Alvarez-Lacalle, B. Dorow, J.-P. Eckmann, E. Moses // Proc. Nat. Acad. Sci. (USA). – 2006. – Vol. 103. – P. 7956–7961. 17. Altmann E. G. Beyond word frequency: Bursts, lulls, and scaling in the temporal distributions of words / E. G. Altmann, J. B. Pierrehumbert, A. E. Motter // PLOS ONE. – 2009. – Vol. 4. – e7678 (7 pp.). 18. Altmann E. G. On the origin of long-range correlations in texts / E. G. Altmann, G. Cristadoro, M. D. Esposti // Proc. Nat. Acad. Sci. (USA). – 2012. – Vol. 109. – P. 11582–11587. 19. Fluctuations of letter and symbol frequencies in Ukrainian and Russian texts / O. S. Kushnir, A. M. Bayovskyi, L. B. Ivanitskyi, S. V. Rykhlyuk // Mater. VII Ukr.-Polish Sci.-Pract. Conf. "Electron. and Inform. Technol.". – Lviv : LNU, 2015. – P. 76–79. 20. Statistical distribution and fluctuations of sentence lengths in Ukrainian, Russian and English corpora / O. S. Kushnir, O. S. Bryk, V. E. Dzikovskyi, L. B. Ivanitskyi, I. M. Katerynchuk, Ya. P. Kis // Bulletin Lviv Nat. Politechnic Univ., Ser. "Inform. Syst. and Networks". – 2016. – No 854. – P. 228–239. 21. Eliazar I. The growth statistics of Zipfian ensembles: Beyond Heaps' law / I. Eliazar // Physica A. – 2011. – Vol. 390. – P. 3189–3203. 22. Simon H. On a class of skew distribution functions / H. Simon // Biometrika. – 1955. – Vol. 42. – P. 425–440.

23. Barabási A.-L. The origin of bursts and heavy tails in human dynamics / A.-L. Barabási // Nature. – 2005. – Vol. 435. – P. 207–211. 24. Chen Y. S. Exponential recurrence distribution in the Simon-Yule model of text / Y. S. Chen // Cybernetics and Systems. – 1988. – Vol. 19. – P. 521–545. 25. Zanette D. H. Dynamics of text generation with realistic Zipf distribution / D. H. Zanette, M. A. Montemurro // J. Quant. Linguist. – 2005. – Vol. 12. – P. 29–40. 26. Keyword detection in natural languages and DNA / M. Ortuño, P. Carpena, P. Bernaola-Galván, E. Muñoz, A. M. Somoza // Europhys. Lett. – 2002. – Vol. 57. – P. 759–764. 27. Herrera J. P. Statistical keyword detection in literary corpora / J. P. Herrera, P. A. Pury // Eur. Phys. J. – 2008. – Vol. 63. – P. 135–146. 28. Level statistics of words: Finding keywords in literary texts and symbolic sequences / P. Carpena, P. Bernaola-Galván, M. Hackenberg, A. V. Coronado, J. L. Oliver // Phys. Rev. E. – 2009. – Vol. 79. – 035102(R) (4 pp.). 29. On the statistics of inter-word distances and the problem of recognition of content words / O. S. Kushnir, A. V. Volosko, L. B. Ivanitskyi, S. V. Rykhlyuk // Electron. Inform. Technol. – 2016. – No 6. – P. 155–164. 30. On the explanation of burstiness in the statistics of linguistic elements: waiting times of letter n-grams / O. S. Kushnir, M. A. Alfavitskyi, V. E. Dzikovskyi, L. B. Ivanitskyi, I. M. Katerynchuk, O. I. Sharha // Mater. VII Ukr.-Polish Sci.-Pract. Conf. "Electron. and Inform. Technol.". – Lviv : LNU, 2015. – P. 84–89. 31. The effect of long-term correlations on the return periods of rare events / A. Bunde, J. F. Eichner, S. Havlin, J. W. Kantelhardt // Physica A. – 2003. – Vol. 330. – P. 1–7. 32. Vajna S. Modelling bursty time series / S. Vajna, B. Tóth, J. Kertész // New J. Phys. – 2013. – Vol. 15. – 103023 (17 pp.). 33. Goh K.-I. Burstiness and memory in complex systems / K.-I. Goh, A.-L. Barabási // Europhys. Lett. – 2008. – Vol. 81. – 48002 (5 pp.). 34. Altmann E. G. Recurrence time analysis, long-term correlations, and extreme events / E. G. Altmann, H. Kantz // Phys. Rev. E. – 2005. – Vol. 71. – 056106 (9 pp.). 35. Statistics of return intervals in long-term correlated records / J. F. Eichner, J. W. Kantelhardt, A. Bunde, S. Havlin // Phys. Rev. E. – 2007. – Vol. 75. – 011128 (9 pp.). 36. Cattuto C. A Yule-Simon process with memory / C. Cattuto, V. Loreto, V. D. P. Servedio // Europhys. Lett. – 2006. – Vol. 76. – P. 208–214. 37. Ferrer i Cancho R. Two regimes in the frequency of words and the origins of complex lexicons: Zipf's law revisited / R. Ferrer i Cancho, R. V. Solé // J. Quant. Linguist. – 2001. – Vol. 8. – P. 165–173. 38. Santhanam M. S. Return interval distribution of extreme events and long-term memory / M. S. Santhanam, H. Kantz // Phys. Rev. E. – 2008. – Vol. 78. – 051113 (9 pp.). 39. Long-term memory: A natural mechanism for the clustering of extreme events and anomalous residual times in climate records / A. Bunde, J. F. Eichner, J. W. Kantelhardt, S. Havlin // Phys. Rev. Lett. – 2005. – Vol. 94. – 048701 (4 pp.). 40. Gerlach M. Scaling laws and fluctuations in the statistics of word frequencies / M. Gerlach, E. G. Altmann // New J. Phys. – 2014. – Vol. 16. – 113010 (19 pp.). 41. Improving statistical keyword detection in short texts: Entropic and clustering approaches / C. Carretero-Campos, P. Bernaola-Galván, P. Ch. Ivanov, P. Carpena // Phys. Rev. E. – 2012. – Vol. 85. – 011139 (6 pp.). 42. Moreno-Sánchez I. Large-scale analysis of Zipf's law in English texts / I. Moreno-Sánchez, F. Font-Clos, A. Corral // PLOS ONE. – 2016. – Vol. 11. – e0147073 (19 pp.). 43. Kushnir O. S. New text-length scaling effects in statistics of natural texts / O. S. Kushnir, L. B. Ivanitskyi, S. V. Rykhlyuk // Mater. VII Ukr.-Polish Sci.-Pract. Conf. "Electron. and Inform. Technol.". – Lviv : LNU, 2015. – P. 80–83. 44. Ferrer i Cancho R. Zipf's law from a communicative phase transition / R. Ferrer i Cancho // Eur. Phys. J.: B. – 2005. – Vol. 47. – P. 449–457. 45. Long-range correlations in nucleotide sequences / C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, H. E. Stanley // Nature. – 1992. – Vol. 356. – P. 168–170.

19. Кунанець Н. Е., Мацюк Г. Р. Тезаурус предметної області "Розумне місто"

THE THESAURUS OF "THE SMART CITY" SUBJECT AREA

Natalija Kunanets
Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, Ukraine
Galuna Matsiuk
Ternopil Ivan Puluj National Technical University, 56, Ruska str., Ternopil, 46000, Ukraine

The sphere of information and communication technologies (ICT) has been actively developing recently, its conceptual apparatus is constantly changing. Certain terms were used only in a narrow circle of specialists twenty years ago but now they are widely used in everyday life. This is caused by the wide spread of new information technologies in various spheres of life, the emergence of new and innovative research areas in science and technology.

Scientific and technological progress, contributing to the rapid development of science and technology, introduces significant changes in the linguistic model of the world, causing the need to update terminological systems, the necessity of systematization and codification of existing terminology. The dictionaries of professional terminological systems are being formed, which are the result of scientific researches and theoretical developments concerning their forming.

The problem of forming terminological dictionaries, including English-Ukrainian, is very important in modern applied linguistics as bilingual terminological dictionaries are of great use in translation and, consequently, the greater use of English professional literature and promotion of the Ukrainian scientists' research results in the international scientific community. In Ukraine there are no bilingual terminological dictionaries presenting the conceptual apparatus of "The Smart City" subject area, that determins the importance of this area of research.

Cognitive and information component in the formulation and solution of problems of "The Smart City" subject area is based on the development of the intellectual analysis of natural language text. An effective dictionary base is the main element of semantic analysis.

Providing the structures with classified terminological systems as electronic thesauri will bring them to the formation of natural language machine fond and development of national terminological data banks. Thesaurus of "The Smart City" subject area will make its contribution to the system.

The structural terminological system of "The Smart City" subject area belongs to a class of lexical and graphical systems that are carriers of lexical and graphical effects. So the lexical and graphical systems based on the V. Shirokov interpretation theory has been presented in the article.

The forming of "The Smart City" subject area thesaurus has been considered. It supplies a wide range of terms and expressions used in the study of intelligent city systems, communication systems, acting as a terminological tool for developers, designers and manufacturers. It can also be used by managers at all levels and in all sectors of " the smart urban communities".

The ontologies are used for structuring knowledge in different subject areas in modern information and technology developments. One of the main tasks of development of the intellectual systems of the subject area ontology is the system of concepts and relationships between them.

An urgent task is to systematize the basic concepts of a complex system "The Smart City" to provide what we used ontological approach in selecting methods and techniques allowing the analysis of the semantic domain information for its effective use, representation and transformation.

The principles of ontological approach to the domain and specific knowledge gaining from the natural language texts, their formal and logical presentation and applied processing on the example of "The Smart City" definition have been analyzed.

It is shown that ontology of "The Smart City" subject area is formed as a categorical and conceptual apparatus of the subject area based on the study of sources about "The Smart City".

The fragment of "The Smart City" subject area ontology has been presented.

Key words: thesaurus, "The Smart City" subject area, ontology, ontohraf.

1. Shcherba L. V. Selected works on linguistics and phonetics / L. V. Shcherba. – Leningrad: Publishing House of Leningrad State University, 1958. – 182 p. 2. Rosenthal D. E., Telenkova M.A. Dictionary of the difficulties of the Russian language: Ok. 30 000 words. – 3rd ed., Ext. – Moscow: Rus. 1970. 3. Vinogradov V. V. The main types of lexical meanings of the word, Selected works. Lexicology and lexicography. – Moscow, 1977. – P. 162–189. 4. Vyskushenko S. A. The problem of developing thematic bilingual terminology dictionary English professional language Livestock / SA Vyskushenko // Scientific Herald Volyn National University. Lesya Ukrainian. Philology. Linguistics: Coll. Science. works. – Lutsk, Volyn National University them. Ukrainian Lesya, 2011. – No. 3. – Part 2. – P. 194–197. 5. Kudashev I. S. Designing of translation dictionaries of special vocabulary: monograph / I. S. Kudashev. – Helsinki, 2007. – 443 p. 6. Predrag V. English-Ukrainian dictionary of educational methodical comments and grammatical tables. – Kyiv: Tower, 2002. – 424 p. 7. Radziyevska C. The principle conclusion of the Anglo-Ukrainian dictionary listing of nanoscience / S. Radziyevska, I. Chekman // Terminology Bulletin: Coll. Science. works. – Kyiv Institute of Ukrainian National Academy of Sciences, 2013. – Vol. 2 (1). – S. 171–179. 8. Seliverstov K. T. Preparation branch dictionaries: Achievements and features styling techniques translation dictionary [electronic resource] / K. T. Seliverstov // Archives of Ukraine: nauk. and practical. magazine. – 2013. – No. 5. – P. 25–62. – Access: http://nbuv.gov.ua/j-pdf/ay_2013_5_4.pdf 9. Shirokov V. A. Elements of lexicography. – Kyiv: Trust, 2005. – 304 p. 10. Shirokov V. A. Information theory lexicographical systems. – Kyiv: Trust, 1998. – 331 p. 11. Shirokov V. A. Phenomenology lexicographical systems. – Kyiv: Science. opinion, 2004. – 327 p. 12. A. Kucherenko Building a modern national civil protection terminology / Kucherenko // Journal of Nat. Univ "Lviv Polytechnic". Series "Problems of Ukrainian terminology." – 2009. – No. 648. – S. 92 – 94. 13. Symonenko L. O. Biological terminology: the formation and operation: manual. guidances. / L.O. Symonenko; Uman. state. ped. University of them. P. Tychyna. – Uman: RVTS "Sofia", 2006. – 103 p. 14. Shunevych B. Modern methods of selection of terms and signing new translation dictionaries terminological / B. I. Shunevych // Journal of Zhytomyr State University. Ivan Franko. – Vol. 38. – P. 90–93. 15. Kulchytsky I. M. Computer-technological aspects of modern lexicographical systems / I. M. Kulchytsky; Eng. mov.-inform. Fund NAS of Ukraine. – Kyiv: Nat. b-ka Ukraine. Vernadsky NAS of Ukraine, 2002. – 59 p. 16. Kulchytsky I. Displays structural relationships registry terminological dictionaries databases / I. M. Kulchytsky // Visn. Nat. Univ "Lviv. Polytechnic". – 2010. – No. 689. – P. 271–280.

17. Kulchytsky I. M. Electronic glossary arhivistyky / I. M. Kulchytsky // Visn. Nat. Univ "Lviv. Polytechnic". – 2009. – No. 653. – P. 148–156. 18. Dmitruk M. V. Formation of Ukrainian veterinary vocabulary: Abstract Dis. for obtaining Sciences. degree candidate. philology. Science / M. V. Dmitruk. – Zaporozhye: the GG, 2001. – 20 p. 19. Serbenska O. A. Ukrainian legal terminology development after the Great October Socialist Revolution: Dis. ... Candidate of Philology: 10.02.01 / Alexander Antonivna Serbenska. – Lviv, 1965. – 465 p. 20. Antonovych E. A. Russian-Ukrainian dictionary directory of engineering graphics, design and architecture: Textbook / EA A., Y. Vasylyshyn, VA Shpilchak. – Lviv, "Mir", 200 – 240 p. 21. 21. Palagin AV, Yakovlev Yu. S. System integration of computer hardware. – Vinnitsa: UNIVERSUM, 2005. – 680 p. 22. Morkovkin V. V. Ideographic dictionaries / V. V. Morkovkin. – Moscow: Izd-vo Mosk. University, 1970. – 71 p. 23. Karaulov Yu. N. Linguistic construction and thesaurus of the literary language / Yu. N. Karaulov. – Moscow: Science, 1981. – 363 p. 24. Bally Sh. General Linguistics and French Language Questions / Sh. Bally. – Moscow: Foreign Literature Publishing House, 1955. – 416 p. 25. Palagin A. V. Ontological methods and means of processing subject knowledge: monograph / A. V. Palagin, S. L. Kryvyi, N. G. Petrenko. – Lugansk: publishing house of the VNU. V. Dal. – 2012. – 324 p. 26. Potapova EV Linguistic ontology "Physics of magnetic phenomena" and its applications: dis. ... cand. Tech. Sciences, – Kiev, 2013. – 180 p. 27. D. V. Lande Elements of Computational Linguistics in legal science. – K .: NDIIP NAPrN Ukraine, 2014. – 168 p. 28. Shirokov V. A. Information theory lexicographical systems. – Kyiv: Trust, 1998. – P. 80. 29. PAS 180 Smart city technology . – Access mode: http://www.bsigroup.com/smart-cities/Smart-Cities-Standards-and-Publication/PAS-180-smart-cities-terminology/ 30. PAS 181 Smart city framework. – Access mode: http://www.bsigroup.com/en-GB/smart-cities/Smart-Cities-Standards-and-Publication/PAS-181-smart-cities-framework/ 31. PAS 182:2014 Smart city concept model. Guide to establishing a model for data interoperability. – Access mode: http://shop.bsigroup.com/Product Detail/?pid. 32. PAS 183 Smart cities – Guide to establishing a decision-making framework for sharing data and information services. – Access mode: www.cibse.org/getmedia/96c2ff0e-73b6-4544-aebd-7debabe211d4/Draft-PAS-183-on-Smart-cities.pdf.aspx 33. PAS 8101: 2014 Smart cities. Guide to the role of the planning and development process. – Access mode:http://shop.bsigroup.com/ProductDetail/?pid=000000000030294642. 34. ISO 37120:2014 Sustainable development of communities -- Indicators for city services and quality of life(2014). – Access mode: https://www.iso.org/obp/ui/#iso:std:62436:en. 35. ISO/TR 37150:2014(en) Smart community infrastructures – Review of existing activities relevant to metrics. – Access mode:: https://www.iso.org/obp/ui/#iso:std:iso:tr:37150:ed-1:v1:en. 36. ISO/TS 37151:2015(en) Smart community infrastructures – Principles and requirements for performance etrics [Електронний ресурс]. – Access mode: https://www.iso.org/obp/ui/#iso:std:iso:ts:37151:ed-1:v1:en . 37. ISO/PRF TR 37152 Smart community infrastructures – Common framework for development and operation – Ad hoc group report [Електронний ресурс]. – Access mode: http://www.iso.org/iso/catalogue_detail.htm?csnumber=66898. 38. A collaborative approach to smart city transformation. – Access mode: smartdubai.ae/whitepaper/Smart_Dubai_ WhitePaper.pdf . 39. Al-Hader, M., & Rodzi, A. (2009). The smart city infrastructure development & monitoring. Theoretical and Empirical Researches in Urban Management, 4(2), 87–94. 40. Barrionuevo J. M., Berrone P., Ricart J. E. Smart Cities, Sustainable Progress / J. M. Barrionuevo, P. Berrone, J. E. Ricart // IESE Insight. – 2012. – Vol. 14. – P. 50–57. 41. Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G. and Portugali, Y. (2012) Smart cities of the future. European Physical Journal Special Topics 214: 481–518. 42. Benevolo C., Dameri R. P. and D'Auria B. Smart Mobility in Smart City. – Access mode: file:///C:/Users/Admin/Downloads/9783319237831-c2%20(1).pdf 43. Borja, J. (2007). Counterpoint: Intelligent cities and innovative cities. Universitat Oberta de Catalunya (UOC) Papers: E-Journal on the Knowledge Society, 5. – Access mode: http://www.uoc.edu/uocpapers/5/dt/eng/mitchell.pdf. 44. Building a Smart + Equitable City. – Access mode: http://www1.nyc.gov/assets/forward/documents/NYC-Smart-Equitable-City-Final.pdf 45. Caragliu A. Smart Cities in Europe / Andrea Caragliu, Chiara Del Bo, and Peter Nijkamp // Journal of Urban Technology. – 2011. – No. 2. – P. 65–82. 46. Caroline Colldahl, Sonya Frey, Joseph E. Kelemen. Smart Cities: Strategic Sustainable Development for an Urban World. - School of Engineering Blekinge Institute of Technology Karlskrona, Sweden, – 2013. 47. Eco-City Movement. – Режим доступу: http://ecocity.ncr.vt.edu/ecocitymovement.html 48. Deakin M. From Intelligent to Smart Cities / Mark Deakin, Husam Al Waer // Intelligent Buildings International. – 2011. – No. – P. 140–152. 49. Getting Smart About Smart Cities. – Access mode: http://sustainablecommunitiesleadershipacademy.org/resource_files/documents/Smart%20Cities%20RG%20(2).pdf 50. Giffinger R. Smart Cities Ranking: An Effective Instrument for the Positioning of Cities? / R. Giffinger // ACE Architecture City and Environment. – 2010. – Vol. 4, No. 12. – P. 7–25. 51. Global Innovators: International Case Studies on Smart Cities Smart Cities Study. – October, 2013. 52. Gontar B., Gontar Z., Pamuła A. Deployment of Smart City Concept in Poland. Selected Aspects. – Access mode: http://eltalpykla.vdu.lt:8080/bitstream/handle/1/963/ISSN2335-8750_2013_N_67.PG_39-51.pdf?sequence=1&isAllowed=y. 53. Green Economy Report. – Access mode: http://web.unep.org/greeneconomy/resources/green-economy-report. 54. Green Technology – What is it? – Access mode: http://www.green-technology.org/what.htm 55. Gruber T. R. A Translation Approach to Portable Ontology Specifications // Knowledge Acquisition. – 1993. – N 5, P. 199 – 220. 56. ICT4 Green by Donato Toppeta.

How ICT can address Sustainability. – Access mode: https://ict4green.wordpress.com/category/smart-city-2/ 57. Jadoul M. (2014) Smart cities are built on smart networks / Marc Jadoul. – Access mode: https://insight.nokia.com/smart-cities-are-built-smart-networks 58. Jahan Selim Human Development Report. Work for Human Development / Selim Jahan. – New York, 2015. – 288 p. 59. Kearmey A. T. Global cities 2016 / Mike Hales, Erik R. Peterson, Andrés Mendoza Peña, Nicole Dessibourg. – NY, 2016. – 13 p. 60. Komninos N. Smart Cities and the Future Internet in Europe / N. Komninos, Schaffers H. // Journal of Knowledge Economy. – 2013. – No. 2. – P. 119–134. 61. Low energy, passive and zero-energy houses. – Access mode: https://www.our-energy.com/low_energy_passive_and_zero_energy_houses.html 62. Making cities smarter Guide for city leaders: Summary of PD 8100 – Access mode: https://www.bsigroup.com/LocalFiles/en-GB/smart-cities/resources/BSI-Making-cities-smarter-Guide-for-city-leaders-UK-EN.pdf 63. Manville Catriona, Cochrane Gavin, Cave Jonathan Mapping Smart Cities in the EU/ Catriona Manville, Gavin Cochrane, Jonathan Cave. – Brussels, 2014. – 196 p. 64. Recommendations to the European Council: Europe and the Global Information Society, M. Bangemann and others, 24–25 June, Korfu, 1994. 65. Report on the WSIS Stocktaking 2016. – ITU 201. – 177 p. 66. Sekhar N. Kondepudi. Smart sustainable cities: An analysis of definitions. – ITU, 2014. 67. Smart cities: Background paper. October 2013. – Access mode: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/246019/bis-13- 1209- smart- cities-background-paper-digital.pdf 68. Smart Cities Maturity Model and Self-Assessment Tool. – Access mode: www.scottishcities.org.uk/site/assets/files/1103/smart_cities_readiness_assessment_-_guidance_note.pdf 69. Smart cities. Ranking of European medium-sized cities. – Access mode: http://www.smart-cities.eu/download/smart_cities_final_report.pdf 70. Smart cities readiness guide. The planning manual for building tomorrow's cities today. – Access mode: www.swenergy.org/Data/Sites/1/media/documents/programs/government/SmartCitiesCouncil- ReadinessGuide-11.18.13a.pdf 71. Smart Cities: Technological Challenges and Issues Challenges and Issues. – Access mode: http://www.cse.wustl.edu/~jain/talks/ftp/smrtcit.pdf 72. Smart Cities. Transforming the 21st century city via the creative use of technology. – Access mode: http://publications.arup.com/publications/s/smart_cities 73. Smart city opportunities for London. – Access mode: https://www.london.gov.uk/sites/default/files/arup-gla_smart_city_opportunities_for_london.pd 74. "Smart+ Connected City Services", Cisco Consulting Services, 2011, http:// www.cisco.com/web/about/ac79/docs/ps/Busan-Green-u-City_IBSG.pdf 75. Smart Economy: Economía Inteligente. – Access mode: https://web.ua.es/en/smart/smart-economy-economia-inteligente.html 76. Smarter Cities. New cognitive approaches to long-standing challenges. – Access mode: http://www.ibm.com/smarterplanet/ru/ru/smarter_cities/overview/ 77. Smart Sensor Networks: Technologies and Applications for Green Growth. – Access mode: http://dx.doi.org/10.1787/5kml6x0m5vkh-en 78. The Internet of Things: A Proto-technate by Enrique Lescure. – Access mode: https://eoshorizon.wordpress.com/tag/smart-cities/ 79.The Role of Standards in Smart Cities. – Access mode: https://www.bsigroup.com/LocalFiles/en-GB/smart-cities/resources/The-Role-of-Standards-in-Smart-Cities-Issue-2-August-2014.pdf 80.The Smart Cities Mission – Access mode: www.slideshare.net/neersee/smart-city-presentation-50456375?next_slideshow=1 81.The Smart City – City of Stockholm. – Access mode: http://international.stockholm.se/city-development/the-smart-city/ 82.The Water Footprint Assessment Manual. Setting the Global Standard. – London, 2011. – 228 p. 83. Towards nearly zero-energy buildings. – Access mode: https://ec.europa.eu/energy/sites/ener/files/documents/nzeb_executive_summary.pdf 84.Trends in Smart City Development. – Access mode: www.nlc.org/Documents/Find%20City% 20Solutions/Research% 20Innovation/Smart%20Cities/Trends%20in%20Smart%20City%20Development.pdf 85. United States Green Buildings Council. – Access mode: http://www.usgbc.org/articles/what-green-building 86.What is Green Infrastructure? – Access mode: https://www.epa.gov/green-infrastructure/what-green-infrastructure

20. Білятинська І. М. Інтегровані лексикографічні системи та методи їх побудови

LEXICOGRAPHICAL INTEGRATED SYSTEMS AND METHODS OF CONSTRUCTING

Iryna Bilyatynska
Computer Systems department of the Secretariat of the Verkhovna Rada of Ukraine, 6/8, Bankova Str., Kyiv, 01008, UKRAINE, E-mail: ira121191@ukr.net

Linguistic research is always topical, because it is a means of communication, ways of knowing the world and reflecting the results of intellectual activity. One of the traditional means of storage, display and research results of the study of linguistic phenomena is the dictionary. Usually, traditional paper dictionary describes some aspect of linguistic substance, but an integral description of linguistic phenomena becomes possible by lexicographical systems. Lexicographical system is the achievement of the computer lexicography, which is the young science of compiling dictionaries by using modern information technology. The software products allow combining the description of a finite number of linguistic phenomena by enabling iner-integration.

Lexicographical system describes the general type of formal structures. Different information systems, databases, dictionaries in traditional or electronic form, encyclopedias, etc. are examples of lexicographical systems in science and technology. Lexicographical system allow considering the specific linguistic phenomena, reflecting the most extensive form and content of units of language in the process of designing information-linguistic objects.

During the construction of lexicographical systems it is commonly accepted to use architecture ANSI / X3 / SPARK. According to the architecture of ANSI / X3 / SPARK lexicographical system consists of three levels. There are conceptual, internal and external. Integration can be done at all three levels simultaneously, or process selective levels depending on the characteristics of components and integration purposes. The result of the integration of a finite number of lexicographical systems is an integrated system that is an amalgamation of several software tools working together. The software tools combination provides its user with panoramic information while each separate lexicographical source exposes a limited message. Lexicographical system, involved in integrating, is the objects of lexicographical environment.

We believe that the integration of lexicographical systems can be made as a result of the following steps:

1. Build a conceptual model of lexicographical systems (components of integration) that has become the objects of integration.

2. Analyze these conceptual models.

3. Identify the phenomena, which are described simultaneously in two or more components of integration. Determine how these descriptions overlap, and which of them are supplemented. Consequently, determine which new features of the described effects can be obtained as a result of the integration.

4. Build a new conceptual model that is based on the findings. This conceptual model is the integration of conceptual models built in step 1.

5. Evaluate the findings of the conceptual model and determine the level at which it is advisable to implement further integration of lexicographical system. In cases when none of these options is applicable, it is reasonable to implement a new lexicographic system with its own internal and external levels of representation.

6. Make a choice of the type of integration software implementation lexicographical system in accordance with the analysis carried out in previous steps.

The article also focuses on the analysis of the main challenges that may accompany the integration process. In particular, we concentrate our attention on the features of unification of lexicographical systems, based on a variety of lexicographical effects and data structures with different architecture using different technologies and approaches to designing. At some point, integration of lexicographical systems can be based on the integration of their software implementations; hence in the article we examine the integration of complex information systems. We conducted an analysis of the prevailing modern approaches to the integration of information systems, including classification according to the data structure, classification according to the level of implementation and modern web integration.

Conducting the research, we have concluded that during the process of integration it is necessary to consider all levels of abstracting the data of lexicographical systems involved in integration: to analyze the conceptual models, ways of data presenting on the inner level, architectures implementing external level in order to determine the level at which the integration can be successfully implement. It is often the case that the process of integration should cover two or all three levels. Sometimes, this analysis demonstrates the impossibility of integrating the output information systems which leads to the task modification with the intention to develop the entirely new object.

Key words: lexicographical system integration, information systems, types of integration, integration of electronic dictionaries, lexicographical integrated system.

1. Steel T. Interim Report: ANSI/X3/SPARC Study Group on Data Base Management Systems 75-02-08 / T. Steel // ACM SIGMOD Record / T. Steel. – New York: ACM, 1975. – Т. 7. – No. 2. – С. 1–140. 2. Рабулець О. Г. Інтегровані лексикографічні системи : автореф. дис... канд. техн. наук: 05.13.06 / О. Г. Рабулець; НАН України. Нац. б-ка України ім. В. І. Вернадського. – К., 2002. – 18 с. – укр. 3. Словники України. Інтегрована лексикограф. система (версія 4.1) [Електронний ресурс] : словозміна, транскрипція, фразеологія, синонімія, антонімія / [В. А. Широков та ін.] ; Нац. акад. наук України, Укр. мовно-інформ. фонд. – Електрон. дан. – К. : Довіра, 2010. – 1 електрон. оптич. диск. – Систем. требования: Операційна система – MICROSOFT WINDOWS 7/ MICROSOFT WINDOWS VISTA / MICROSOFT WINDOWS SERVER 2008 / MICROSOFT WINDOWS SERVER 2003 / MICROSOFT WINDOWS XP SP3 ; програмне забезпечення MICROSOFT ,NET 4.1; процесор INTEL PENTIUM/CELERON/ XEON, AMD K6/ATXLON / DURON або сумісний з ними процесор, тактова частота якого складає 1 ГГЦ і вище; оперативна пам'ять – не менше 512 МБ ; вільне місце на жорсткому диску – не менше 100 МБ; дисковод для CD/DVD-дисків – для встановлення програми. – Назва з етикетки диску. 4. Широков В. А. Комп'ютерна лексикографія / В. А. Широков. – К.: Наукова думка, 2011. – 351 с.

21. Шаховська Н. Б., Гірак Х. Ю. Шкалювання емоційно забарвлених слів для використання у методах класифікації тональності

RANGING OF EMOTIVE WORDS FOR THE USE IN THE METHODS OF TONE CLASSIFICATION

Nataliya Shakhovska, Khrystyna Hirak

Information Systems and Networks Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, Ukraine

The processing of large volumes of text data left by users in social networks in order to get the author's thoughts takes a lot of time and effort. Especially if you need to analyze thousands and millions of user profiles. This problem has led to the development of methods and tools for analyzing the semantic color of text posts in social networks. Such methods will allow to analyze the author's opinion, classify them according to certain criteria in order to predict the future behavior and needs. On the whole, it is possible to divide the existing approaches of tonality classification into the following categories: approaches based on the rules; approaches based on dictionaries; template based approaches; frequency methods; machine learning.

Affective lexicons-based approaches use a list of words with the meaning of the tonality (valency) of each word. The most common way of providing a document (s) in the tasks of computer linguistics and search is in the form of a set of words; N-gram data sets; vectors. The article analyzes the methods for determining the coefficients of words importance (the Fisher's method, the pair comparison, the Purto hypothesis, the Bayesian classifier, the linear classifiers, the method of Rocco, the method of k-nearest neighbors, the method of reference vectors, the method of modeling of maximum entropy, Word2Vec), and the approach of averaging the results of methods has been put forward.

The authors offered the following algorithm for determining the tonality of posts:

1. Processing the text data in advance. At this stage, all html tags (if we work with hypertext), punctuation marks, symbols are deleted. This operation is realized on the recited external library of the programming language php "Beautiful Soup". Then in the text there are so-called "stop words" – these are frequent words in the language, which basically do not carry any meaning load (for example, in English, such words as "the, at, about, ..."). Stop words are deleted using the NLTK package. After processing the data a set of words (array of words) is obtained. At this stage, it is possible to further improve the structure of words, but this will be done in further research. Meaning that three posts (e.g. comments on the news) are given in the social network with the following pre-processed vectors of words:

• [performance, roof, automobile]
• [autopilot, capacity, battery]
• [seats, decor, speed]

2. Presentation in the form of a vector. For this purpose, the methods of words ranking were examined and the basis-dictionary was compiled, taking into account expert assessments for further research. So, on this stage it is needed to submit text in the form of a vector of numbers (selectively you can use the Dahl or Zheleznyak dictionaries), that is, to replace the words from the text with the index from a pre-designed dictionary. For greater clarity, it is offered to combine all the words from the list in point 1 and sort them by the method of the Fisher's scale:

[automobile, battery, speed, autopilot, performance, capacity, decor, seats, roof].

By replacing the previous vectors with the index of words in the dictionary, the following results are obtained:

• [1,0,0,0,1,0,0,0,1]
• [0,1,0,1,0,1,0,0,0]
• [0,0,1,0,0,0,1,1,0]

The resulting vectors are called "word vectors" or "feature vector". Thus, vectors for each post in the social network are obtained.

3. Retrieval of the integral value. With the calculated vectors of words, and, accordingly, calculated weight coefficients of the importance of each word, one can begin to determine the integral estimate of each post as well as the classification of texts, etc. These studies are planned to be conducted in further scientific works.

Consequently, the software module "Semant-1" has been developed in order to implement the above-described algorithm; this module can be effectively used for tasks of semantic ranking of words.

Key words: emotional coloring, colored words, tonal dictionaries, valence of a word, frequency analysis, the scale of Fishburne, pair comparison, hypothesis of Purto.

1. Pang B. Opinion Mining and Sentiment Analysis / B. Pang, L. Lee // Foundations and Trends in Information Retrieval: Vol. 2. No. 1–2, 2008. 2. Danyliuk I. H. Technolohiya avtomatychnogo vyznachennya tematyky tekstu [Tekst] / I. H. Danyliuk // Lingvistychni studiyi: Zb. nauk. Prac. Vypusk 17 / Ukl. Anatoliy Zahnitko (nauk. red.) ta in. –

Donetsk: DonNU, 2008. – S. 290–293 3. Medykovsky M. O. Doslidzhennya efectyvnosti vyznachennya vahovych koeficientiv vazhlyvosti / Medykovsky M. O., Shunevych O. B. // Visnyk Khmelnyckogo universytetu. – 2011. – No. 5. – S. 176–182. 4. Lytvyn V. V. Metod kvazireferuvannya tekstovych dokumentiv na osnovi ontolohiji predmetnoji oblasti / V. V. Lytvyn, T. I. Cherna, V. M. Kovalevych / Vidbir i obrobka informaciyi, Vyp. No. 41 (117). – 2014. – S. 100–108 5. Chomiv B. A. Komparatyvnyj analiz matematychnych modeley na zasovi ociniuvannya opiniyi v tekstovych danych internet resursiv / B. A. Chomiv, S. A. Lupenko, A. S. Sverstiuk // Visnyk Chmelnyckogo nacionalnogo universytetu. – 2011. No. 6. – S. 7–16. 6. Chalaya L. Ye. Miery vazhnosti konceptov v cemanticheskoy seti ontolohicheskoy bazy znaniy / L. Ye. Chalaya, Yu.Yu. Sheviakova, A. Yu. Shafranenko // Materialy druhoyi mizhnar. nauk.-techn, konf. "Suchasni napriamy rozvytku informacijno-komunikacijnych technologiy ta zasobiv upravlinnya". – K.: KDAVT, 2011 – S. 51. 7. Shahovska N. B., Noha R. Yu. Analitychnyj ohlyad metodiv ta zasobiv opraciuvannya tekstovoji informaciji Informacijni systemy ta merezhi No. 715. – L. : Vydavnyctvo Natsionalnoho universytetu "Lvivs'ka Politechnika", 2011. – S. 215–223. H. Wu and R. Luk and K. Wong and K. Kwok. "Interpreting TF-IDF term weights as making relevance decisions" // ACM Transactions on Information Systems, 26 (3). 2008. 12. Katrin ERK. Vector space models of word meaning and phrase meaning: A survey. Language and Linguistics Compass, 2012, 6.10: 635–653. 8. Online resource TF-IDF [Access mode]. [https://ru.wikipedia.org/wiki/TF-IDF]. 9. Online resource Okapi [Access mode]. [https://ru.wikipedia.org/wiki/Okapi_BM25]. 10. Online resource: [Access mode]. [https://habrahabr.ru/post/149605/]. 11. Online resource: [Access mode]. [http://nlpx.net/archives/179].

22. Бабич С. В., Турбал Ю. В. Методи формування матриць розкладів на основі використання модифікованих перманент

METHODS OF FORMING MATRICES OF TIMETABLES BASED ON THE USE OF MODIFIED PERMANENT

S. Babych, Y. Turbal
National University of Water and Environmental Engineering

The specific problem of forming of the timetable of lessons of the institution is a lot of criteria and dependence from different computing resources. This technique is similar to division of tasks of constraint satisfaction. In the basis of the construction of timetable is the scheduling theory, which is widely used in the organization of work of enterprises and in forming of school timetable. Those problems also intersect with the section of dynamic programming in control theory and the theory of computing.

The paper proposes a new configuration approach to analysis of timetables matrix, based on using of combinatorial properties of algebraic structures, including permanent.

All well-known criteria which must meet a timetable: students should not have break between lessons, teachers should not have break between lessons, the number of working days of the teacher should be minimal, the wishes of teachers should be considered the most important etc. Obviously, the optimal timetable is the best choice for each day of the training week. Therefore confine ourselves to one training day.

If for each teacher puts in correspondence some positive integer (number, weight, etc.), the timetable will be a matrix of dimension $3 \times n$, where 3 – number of lessons of a teacher a day, n – the number of groups (subgroups), for which consists the timetable. We will call that matrix "matrix of timetable" and mark R. Obviously that during the analysis of matrix of timetable it is arisen a number of tasks related to the requirements of the timetable. The task of this paper is to develop a methodology admissibility verification matrix timetable and algorithmic aspects of their construction. During this we use the concept of ternary, binary and unary configurations.

Ternary configuration of timetable formed by the element of matrix of timetable d, call the set of matrix elements *R*, that$a_{ij} = a_{kl} = a_{mp} = d$, *i, k, m* ϵ {1, 2, 3}, $1 \le j$, $p \le n$. Similarly, we can introduce the concept of binary and unary configurations (binary configuration cannot be a ternary element). Natural for a ternary configuration is a condition where $i \ne k \ne m$ (The teacher cannot carry out a few lessons at the same time). The same condition holds for binary configuration.

Thus, the matrix of timetable is a set of ternary, binary, unary configurations and zeros (zero means that no lessons). It is obvious that a set of configurations must be such that matrix of timetable could be formed at all (in one group should be no more than 3 lessons per day). The matrix of timetable call admissible if the line does not have two identical elements. To construct the algorithms of forming matrix timetables used modified permanent incidence matrix that represents the sum of all possible products of matrix elements, each of which contains one element from each row and different columns and item streaming column (the column corresponding to the streaming element) cannot be the product together with elements of other lines that correspond to the same stream.

Note that in the absence of flow elements modified permanent is a normal permanent.

Allowable matrix is formed gradually, forming first, second, third string (etc.) by permutations of elements of columns. To form the first line of permissible matrix of timetable, it is necessary that a configuration of sets of columns $\{R_1, R_2, ..., R_n\}$ contains a system of distinct representatives. But a system of distinct representatives exists only when the permanent of incidence matrix of configuration is different from zero. So we examine a permanent of matrix A. If it is different from zero, than a system of distinct representatives exists: $a_{11}, a_{12}, ..., a_{1n}$. This system of distinct representatives forms a first line of acceptable matrix of timetable. In order to form a second line of acceptable matrix it is necessary that system of distinct representatives of configuration sets exist $\left\{\frac{R_1}{a_{11}}, \frac{R_2}{a_{12}}, ... \frac{R_n}{a_{1n}}\right\}$, where: $a_{11}, a_{12}, ..., a_{1n}$.- is the first line of acceptable matrix of timetable. We continue the relevant arguments. Note that the initial process of forming configurations can be either automatic, for reasons of minimal number of working days of a teacher, or be the result of an agreement with a teacher (for example, for some teachers it is difficult to carry out three lessons in one day, then you can create the corresponding binary or even unary configuration if everyone will be satisfied).This approach solves the problem of minimizing of breaks between lessons of teachers at the early stage of solving the problem: matrix of timetable is a set of optimized (without breaks between lessons) configurations.

Key words: scheduling, task scheduling, configuration approach, permanent matrix.

1. Dechter R. Constraint Processing / R. Dechter. – Kaufmann, 2003. – 481 p. 2. Беллман Р. Динамическое программирование / Р. Беллман. – Издательство иностранной литературы, 1960. – 400 с. 3. Давыдов С. В. Система автоматического построения расписания учебных зянятий. – М., 1999. – 320 с. 4. Кузьмичев А. Б. О подходе к автоматизации составления расписания в учебном заведении / А. Б. Кузьмичев // Техника машиностроения. – 2014. – No. 3. – С. 23–26. 5. Попов Г. А. Формализация задачи составления учебного расписания в высшем учебном заведении. – Астрахань, Вестник АГТУ. – 2006. – No. 1(30). 6. Конвей Р. В. Теория расписаний / Р. В. Конвей, В. Л. Максвелл, Л. В. Миллер. – М.: Наука, 1975. – 389 с. 7. Танаев В. С. Теория расписаний. Одностадийные системы. / В. С. Танаев, В. С. Гордон, Я. М. Шафранский. – М.: Наука, 1984. – 345 с.

23. Литвин В. В., Угрин Д. І., Іллюк О. Д., Білоус С. В., Рибчак З. Л. Система оптимізації маршрутів туризму на основі модифікації генетичного та мурашиного алгоритмів

THE SYSTEM OF OPTIMIZATION TOURISM ROUTES BASED ON MODIFICATION OF GENERIC AND ANT ALGORITHMS

V. Lytvyn[1], D. Ugryn[2], O. Illyuk[2], S. Bilous[3], Z. Rybchak[1]

[1] Lviv Polytechnic National University, Information Systems and Networks Department, [2]Chernivtsi faculty NTU "Kharkiv Polytechnic Institute", Department of Information Systems [3] Company Ceit.dev, E-mail: [1]Vasyl.V. Lytvyn@lpnu.ua, zoriana.l.rybchak@lpnu.ua , [2]ugrind@mail.ru, [3]bilous@ceit.com

Computational intelligence algorithms have been increasingly used to solve combinatorial optimization problems recently. These algorithms have several advantages: practical application, flexibility in settings and allow very effective results which help to find the best solution in a short period of time. These include ant algorithms and genetic algorithms.

There is a need to use algorithms that not only carry out the main task of searching the shortest route, but also execute related needs in cases of problems in transportation route optimization tasks used in the decision search for the optimal route taking into account the cost of resources. It is proposed to use multiple algorithms, such as such as genetic algorithm and ant algorithm, to execute such a task.

So far, an ant algorithm has been used reference to perform transport tasks. This algorithm is classified as swarm algorithms. The article suggests using ant algorithms to solve transport problems, as for finding the shortest path ants leave pheromone traces in the search for food sources. As a result of these actions on the way, which has the largest concentration of pheromones, more and more ants will move considering this path as the shortest. This approach allows you to find the best solution in most transportation problems. The purpose of this article is to provide analysis of current context representation and processing methods and delineate unresolved problems and promising areas for research.

However, one ant algorithm will not be enough to perform transport tasks putting taking into account the use of resources by calculation of costs. There was a need to use a genetic algorithm, particularly its functions of crossing two decisions. Genetic algorithm allows to cross several solutions of ant algorithm creating a descendant that contains a single unified solution with set properties, such as fuel consumption, time consuming, distance, etc.

In practice, the solution of transport problems, taking into account the cost of resources is used. A combination of two or more algorithms or a hybrid algorithm creation will fulfill the objectives of the problem. However, the use of such a decision requires a technical background, according to current trends in the world.

Modern knowledge about mobile technology can build various models of transportation problems solution in many areas, including the tourism sector. Today the need for route optimization travel agencies is important. You can find a basic explanation of this statement. Technology development is growing day by day. New roads have been built, new tourist resorts have been constructed, interaction channels of tourist centers have been created, so transport network as a whole has been increased. As a result, optimization of new routes is an urgent problem. Therefore, an intensive study of the scope and development of new solutions is very promising direction.

Modeling and design of limitation in time and expenses which are based on the modeling of collective intelligence are used in implementing these methods. These include: the method of ant colonies (Ant Colony Optimization, ACO), a method of crossing two solutions genetic algorithm (Crossbreeding solutions genetic algorithm, CSGA), modifications operators (modification operators, MO) and other methods. These methods are effectively used to solve different problems: ACO is used for solving the shortest route search, CSGA used for solving problems of transportation, MO used for clustering and data objects. All in all, the above listed methods are used to create a hybrid algorithm.

The usage of programming methods combining genetic algorithm and ant algorithm allows to carry out the tasks of the article. Analyzing the behavior of ant colonies, such as search for the shortest route by leaving mating pheromones and crossing features of two solutions of genetic algorithm methods and algorithms have been developed for such operations: search for the optimal route, costing resources, search distance, time, route, storing executed routes. The paper presents a description of created system for mobile phones with operating system IOS, which performs all operations listed above. Mobile application is tested by "first test" method. It is done in several stages. The first phase included the collection of information and building a hybrid model with the modified crossing operators and initialization. Mathematical models of algorithms were built on the second stage. They take into account the features of the hybrid algorithm. Algorithms models were created also in this stage. Modifiers, which served as costing resources, namely consumption, path length, timing, number of tourists were built in the third stage. Tourists service becomes an important issue in the construction of the algorithm, as meeting all the needs of tourists (not compensating this process by finding the shortest path and costing of all resources) was the main goal of the study. A functions combination of genetic algorithm taking into account the modified operators were used to solve this problem.

The article analyzes modern methods of route optimization which are used to solve transportation problems with the help of which solutions to transport tourists designed between the settlements taking into account efficient use of resources. These methods include genetic and ant algorithms by which the optimal route search and placement of collection points were performed.

Based on created operators modifications of initialization and crossing, a system that solves the transportation problem in the tourism sector with regards to placement of collection points and transport tourists in vehicles was built.

In the study of the effectiveness of the algorithm as a numerical experiment and testing of the effectiveness of mobile application by "first test" method was found that the system builds a route from one settlement to another in which an error and distance is the smallest.

Key words: ant colony genetic algorithm, hiking trails, a mobile application testing.

1. Kazharov A. A., Kurejchy`k V. M. Murashy`ni algory`tmy` dlya vy`rishennya transportny`x zadach / Rosijs`koyi akademiyi nauk. Teoriya i sy`stemy` upravlinnya. – 2010. – S. 32–45. 2. Yemel`yanova T. S. Rishennya etalonny`x transportny`x zadach z klasterny`m roztashuvannyam kliyentiv iz vy`kory`stannyam genety`chny`x algory`tmiv // Nechitki sy`stemy` i obchy`slennya (NSMV-2008): naukova konf. z mizhnar. uchast. – 2008. – S. 195–199. 3. Gladkov L. A. Genety`chni algory`tmy`: Navchal`ny`j posibny`k // L. A. Gladkov, V. V. Kurejchy`k, V. M. Kurejchy`k. M .: Fizmat, 2006. – S. 320. 4. Goryachev Yu.V. Genety`chni algory`tmy` bagatokry`terial`noyi konfliktnoyi opty`mizaciyi. / M.: 2001. – S. 102. 5. Kurejchy`k V. V., Zaruba D. V., Zaporozhecz` D. Yu. Zastosuvannya genety`chnogo algory`tmu rozv'yazannya zadachi try`vy`mirnoyi upakovky` / Novy`ny` PFU. Texnichni nauky`. – 2012. – S. 8–14. 6. Bova V. V., Kurejchy`k V. V. Integrovana pidsy`stema gibry`dnogo i kombinovanogo poshuku v zadachax proektuvannya ta upravlinnya // PFU. Texnichni nauky`. – 2010. – S. 37–42. 7. Kurejchy`k, V. M. Poshukova adaptaciya: teoriya i prakty`ka / V. M. Kurejchy`k, B. K. Lebedyev, O. K. Lebedyev. – M.: Fizmat, 2006. – S. 272. 8. Yermeyev A. V. Rozrobka i analiz genety`chnogo ta gibry`dnogo algory`tmu dlya rishennya zadach dy`skretnoyi opty`mizaciyi / Avtoref. dy`s. kand. tex. nauk. Omsk , 2000. – S. 22. 9. Gvozdyev S. E. Matematy`chne programuvannya / Novosy`birs`k: NGASU – 2001. – S. 96. 10. Bobary`kin V. A. Matematy`chni metody` rishennya avtotransportny`x zadach / SZPI – 1986. – S. 83. 11. Alyeksyeyev A. O. Transportna zadacha po kry`teriyu chasu pry` obmezhenij kil`kosti transportny`x resursiv / Matematy`chni metody` opty`mizaciyi i upravlinnya v skladny`x sy`stemax. KGU – 1984. – S. 60–65. 12. Verxovs`ky`j B. S. Zadachi linijnogo programuvannya ty`pu transportny`x / DAN SSSR, 1963, t. 151. – #3. – S. 515–518.

24. Шаховська Н. Б., Висоцька В. А., Скотар О. О. Розроблення архітектури інтелектуальної системи на основі інноваційних методів навчання студентів

DEVELOPMENT OF THE INTELLECTUAL SYSTEM ARCHITECTURE BASED ON INNOVATIVE METHODS FOR TEACHING STUDENTS

Natalya Shakhovska[1], Victoria Vysotska[2], Oleh Skotar[3]
[1,2]Information Systems and Networks Department, [3]Software Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, 79013, UKRAINE, E-mail: [1]natalya233@gmail.com, [2]Victoria.A. Vysotska@lpnu.ua, [3]oleh.skotar@gmail.com

Lately as in Ukraine and so in other European countries there is an interest student's loss tendency in scientific research and education in general. The main reason for such development of events is the motivation lack. Students are not able to concentrate on learning. The reason is the rapid development of information technology. Because of this, people do not have neither the time nor desire to read books, to take notes of lectures and to perform traditional laboratory work. In particular this applies students in areas related to information technology. IT sector is innovative. And it is natural that traditional teaching methods (such as lectures and books) are not always effective. These methods are weak, primarily because the student is passive in the learning process. The person is the only one observer. And so he can lose attention easily to the material. Student must constantly maintain a high level of concentration on a lecture or a book that quickly tires. Therefore, the traditional system needs to be updated for encouraging such students to study and research, making it more interesting and urgent. In this complex issue comes to the aid gamification. These techniques tend to involve the natural human instincts: competition, achievement, status, expression, and problem solving. This article describes a system in which students can learn the educational program by creating so-called bots for games. A game refers to the software, which realizes environment for bots competition among themselves in different conditions. So the research object is student's education system. This paper analyzes the current system, its advantages, disadvantages and alternatives. Also, innovative methods of teaching and evaluation of students are considered that are the subject of this research. Writing bots for games as learning provides an opportunity to make planned work on the site online, and also the possibility competes in the implementation effectiveness with the other students. All this greatly affects the motivation of students and involves them in active learning required subject area, and encourages to a scientific research. So, virtual computer laboratory will interest students in learning, will provide easy access to materials, provide opportunities for self-realization, and will make students evaluation much simple and objective. The system development (for an existing system supplement in the universities) will give constant and quick access to necessary materials for users (for example, lecture notes, articles, video lectures). Also, it will simplify, speed up and make more objective assessment of knowledge, improve the situation, interest students in learning, and allow them to evaluate properly. Advantages of system for analyzing, using and developing artificial intelligence (SAUDI) are the following:

- gamification will encourage students to study and research;
- system success tracking will allow objectively evaluate students without the human factor intervention;
- easy access to learning materials and various methods of filing them, which allow to learn the material for students more convenient;
- • gamification provides the opportunity in self-realization for students and good abilities.

SAUDAI disadvantages are the following:

- this computer system (technical factors are influenced critically in its work: a problem on the server or internal error);
- the system can not completely replace traditional approach, because it will not officially be replaced in the near future;
- sometimes human factor lack does not allow properly assess student.

The analysis showed that SAUDAI must be used in combination with the traditional system. This is not to replace teachers, but it is a useful tool in their hands. Using the online system will allow students to easily access to all the necessary materials and discuss problem questions with other students and teachers. A main goal of the system is to provide an innovative method of fixing and testing of mastered knowledge by students in an interesting way for them. Also, it is to provide opportunities for students for self-realization. So, SAUDAI use the idea for online teaching and student learning environment. SAUDAI greatly expands their functionality and introduces gamification in the knowledge validation process, making it extremely efficient and interesting.

Key words: innovative teaching methods, students studying, IT, game.

1. Holoshchuk R. Matematychne modelyuvannya protsesiv dystantsiynoho navchannya / R. Holoshchuk, V. Lytvyn, L. Chyrun, V. Vysotska // LPNU Visnyk. – Lviv, 2003. – No. 489. – P. 100-109. 2. Rashkevych Y. Bolons'kyy protses ta nova paradyhma vyshchoyi osvity / Y. Rashkevych. – Lviv: LPNU. – 2014. – 168 p. 3. Goloshchuk R. Ínteraktivna vzaêmodíya ta zvorotniy zv'yazok u sistemí distantsíynogo navchannya /

R. Holoshchuk, V. Vysotska // LPNU Visnyk. – Lviv, 2002. – No. 464. – P. 44–53. 4. Berko, A. Intranet arkhitektura intelektual'nykh system elektronnoho navchannya / A. Berko, V. Vysotska // LPNU Visnyk. – Lviv, 2001. – No. 438. – P. 3–10. 5. Vysotska, V. A. Systema opratsyuvannya struktury elektronnoho pidruchnyka / V. A. Vysotska // LPNU Visnyk. – Lviv, 2003. – No. 489. – P. 49–63. 6. Methods based on ontologies for information resources processing / [V. Lytvyn, V. Vysotska, L. Chyrun, D. Dosyn] // LAP, Germany. – 2016. – 324 c. 7. Berko A. Systemy elektronnoyi kontent-komertsiyi: monohrafiya / A. Berko, V. Vysotska, V. Pasichnyk. – Lviv: LPNU, 2009. – 612 pp. 8. Matematychna lingvistyka / [V. Pasichnyk, Y. Shcherbyna, V. Vysotska, T. Shestakevych]. – Lviv : Novyy svit-2000, 2012. – 359 p. 9. Vysotska V. Metody i zasoby opratsyuvannya informatsiynykh resursiv v systemakh elektronnoyi kontent-komertsiyi : avtoreferat dysertatsiyi na zdobuttya naukovoho stupenya kandydata tekhnichnykh nauk : 05.13.06 – informatsiyni tekhnolohiyi / V. Vysotska; LPNU. – Lviv, 2014. – 27 pp. 10. Vysotska V. Linguistic Analysis of Textual Commercial Content for Information Resources Processing / V. Vysotska // TCSET'2016. – Lviv–Slavske, Ukraine, 2016. – P. 709–713. 11. Bisikalo O. Vyyavlennya klyuchovykh sliv na osnovi metodu kontent-monitorynhu ukrayinomovnykh tekstiv / O. Bisikalo, V. Vysotska // Naukovyy zhurnal "Radioelektronika. Informatyka. Upravlinnya.". – No. 1(36). – Zaporizhzhya: ZNTU. – 2016/1. – P. 74–83. 12. Chyrun L. Informational resources processing intellectual systems with textual commercial content linguistic analysis usage constructional means and tools development / L. Chyrun, V. Vysotska, I. Kozak // Econtechmod. – Lublin; Rzeszow, 2016. – 5(2). – P. 85–94. 13. Vysotska V. Analysis of business processes in electronic content-commerce systems / V. Vysotska, L. Chyrun, P. Kozlov // Econtechmod. –5(1). – Lublin–Rzeszyw, 2016. – P. 111–125. 14. Vysotska V. Design and analysis features of generalized electronic content-commerce systems architecture / V. Vysotska, L. Chyrun, P. Kozlov // Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska. – IAPGOS, Poland, 2016. – 6(2). – P. 48–59. 15. Vysotska V. Process Analysis in Electronic Content Commerce System / V. Vysotska, R. Hasko, V. Kuchkovskiy // CSIT'2015. – Lviv: Lviv Polytechnic, 2015.– P.120-123. 16. Lytvyn V. Designing Architecture of Electronic Content Commerce System / V. Lytvyn, V. Vysotska // CSIT'2015. – Lviv: Lviv Polytechnic, 2015. – P.115–119. 17. Vysotska V. Analysis Features of Information Resources Processing / V. Vysotska, L. Chyrun // CSIT'2015. – Lviv: Lviv Polytechnic, 2015. – P.124–128. 18. Vysotska V. The Means Structure of Information Resources Processing in Electronic Content Commerce Systems / V. Vysotska, L. Chyrun // JISCT. –3(3). – Punjab, India, 2015. – P. 241–248. 19. Alyeksyeyeva K. Tekhnolohiya upravlinnya komertsiynym web-resursom na osnovi nechitkoyi lohiky / K. Alyeksyeyeva, A.. Berko, V. Vysotska // Naukovyy zhurnal "Radioelektronika. Informatyka. Upravlinnya.". – No. 3(34). – Zaporizhzhya: ZNTU. – 2015/3. – P. 71–79. 20. Vysotska V. Kontseptual'na model' opratsyuvannya informatsiynykh resursiv v systemakh elektronnoyi kontent-komertsiyi / V. Vysotska, L. Chyrun // Matematychni mashyny i systemy. Naukovo-tekhnichnyy zhurnal. – 2015. – No. 3. – P. 179 – 190. 21. Chyrun L. Osoblyvosti metodiv kontent-analizu tekstovykh masyviv danykh web-resursiv v mezhakh rehionu / L. Chyrun, V. Kuchkovskyy, V. Vysotska // LPNU Visnyk. – No. 829. – Lviv, 2015. – P. 296–320. 22. Alyeksyeyeva K. Osoblyvosti protsesu upravlinnya web-resursom komertsiynoho kontentu na osnovi nechitkoyi lohiky / K. Alyeksyeyeva, A. Berko, V. Vysotska // LPNU Visnyk. – No. 826. – Lviv, 2015. – P. 201–211. 23. Kuchkovskyy V. Zastosuvannya metodiv Internet-marketynhu dlya analizu Web-resursiv v mezhakh rehionu / V. Kuchkovskyy, V. Vysotska, S. Nytrebych, R. Olyvko // LPNU Visnyk. – No. 832. – Lviv, 2015. – P. 129–164. 24. Vysotska V. Upravlinnya Web-proektamy elektronnoho biznesu dlya realizatsiyi komertsiynoho kontentu / V. Vysotska, A. Noha, P. Kozlov // LPNU Visnyk. – No. 814. – Lviv, 2015. – P. 421-434. 25. Kozlov P. Suchasni tekhnolohiyi upravlinnya Web-resursamy v informatsiyniy systemi analizu servisu tsyfrovoyi dystrybutsiyi / P. Kozlov, V. Vysotska, L. Chyrun // LPNU Visnyk. – No. 832. – Lviv, 2015. – P. 103–128. 26. Shakhovska N. Metody ta zasoby dystantsiynoyi osvity dlya zaokhochennya i zaluchennya suchasnoyi molodi do provedennya samostiynykh naukovykh doslidzhen' / N. Shakhovska, V. Vysotska, L. Chyrun // LPNU Visnyk. – No. 832. – Lviv, 2015. – P. 254–284.

25. Ковалюк Т. В. Узгодження вимог професійних та освітніх ІТ стандартів до компетентностей випускників ІТ спеціальностей ВНЗ

HARMONIZING THE REQUIREMENTS OF PROFESSIONAL AND EDUCATIONAL IT STANDARDS WITH THE COMPETENCIES OF GRADUATES OF HIGH EDUCATION IT SPECIALTIES

Tetiana Kovaliuk

Automated Systems of Information Processing and Management Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37, ave. Peremohy, 03056, Ukraine

The article covers necessity, concept and role of professional standards for the system of IT-industry and IT-education, international practice of their use. Purpose of this article is to contemplate ways to harmonize IT education and IT Professional Standards in terms of formation, development, accumulation, renewal, assessment of IT professionals' qualifications and competencies in accordance with European e-Competence Framework (e-CF) and Ukrainian stakeholders' requirements. Expediency analysis requirements of stakeholders in the development of IT

education content is presented. A structure and a content of professional IT standards in Ukraine are described. It is proved that professional standards provide the sphere of education with the necessary information about areas and objects of professional activity of graduating students, its kinds, tasks and necessary competences of future specialists. Ukrainian Educational IT Standards were developed without consideration of IT industry needs and European e-Competence Framework. It is shown, that development of professional standards based on competence approach, European e-Competence Framework and trends in IT education is a critical problem in Ukraine. In accordance to structure and content of industry standards of higher education, Professional standards might be the base to define competences required for realization of typical tasks of activity and operational functions of graduates with bachelors' and masters' degree. Requirements to basic knowledge specified in professional standards give an opportunity to define a list of disciplines in curricula which students of IT specialties will obtain. General characteristic of Educational Standards of Ukraine provides description of subject area of activity with the definition of study and activity objects, study purpose, theoretical content of subject area, types of professional activity, methods, methodologies and technologies, which should be obtained by graduates of higher educational institution, tools and equipment to obtain those knowledge and skills. Professional IT standards are based on international information technology standards ISO/IEC 15288:2008, ISO/IEC 12207:2008, European framework of competences (e-CF). IT professions for which professional standards are developed in Ukraine correspond to nomenclature of professional profiles of European e-Competence Framework, namely: Information Systems Specialist, Software Developer, IT Project Manager, IT Product Manager, and Information Resources Specialist. Professional standard "Information systems specialist" and "Software developer" professional standard are considered. Labor functions and labor actions of Professional standard "Information systems specialist" correspond e-CF descriptors of ICT competences relevant to business processes in information systems: PLAN, BUILD, RUN, ENABLE, MANAGE. "Software developer" professional standard defines the main purpose of professional activity namely development, debugging, testing, modification of software.

Key words: IT-industry, IT-education, professional standard, educational standard, competence, proficiency levels.

1. Zakon Ukrainy No. 1556-VII "Pro vyshchu osvitu" vid 01.07.201. 2. Postanova Kabinetu Ministriv Ukrainy vid 29.04.15 roku No. 266 "Pro zatverdzhennya pereliku haluzey znan' i spetsial'nostey, za yakymy zdiysnyuyet'sya pidhotovka zdobuvachiv vyshchoyi osvity". 3. Professyonal'nye standarty v oblasty informatsyonnykh tekhnolohyy. — M.: APKYT, 2008. – 615 s. 4. Mashukova N. D. Osnovnye napravleniya sovershenstvovaniya kachestva rabochey syly. [Elektronnyy resurs] – Rezhym dostupu: www.myshared.ru/slide/179699/. 5. M. L. Anshyna. Razrabotka novoho professional'noho standarta "Menedzher po informatsionnym tekhnolohiyam" / M. L. Anshyna. , N. S. Vol'pyan, A. Y. Oleynyk, B.B Slavyn// Kachestvo. Innovatsyy. Obrazovanye. – M.: Yzdatel'stvo "Izvestyya", 2014. – No. 2 (105), 2014. S. 55 – 62. 6. Khoreva L. V., Vasyna E. V. Obrazovatel'nye i professional'nye standarty v natsional'noy systeme obrazovaniya // Kreatyvnaya ékonomyka. – 2011. – No. 2 (50). – S. 45–51. 7. Sukhomlyn V. A. Professional'nye standarty y obrazovanye. Perpendykulyarnyy vzhlyad. – M.: VMyK MHU im. Lomonosova, "MAKS-press", 2008. 8. Zherebyna O. Professional'nye standarty v oblasty IT: "instruktsyya po prymeneniyu" [Elektronnyy resurs] – Rezhym dostupu: www.apkit.ru/files/ITStandarts_Zherebina.doc. 9. Professional'nye standarty v oblasty UT. [Elektronnyy resurs] – Rezhym dostupu: http://www.apkit.ru/committees/education/meetings/ standarts.php. 10. Nakaz MON Ukrainy vid 01.06.2016 No. 600 "Pro zatverdzhennya ta vvedennya v diyu Metodychnykh rekomendatsiy shchodo rozroblennya standartiv vyshchoyi osvity". 11. Tuning Educational Structures in Europe – Final Report Pilot Project. [Electrinic resource] – Access mode: http://tuningacademy.org/wp-content/uploads/2014/02/TuningEUI_Final-Report_EN.pdf. 12. Profesiynyy standart. Fakhivets' z informatsiynykh system. [Elektronnyy resurs]. Rezhym dostupu: http://mon.gov.ua/ content/%D0% 9D%D0%BE%D0% B2%D0%B8%D0%BD%D0%B8/2016/03/15/5-ps-spes-infosystems-13.12.2014.pdf. 13. Profesiynyy standart. Fakhivets z rozrobky programnoho zabezpechennya. [Elektronnyy resurs] – Rezhym dostupu: http://mon.gov.ua/content/%D0%9D%D0% BE%D0%B2%D0% B8%D0%BD%D0%B8/2016/03/15/6-ps-rozrobnik-pz-13.12.2014.pdf. 14. Proekty standartiv vyshchoyi osvity [Elektronnyy resurs]. – Rezhym dostupu: http://mon.gov.ua/activity/education/reforma-osviti/naukovo-metodichna-rada-ministerstva/proekti-standartiv-vishhoyi-osviti.html. 19. Postanova Kabinetu Ministriv Ukrayiny vid 23.11.2011 r. No. 1341 "Pro zatverdzhennya natsional noyi ramky kvalifikatsiy". 20. Kovaliuk T. V. Profesiyni standarty v galuzi informatsiynykh tekhnolohiy ta yikh harmonizatsiya z IT-osvitoyu Ukrainy. Kovaliuk T. V., Kobets' N. M. // Inzheneriya programnoho zabezpechennya. Naukovyy zhurnal NAU. No. 21 (1), 2015. c. 46–55. 21. Curricula Recommendations. [Electronic resource] – Access mode: http://www.acm.org/education/curricula-recommendations. 22. Postanova KMU No. 53 vid 1.02.2017 "Pro vnesennya zmin do postanovy Kabinetu Ministriv Ukrainy vid 29 kvitnya 2015 r. No. 266". 23. MSIS 2016. Global Competency Model for Graduate Degree Programs in Information Systems. The Joint ACM/AIS MSIS 2016 Task Force. December 5, 2016. [Electronic resource]. – Access mode: https://app.box.com/s/42c3kzn2dlxyyjm5udawlh9a6n8dx46x.