[Vieira et al., 2005] Vieira, G.E., Cesar, O. Jr.: A conceptual model for the creation of supply chains models. Proceedings of the 37th conference on Winter simulation, pp. 2619 – 2627, Orlando, Florida, (2005)

Authors' Information

Jacek Jakieła, Ph.D., Eng. *– Department of Computer Science FMEA RUT; W. Pola 2, 35-959 Rzeszow, Poland; e-mail: jjakiela@prz.edu.pl*

Major Fields of Scientific Research: Software Development Methodologies, Agent and Object-Oriented Business Modeling, Internet Enterprises Models, Computational Organization Theory and Multi-Agent Simulation of Business Architectures.

Paweł Litwin, Ph.D., Eng. *– Department of Computer Science FMEA RUT; W. Pola 2, 35-959 Rzeszow, Poland; e-mail: plitwin@prz.edu.pl*

Major Fields of Scientific Research: Applications of Neural Networks in Mechanics, Computer Simulations, Finite Element Method.

Marcin Olech, M.Phil., Eng. *– Department of Computer Science FMEA RUT; W. Pola 2, 35-959 Rzeszow, Poland; e-mail: molech@prz.edu.pl*

Major Fields of Scientific Research: Multi-agent Simulation, Application of Artificial Intelligence in Industry.

SEMANTICALLY RICH EDUCATIONAL WORD GAMES ENHANCED BY SOFTWARE AGENTS

Boyan Bontchev, Sergey Varbanov, Dessislava Vassileva

Abstract*: With steadily evolving new paradigms for technology enhanced learning, educational word games such as quizzes, puzzles and quests raise new appeal and motivation for students in following game based educational processes. Traditional word games may be applied more successfully to game based learning in given scientific domain provided they are highly oriented to the content and problems of that domain. Such games may be more efficient if they include artificial agents simulating opponents, advisors or collaborators of the player. Authors present a semantic structuring model of learning content for logic and word games serving for educational purposes and, next, show the place of artificial agents within the game construction and possible ways of agent's realization. There are given results from practical experiments with playing a memory game using the semantic content model, without and with agents integrated into the game.*

Keywords*: Agents, quiz, word game, e-learning, game based learning.*

Introduction

The last developments of the Internet technologies and their usage for technology enhanced learning require blending traditional teaching and training with online processes. New trends in evolution of modern e-learning comprise various approaches of applying educational games in a complement to the traditional instructional learning design [Dempsey et al, 1996]. Word and logic games are considered as a rather effective mean for retaining interest of learners by attracting their attention for much more time than other approaches. Educational games appeared to be not only entertaining but also cognitive and educative for today learners, as far as they have precise definition of goals, constraints, built-in rules and consequences [Prensky, 2006]. Such logic games help learners to improve their knowledge in given domain by challenging them to solve various problems or to use practices within the domain area and, thus, to improve creative thinking [Salen and Zimmerman, 2003]. Mark Prensky explains in [Prensky, 2006] why and how games deliver pleasure and emotions while the learner plays a game. Games impose conflicts and competitions by using rules and establishing goals in order to facilitate evolution of creativity, motivate learners and, as well, to satisfy their ego in an interactive and pleasant way.

Nowadays, the mostly applied educational games are represented by quizzes, puzzles, quests and problem solving staging [Batson and Feinberg, 2006], [Ferreira et al, 2008]. Such logic word games make use of textual and even multimedia content of a proper domain and apply predefined rules for solving some problems, sometimes supported by counters or dices according given educational purposes. They may be represented by board games [Bontchev and Vassileva, 2011] which may be played by a single player alone or against a simple or intelligent software agent replacing the real opponent.

The paper describes a principal model for semantically enriched educational word games such as quizzes, puzzles and quests and, next, its realization by using an agent based software architecture. It explains in brief a model of semantic structuring of domain content by means of UML class diagrams. The model is used as a basic paradigm for the development of educational word games. The paper presents how it facilitates the implementation of a positional memory quest board game for matching symbols with minimum number of mouse clicks. Next, it shows how a software agent is used within the memory game as opponent of real player. There are given practical results from an experimental field trial aimed at evaluation of student appreciation of the single user memory game, with and without software agent.

Motivational Background

Educational games used to support e-learning may vary from simple single user word games to complex multi-user collaboration games using augmented reality. The paper is focused on word games like quizzes, puzzles and quests because they suppose a rather simple construction process and, at the same time, are very useful for exercises and self-assessment.

Educational Logic Games

Quizzes are popular as word games with educational goal thanks to their easy implementation. They can be used by students and teachers for self-assessment tests and control exams. Besides assessment of knowledge, they are purposed for producing more fun and increasing motivation for learning. For this reason, some of the educational logic and word games are implemented as combinations between board games and quizzes. This type of games usually uses board rules for navigation within a quiz. Thus depending on his/her position on the board, the player gets a question with a certain complexity [Bontchev and Vassileva, 2010]. Educational quizzes could be developed and managed using different existing tools such as Quiz Center and Quiz Builder [Bontchev and Vassileva, 2011]. These tools have opportunities for automatically evaluating students' answers, some of

them automatically generate questions from curriculum [Guetl et al, 2005] and other use authoring tools to create quizzes [Retalis, 2008]. In some authoring tools there can be created not only simple questions, but also parameterized ones [Feng, 2005].

In recent years, games are often used in e-learning as means of students' attention to being actively involved in learning process and as means of achieving a higher motivation for studying [Aleksieva-Petrova and Petrov, 2011]. Games are also successfully used in adaptive e-learning systems, e.g. like adaptive quizzes such as QuizGuide and QuizJet [Hsiao et al, 2009] and adaptive board games such as ELG [Retalis, 2008]. In the ELG different learning scenarios could be presented in the form of board games, through which students are expected to improve their performance and extend their knowledge. Games created with the ELG can be customized on different parameters such as learner's level of knowledge, preferences and educational goals. Games such as QuizGuide and QuizJet [Hsiao et al, 2009] support students to select self-assessment quizzes most suitable for learner's goals and preferences. These systems use adaptive navigation to increase learners' knowledge by selecting most important topic of training.

Generally educational word and board games use rule-based approaches for navigation through its board and require detailed knowledge of the subject concerned [Retalis, 2008]. It could be used different strategies for selecting a question, its level of difficulties and area of knowledge, depending on the learner/player profile (goals and preferences, learning style and performance) [Bontchev and Vassileva, 2010].

Types of Software Agents for Educational Games

Educational games may have three main different modes of play. These modes are multiple users mode, single user mode and single user against a simulated player mode. The last of these three modes of play appeared to be one of the most preferred by end users. Usually the simulated player is implemented by means of an intelligent software agent. Usage of software agents is an established and effective approach for realizations of various types of simulations. According [Varbanov et al, 2007] intelligent agents may be applied in several aspects and cases of serious games. Main cases of them are as follows:

- The case where a part of the game is implemented as an intelligent agent. The software agent performs a simulation process and its behavior unlike other cases is similar in each game and each player. In this case, the participation of the intelligent agent is hidden for gamers and it is realized as an autonomous system.
- The case in which the intelligent agents acts as opponents of the player. The aim is to simulate the competition and to encourage gamers to achieve better results. The software agents which participate in a game have different behavior and their numbers may vary depending on the level of a user.
- The case where an intelligent agent participates as business partner or/and task collaborator. The player communicates and performs various activities together with this type agent.
- The case where the intelligent agents are included in a game as assistants to players. These software agents are responsible for providing necessary information, guidance, comments and recommendations.

The agent may have various levels of implementation complexity, e.g. from a simple search procedure to controlling multidimensional state space of a virtual world model. Moreover, complex agents may simulate social behavior with knowledge about the game and adaptation to the physical player. As well, agents may cooperate within a community of social agents.

Semantically Rich Educational Games

Logic and board games like word and problem-oriented games need a special organization of course content in order to integrate it to the game. A semantic structuring of the content will enable game engine to extract specific terms and their inter-relationships dynamically during the play process. It follows a brief representation of such a model and a sample content organization based on it.

Semantically-Rich Content Model for Educational Word Games

The model for semantic content organization for educational games was proposed in (Bontchev and Vassileva 2011) and is based on UML class diagrams. It models visually content relationships by class hierarchies, instances, class attributes, metadata and relationships plus axioms. For any class, there may be presented its possible super-classes by means of *IS_A* link. A class instance is linked to its class by an *instance Of* reference. A class may contain a content description text and image (again with public visibility and classifier scope) and a list of class properties and resources. Names and values of both properties and resources may be of private, protected, public of package scope and may be annotated within metadata. Any concept may be related to another term by UML relationships of type dependency, association, aggregation, and composition. As a minimum, a relationship may have name, direction, and roles and cardinality of the inter-related entities.

Figure 1 shows a UML class diagram of some terms from a bachelor course in XML technologies. Terms are semantically structured by their inter-relationships and, possibly, with some instances. For example, the terms “ANY”, “EMPTY” and “PCDATA” are subtypes of “Element type DTD declaration” having with it IS_A relations. “Element type DTD declaration” may compose “Sequence list”, “Choice list” and “Element quantifier”. There is shown only two instances – one of “Choice list” and another of “PCDATA”.

In fact, fig. 1 illustrates only a small part of the whole UML diagram of XML terms plus their relationships and instances. Such semantically-rich content models are comprised by a plenty of specific terms (such as DTD, XML Schema, DOM, SAX, etc.) and their interrelationships. These terms and relationships are to be extracted and integrated (preferably, automatically) in any board games suitable for educational purposes. The section below represents an educational board game using this content and appropriate for agent-based implementation.

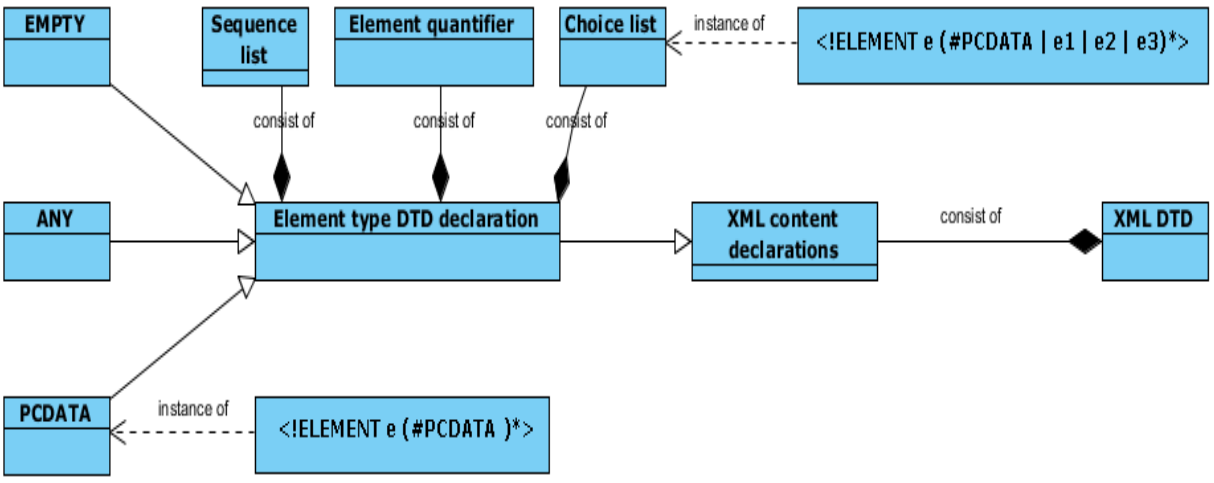


Fig. 1. UML diagram of some XML terms and their relationships and instances

A Memory Game Using the Semantic Content Model

Authors have investigated several educational logic games (both word and board games) extracting content from the semantic model. The simplest case of such games makes use of a single entity (class or instance) without utilizing its possible inheritance and/or association relations to other entities. Simple word games serve as good examples here, e.g. the classical hangman game and anagram games. More attracting games may be constructed by using type information about the entity, its properties and instances.

For a mass invasion of games usage in education, teachers need of simple and rapid process of constructing educational games. Classical word puzzles may be easily constructed as board mini-games (Bontchev and Vassileva, 2010). The building parts of a work puzzle being partial images or letters may be represented as objects, which are be moved to the correct positions while wrong moves are possible. More complex logical quizzes and quest may require not only simple moves but also some other player action types like single and double clicks onto objects and object relations. Fig. 2 provides a snapshot of a positional memory board game for matching XML terms to their instances – both linked into couples by means of *instanceOf* relationships in fig. 1.

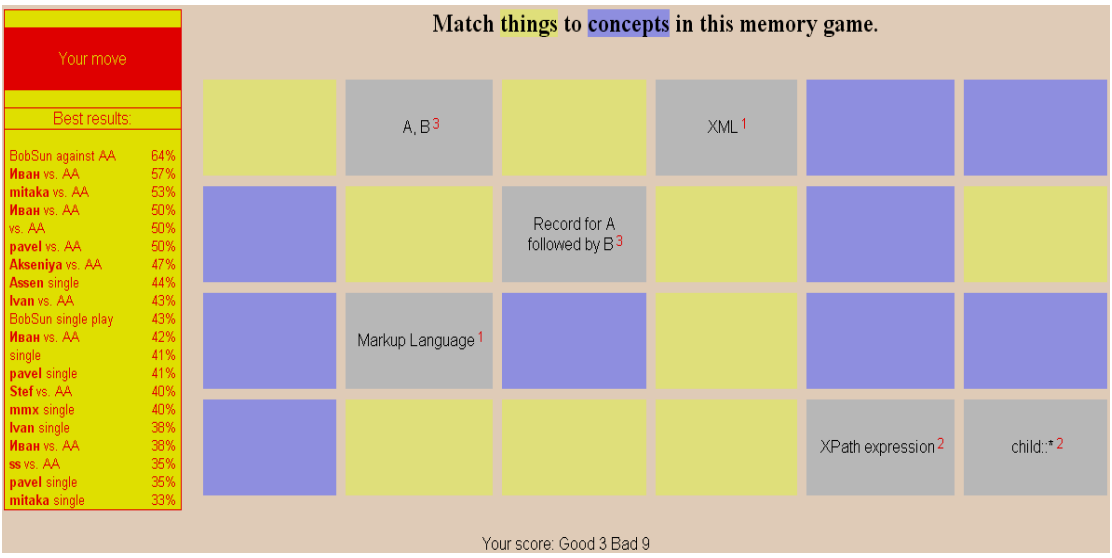


Fig. 2. Positional memory quest board game for matching symbols with minimum number of mouse clicks

The player has to click first onto a blue cell in order to see its hidden term. Next, he/she should click onto a yellow cell supposed to contain the hidden instance of this term. The yellow cell shows its contents and, if the instance matches the term, the player receives a point and both the term and instance are appended by a red suffix showing the number of that match, and continue being displayed. Otherwise, the player looses a point and both the term and instance disappear. The game is over when all the terms of the blue cells are matched to their instances.

Left on fig. 2, there is given a list of players' results until the present moment. Some of them are reached within single user games as explained above, while others are resulted in playing against an artificial agent called AA. In the last case, game play is controlled in turns – the player tries to match a term to its concept and next, the agent does the same. Thus, player is supposed to learn not only from his/her mistakes but from wrong moves of the agent, too.

The Software Agent Architecture

In one of our previous works [Varbanov et al, 2007], we have used a three-layered architecture of the software agent, shown in fig. 3. The first layer contains methods for interfacing the virtual world in which the agent acts. The second layer represents the model of the virtual world, which the agent builds and uses. The third layer contains the agent's decision-making functionality, which applied to the virtual world model implements the decision-making process. For the implementation of those three layers, different technologies were used, accordingly:

- Java methods to implement agent's sensors and effectors, communicating with the game server via CORBA;
- Protégé [Drummond et al, 2005], as knowledge base management system to implement the virtual world model;
- Algernon [Algernon, 2005], as forward-backward rule based inference engine to implement the agent's decision-making functionality.

That approach, although fruitful, left the impression that in many cases the full power of Java, Protégé and Algernon would not be needed. Therefore, this three-layered architecture could be implemented by more "lightweight" means. The ideal solution would consist of maximum lightweight tools for the trivial cases, and enough openness to provide for easy inclusion of heavyweight tools, such as Protégé and Algernon, if needed.

In [Bontchev et al, 2010], we've described architecture based on Persevere Server [Persevere, 2011] - an object storage engine and application server (running on Java/Rhino) that provides persistent data storage of dynamic JSON data in an interactive server side JavaScript environment (fig. 4).

Placing Persevere Server at the base of our intelligent software agent's implementation, the three layers architecture mapping to technologies evolves to:

- JavaScript methods to implement agent's sensors and effectors, communicating with the game server through a standard JSON HTTP/REST Web interface;
- Persistent JavaScript/JSON classes, methods and instances to implement the virtual world model;
- JavaScript methods to implement the agent's decision-making functionality, when the decision-making process is not too complex.

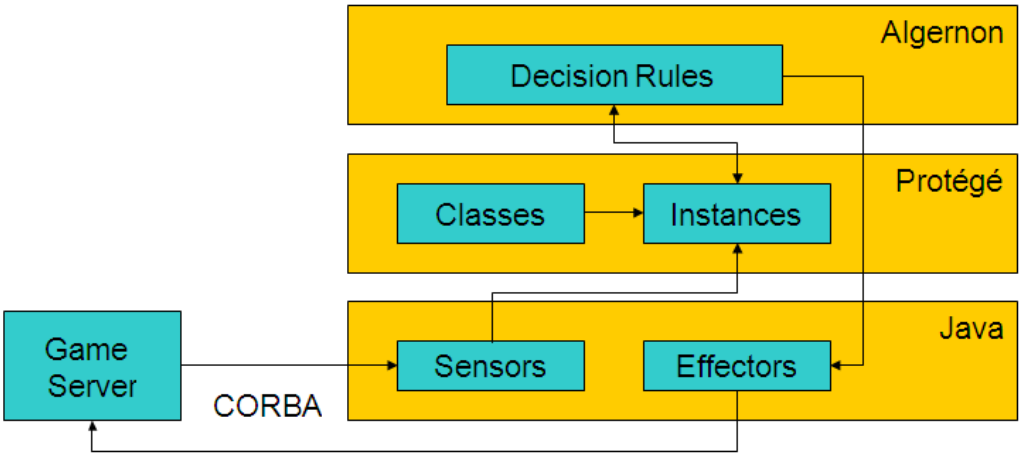


Fig. 3. Three layer agent's architecture used in PRIME project [Bontchev et al, 2010]

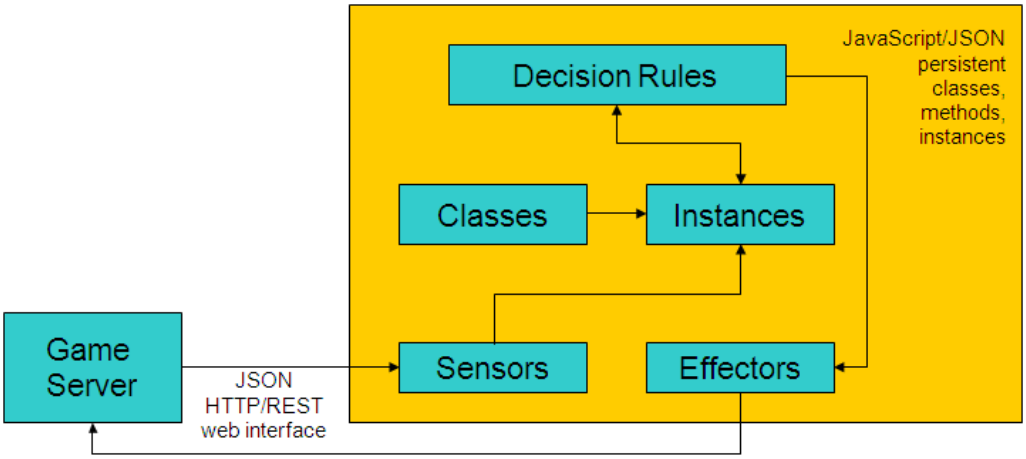


Fig. 4. A lightweight implementation based on Persevere Server [Bontchev et al, 2010]

If the case at hand imposes complex intelligent software agents, Persevere Server provides for easy inclusion of additional Java libraries. In fact, any tool written in Java and having well-documented Java API could be easily included and used.

This architecture is used in the presented here experiments.

Implementation of the Memory Game Using Agent as Opponent

The memory game does not impose high complexity on the participating software agent, comparing with other types of games. Thus the agent's resources are concentrated on adapting a chosen strategy to the user's evolving model, built and exploited by the agent.

Let us consider several specific opponent agents, demonstrating, although quite in a basic way, how the agent's functionality responds to the strategic goal(s) we pursue:

1. Goal: Mimic human player's features
   - STM agent – a simple agent, which knows the proper associations, but simulates short-term memory (STM) capacity, i.e. memorizes only the last N moves. Comparing the user success to the results of agents initialized with different N values, some conclusions about the users STM capacity might be drawn.
   - PK agent – a simple agent, which models partial knowledge of the proper associations, by using parameter 0<P<=1, representing the probability to guess properly the correct association.
   - Note: Both described agents might be adaptable – they can dynamically change their N or P parameters, to keep their success rate close to the human player's.
2. Goal: Stimulate human's performance
   - Average Adaptive agent. The agent "knows" all tile's values from the start, but deliberately makes "wrong" moves to keep in pace with the human's rate of success. Sometimes, to help the user, the agent may make "wrong" moves by turning a tile, most appropriate for the human's next move. This agent might function in tolerant mode (by keeping its success rate a little bit below human's average) or in aggressive mode (by keeping its success rate a little bit above human's average), thus stimulating the user to perform better.

Those and more agent's types can be used in experimenting with different user target groups in order to specify more precisely the relationship between user classes and appropriate agent models and parameters, implementing strategies.

The memory game presented in 3.2 is implemented including an option to play against software agent of the Average Adaptive type.

Practical Experiments and Results

The experimental field trial presented here aimed at evaluation of student appreciation of single user memory games with XML content - without and with software agents. In executed practical experiments participated 30 four-year students of the bachelor program in Software engineering at Sofia University, Bulgaria. After the game play, students filled a questionnaire about the efficiency of such game based learning approach. The questions had answers in Likert scale (Strongly disagree, Disagree, Neutral (neither agree nor disagree), Agree, Strongly agree) as given below:

- Q1: The presence of an agent motivates me to play better and more.
- Q2: The presence of the agent suppresses me and prevents me from playing well.
- Q3: The agent behavior designed within the game simulates successfully another real player.
- Q4: Educational games representing problems to be solved using course content are very useful for self-learning and self-test.
- Q5: I feel educational games will have a positive contribution to University education and technology enhanced learning at all.

As well, the questionnaire contained a question with answers in non-Likert scale:

- Q6: What type of game construction is most suitable for e-learning purposes? – possible answers here are as follows:
    1. Single user games with no agents
    2. Single user games with simple software agents
    3. Single user games with artificial intelligence (AI) agents
    4. Multi-user games with no agents
    5. Multi-user games with simple software agents
    6. Multi-user games with AI agents

Figure 5 provides a view of answers for questions from Q1 to Q5. It is obvious that the majority of learners find software agents do not prevent them from playing well even more – a presence of such an agent does stimulate them within the game play. However, students cannot agree on successful simulation of real player behavior on behalf of a software agent in the course of an educational game (question Q3) – in fact, their opinions are divided into two quasi equal groups. At the same time, all of them share the opinion that educational games representing problems to be solved using course content are very useful for self-learning and self-assessment (Q4). In general, they think games will have a positive contribution to University education and technology enhanced learning at all.

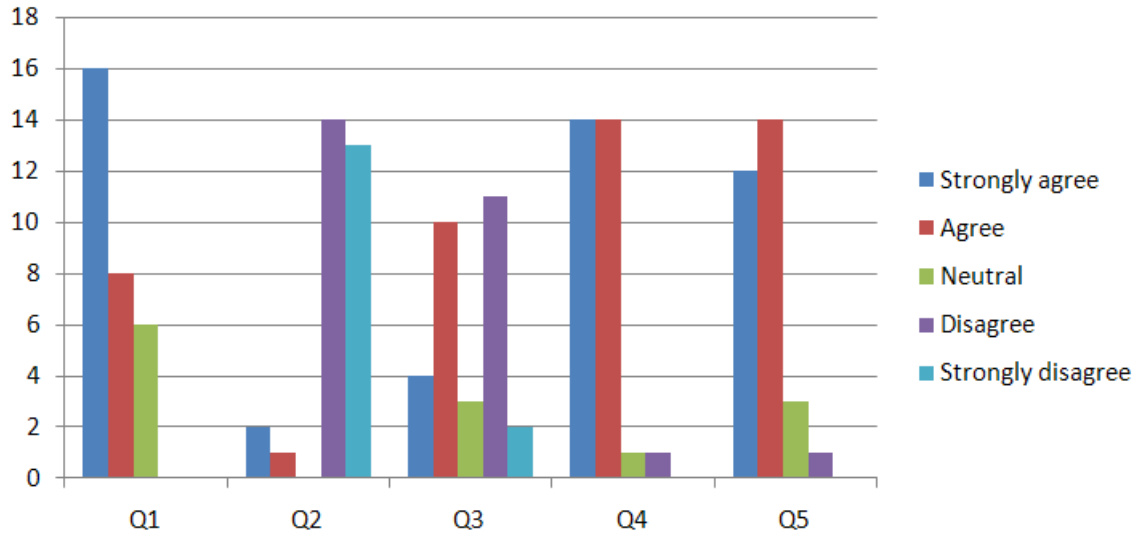


Fig. 5. Results for questions Q1 ÷ Q5

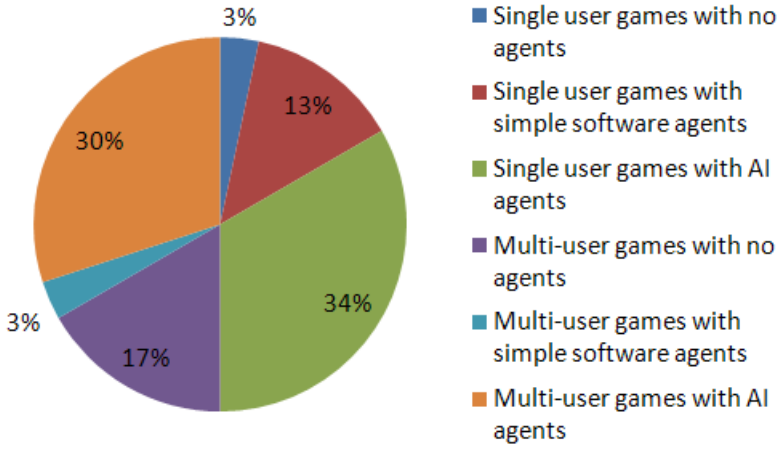


Fig. 6. Answers for question Q6

Figure 6 represents distribution of the answers of question Q6. The majority of students find both single and multi-user games with AI agents most suitable for e-learning purposes. As well, they agree on the assumption that even games with simple software agents will have a positive impact on educational process, especially single user games. The found results sound very encouraging in terms of new development of intelligent agents for educational games.

Conclusions and Future Works

The present paper proposed a way of using semantically structured courseware content in simple educational logic games such as word or board games. Word games as simple or enhanced quizzes, mazes and quests for solving other logical problems are of great importance for modern e-learning as far as they offer a plenty of benefits and are very useful for self-learning and self-testing. They may be combined with some principles of board games in order to represent more complex logical problems which are to be solved by player actions under execution of some rules upon context conditions [Bontchev and Vassileva, 2010]. Thus, the game context could be based on semantic relationships among domain concepts and terms and, as well, may enable inclusion of a

more complex set of e-learning instructions. Moreover, the game context may include some learning activities such as scenarios typical for a given virtual world, for obtaining new knowledge and/or for maintaining cognitive skills. The presented example of a board game for training memory in semantic relationships within the XML domain is only one of the many examples of such games. It relies on extraction of semantically structured content which may be done by hand or automatically, by extraction of domain concepts and their relationships from an UML description file presented as XMI (XML Metadata Interchange) document.

The second focus of the paper was directed on construction and usage of software agents for educational board games relying on open architectures and agent programming. The software application architecture was chosen specially for shortening both the design and implementation phases. Authors plan to develop further the game to evolve in a more complex dynamic model. As well, the applied software agent may be rather primitive in the beginning and, next, to be replaced with an enhanced intelligent agent.

The proposed combination of semantic model, technologies and tools provides a good ground for further development of other word and board games and, on other side, for creation of new, more intelligent software agents. The conducted survey shows that presence of an agent motivates students to play better and more. As well, the authors are highly encouraged by the students' appreciation of the agent behavior designed within the game which may simulate successfully another real player. They plan to develop more intelligent agents in order to incorporate them into new single- or multi-player educational games.

Acknowledgments.

The work reported in this paper is supported by the ADOPTA project funded by the Bulgarian National Science Fund under agreement no. D002/155.

Bibliography

[Algernon, 2005], Algernon Tutorial, http://algernon-j.sourceforge.net/

[Aleksieva-Petrova and Petrov, 2011] Aleksieva-Petrova A., Petrov, M. ADOPTA Model of Learmer and Educational Game Structure, Int. Conf. CompSysTech'11, 16-17. June 2011, Viena, Austria (in print)

[Batson and Feinberg, 2006] Batson, L., Feinberg, S.: Game Designs that Enhance Motivation and Learning for Teenagers, Electronic Journal for the Integration of Technology in Education, Vol. 5, pp. 34-43, 2006.

[Bontchev and Vassileva, 2010] Bontchev, B., Vassileva, D.: Modeling Educational Quizzes as Board Games, Proc. of IADIS Int. Conf. e-Society'2010, ISBN: 978-972-8939-07-6, March 18-21, Porto, Portugal, pp.20-27, 2010.

[Bontchev and Vassileva, 2011] Bontchev, B., Vassileva, D.: A Light-Weight Semantic Content Model for Mobile Game Based Learning, Proc. of IADIS Mobile Learning 2011, Avila, Spain, ISBN:978-972-8939-45-8, pp.177-182.

[Bontchev et al, 2010] Bontchev, B., Varbanov S., Vassileva, D.: Software Agents in Educational Board Games, LATEST TRENDS on ENGINEERING EDUCATION, 7th WSEAS International Conference on ENGINEERING EDUCATION (EDUCATION '10), Corfu Island, Greece, July 22-24, 2010, ISSN: 1792-426X, ISBN: 978-960-474-202-8, pp.387-392.

[Dempsey et al, 1996] Dempsey, J. et al.: Instructional applications of computer games". American Educational Research Association, New York, 1996, pp. 8–12, 1996.

[Drummond et al, 2005] Drummond N., M. Horridge, Knoblauch, H.: Protégé-OWL Tutorial, Proc. of 8th Int. Protégé Conf., Madrid, July 2005 (also at http://protege.stanford.edu/), 2005.

[Feng, 2005] Feng, K.: Joyce: A Multi-Player Game on One-on-one Digital Classroom Environment for Practicing Fractions, Proc. of the Fifth IEEE Int. Conf. on Advanced Learning Technologies (ICALT'05), pp. 543-544, 2005.

[Ferreira et al, 2008] Ferreira, A. et al.: The common sense-based educational quiz game framework "What is it?", ACM Int. Conf. Proceeding Series. Vol. 378, pp. 338-339, 2008.

[Guetl et al, 2005] Guetl, C. et al.: Game-based E-Learning Applications of E-Tester. In P. Kommers & G. Richards (Eds.), Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications 2005, pp. 4912-4917, 2005.

[Hsiao et al, 2009] Hsiao, I-Han, Sosnovsky S., Brusilovsky, B.: Adaptive Navigation Support for Parameterized Questions in Object-Oriented Programming, LNCS, Vol. 5794/2009, ISBN 978-3-642-04635-3, pp. 88-98, 2009.

[Persevere, 2011], Persevere Documentation, http://docs.persvr.org/

[Prensky, 2006] Prensky, M.: Don't bother me, mom, I'm learning!, St. Paul, Minn. MN: Paragon House, 2006.

[Retalis, 2008] Retalis, S.: Creating Adaptive e-Learning Board Games for School Settings Using the ELG Environment, J. of Universal Computer Science, vol. 14, no. 17 (2008), pp. 2897-2908, 2008.

[Salen and Zimmerman, 2003] Salen, K., Zimmerman, E.: Rules of play: Game design fundamentals, Cambridge, MA, USA, October, MIT Press, 2003.

[Varbanov et al, 2007] Varbanov S., Stoykov S., Bontchev B., Software Agents in a Serious Business Game – the PRIME Project Case, LG2007 Learning With Games, eds. Marco Taisch and Jacopo Cassina, Sophia Antipolis, France, 24 – 26 September 2007, pp. 301 – 306, ISDN 978-88-901168-0-3, 2007.

[White et al, 2004] White, M. et al.: ARCO - an architecture for digitization, management and presentation of virtual exhibitions, Proc. of Computer Graphics Int. Conf., pp. 622-625, 2004.

Authors' Information

Boyan Bontchev *– Assoc. Prof., PhD, Sofia University "St. Kl. Ohridski", 125, Tzarigradsko shoes, bl. 2, Sofia 1113, Bulgaria; e-mail: bbontchev@fmi.uni-sofia.bg*

Major Fields of Scientific Research: Software architectures, Technology enhanced learning, Educational games, Adaptive systems

Sergey Varbanov *– Assist. Prof., Institute of Mathematics and Informatics, Acad. Georgi Bonchev Str., Block 8, Sofia 1113, Bulgaria; e-mail: varbanov@fmi.uni-sofia.bg*

Major Fields of Scientific Research: Artificial Intelligence, Artificial Intelligence Applications, Intelligent Agents, Neural Networks, Genetic Algorithms, Educational Games, Serious Games.

Dessislava Vassileva *– Research Assoc., PhD, NIS at Sofia University "St. Kl. Ohridski", 125, Tzarigradsko shoes, bl. 2, Sofia 1113, Bulgaria; e-mail: ddessy@fmi.uni-sofia.bg*

Major Fields of Scientific Research: Software technologies, Technology enhanced learning, Educational games, Adaptive hypermedia systems

- serial, parallel-serial and dichotomy partitioning,
- initial partitioning,
- partitioning optimization;

- packaging:
  - serial and parallel-serial packaging,
  - initial packaging,
  - packaging optimization;
- placement:
  - hierarchical initial placement by multilevel macromodels,
  - multilevel placement optimization by scanning area method with macromodels.

For these problems also were developed special algorithms for escaping from the local extreme.

The proposed algorithms have such properties:

- can be efficient for choosing the appropriate number of partitions to divide the circuit;
- arbitrary division ratio can be chosen for partitioning;
- many same procedures can be used for initial solution and their optimization;
- close to linear computational complexity;
- provide good quality of solutions;
- are appropriate for large and very large-scale problems.

Most likely, the first proposal to use the free hierarchical clustering for partitioning was in [Bazylevych, 1975]. It was further developed in [Bazylevych, 1981] and used for packaging and placement [Bazylevych, 2000, 2002, 2007] with good results. More lately hierarchical clustering, especially enforced, was used for hyper graph partitioning [Garbers, 1990], [Cong, 1993], [Dutt, 1996], [Karypis, 1997], [Saab, 2000] and for others problems.

For all test cases investigated, the results are not worse, and in many cases they are better comparatively with obtained by other known methods. For some cases, the optimal results were received for the first time.

Main stages for solving the problems by hierarchical clustering

For solving the large-scale intractable combinatorial problems at the first step we must perform aggregation. We divide large problems into the set of small ones that are simulated by macromodels. Every macromodel include the fixed number of initial circuit elements. The number of macromodels in aggregated circuit is also very important. It is possible to create multilevel model in such way that the number of macromodels and numbers of their elements at very level of decomposition must be not very large to receive good quality solution. We build multilevel system of hierarchically built macromodels. In such system every subproblem could be solved with the high quality for not large CPU time. The main decisions that we must make in such approach are:

- how to chose the number of elements of basic subproblems that we can solve with high quality in a reasonable running time,
- what must be the number of level in macromodeling,
- what method is desirable to use for solving the basic subproblems,
- how merge the partial solutions of subproblems into one solution of whole problem,
- do we need to use additional optimization (refinement) algorithms or not,
- how to escape from the local extreme at every level of macromodeling ?

One of the first problems that appear here is to create the hierarchical macromodel of initial system. Thereto we must receive multilevel aggregation of circuit. One way is to reveal hierarchical built clusters. For this reason it is

possible in electrical system to use the Optimal Circuit Reduction method [Bazylevych, 1975, 1981]. By this method the problem solving is divided into the following steps (Figure 1):

- the bottom-up free hierarchical circuit clustering;
- the mathematical description of clusters by macromodels;
- the top-down multilevel solving with receiving global initial solutions and theirs local optimizations with macromodels at every level of decomposition.

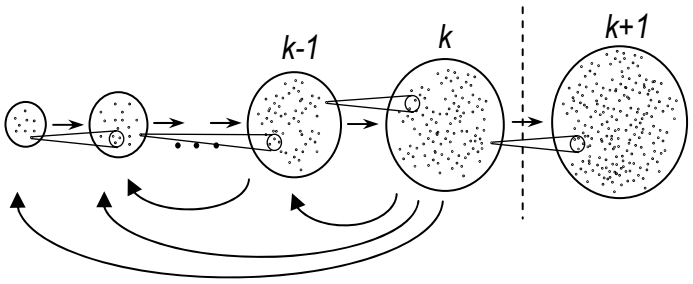


Fig. 1. Bottom-up (left arrows) hierarchical clustering and top-down (right arrows) problem solving

The main feathers of developed approach are:

- the problem size as a whole (the number of variables) increases step by step during the solution process from substantially reduced, initially (level 1) to real (level k);
- the number of tasks which are to be solved increases on each recursive level. However, all of them would be not large, and are properly hierarchically inserted one into the other, and thus can be solved by the same basic procedure with high quality.

The idea was to operate not by original elements, the number of which is extraordinarily high, but by the hierarchically built clusters of arbitrary sizes (not large) that could be mathematically described by macromodels. The (*k+1*)-level (Figure 1) shows that simulating the problem by 0-1 models (binary programming) will significantly increase the number of variables. This case can not simplify the problem solving.

This enables us:

- to essentially decrease the size of the problem, facilitating a solution and reducing the calculation consumption, the large size problem is reduced to recursive solving of small unique tasks;
- to improve the quality of the solution by more easier trapping into the zone of the global optimum. The number of local extrema is significantly smaller.

The Optimal Circuit Reduction Method

The Optimal Circuit Reduction (OCR) method builds the Optimal Reduction Tree T^R (Figure 2). It is a rooted (generally *n*-ary) tree which leaves (level *1*) correspond to the set $P = \{p_1, \dots, p_n\}$ of circuit elements and a root (level *H*) - to all aggregated circuit.

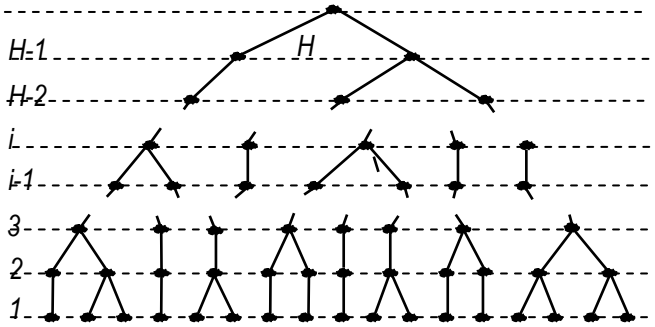


Fig. 2. The Optimal Reduction Tree

The main steps of T^R generation are:

- consider the set C_i of all clusters in the level *i*. At the first level we consider a set of all initial elements $P = \{p_1, \ldots, p_n\}$;
- form a set for all pairs of adjacent clusters for every cluster of the set C_i;
- calculate the merging criterion values for all pairs of adjacent clusters;
- create the ordered list $L(\eta)$ of pairs of adjacent clusters by the chosen merging criterion;
- form the new set C_{i+1} of clusters of the $(i + 1)$-th level. There are several possibilities. In the best case - free clustering - we merge only the maximum number of independent pairs of adjacent clusters with the best value of the merging criterion. It could cause a large tree's height and consequently takes a lot of CPU time. The one possible way to reduce the running time is to take all independent pairs with ε given decreasing (increasing) of the best criterion value. The second way is to merge the first λ of all possible independent pairs, where λ ($0 < \lambda \leq 1$) - is a reduction parameter. It is partially enforced clustering. In the last case when $\lambda = 1$ we merge all clusters. It is enforced clustering. Here a height of the ORT is a minimal and therefore it takes the minimal running time but results could be worse. This case corresponds to the enforced circuit reduction that might not generate good natural clusters, because at every level we must merge together some clusters that do not have good criterion's value;
- form the new $(i + 1)$-th level of the tree T^R by including a set of the new clusters, defined by merging and the rest clusters from the previous level that are not merged.

We must draw attention that we do not have to build binary Reduction Tree obviously. If, for example, one element creates two or more pairs with the same criterion's value, it is possible to join three or more elements together at one step. It reduces the height of tree and, of course, the CPU time. It is not easy to choose the criteria for clusters merging to receive the best clusters. There are many possible merging criteria. The main of them are:

- maximization the full number of the internal clusters' nets;
- minimization the full number of the external clusters' nets;
- maximization the full number of the subtraction of the of internal and external nets.

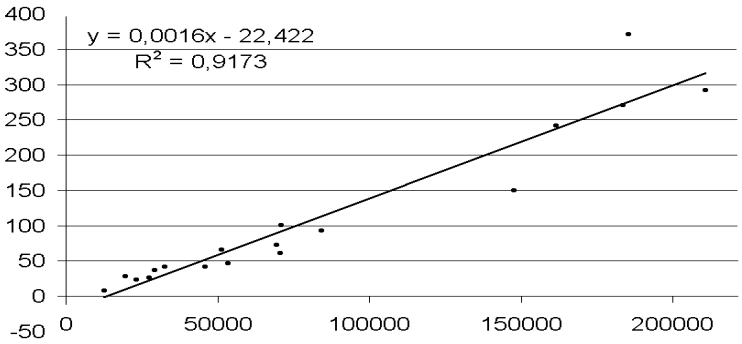


Fig.3. Circuit size vs time for TR building

There are possibilities of exploiting the various mixed merging criteria with the weight coefficients for the individual nets, the element numbers, the sizes of clusters, the time delay, etc. Very important is also the dependency circuit size vs time. As our experiments show, this dependency is close to linear by using the OCR method (Figure 3). Experiments are conducted with the library of the IBM01-IBM18 [Alpert, 1998].

Figure 4 shows the example of the Reduction Tree T^R and the dependency of the cluster external nets'number vs reduction steps starting from element 11 for some circuit with 17 elements. It can help to receive the better dividing the circuit into partitions. The cutting χ_8 at eighth level shows that partitions (1, 2, 3, 4, 5, 6, 6, 7, 8, 9. 10) and (11, 12, 13, 14, 15,16, 17) create the minimal cut with 5 nets. Other cuttings have more external nets. No other partitioning method has such possibility.

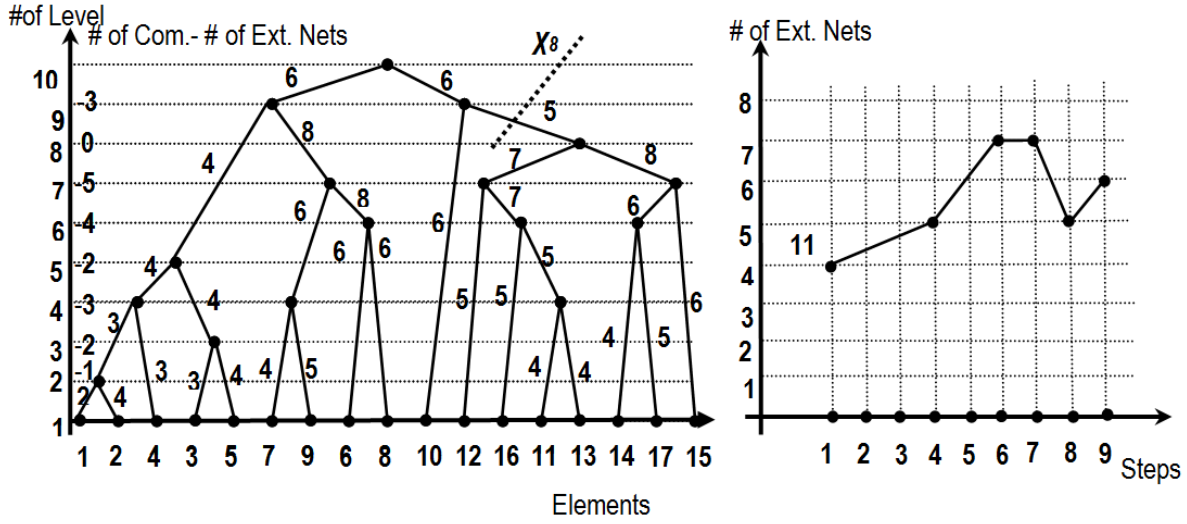


Fig. 4. Example of the Reduction Tree and the diagram dependency of the external nets clusters' number vs reduction steps

Partitioning

It is necessary to obtain the partitioning $P^* = \{P_1, \ldots, P_k\}$ for the set of elements $P = \{p_1, \ldots, p_n\}$ so that the quality function is optimized:

$$Q(P^*) \rightarrow opt\, Q\left(\widetilde{P}\right), \quad \widetilde{P} \in D,$$

while satisfying such or some other constraints:

$$(\forall P_i, P_j \in P^*)\ [|P_i| \approx |P_j|].$$

Set $\widetilde{P}$ is the arbitrary partitioning in the feasible region *D*, *k* – the number of partitions. The solution should also satisfy the following additional conditions:

$$(\forall P_i \in P^*)\,[P_i = \{p_{i1}, \ldots, p_{ini}\},\ p_{ij} \in P;\ i = 1, \ldots, k;\ j = 1, \ldots, n_i];$$

$$(\forall P_i \in P^*)(P_i \neq \varnothing);$$

$$(\forall (P_i, P_j) \in P^*)[P_i \cap P_j = \varnothing].$$

The n_i is the number of elements of *i-th* partition.

By the OCR method we recommend to solve partitioning problem in the two stages: initial partitioning and partitioning optimization.

Initial partitioning

Using the constructive method, it is desirable to find an initial solution at the first stage, which must be improved by the iterative method at the second stage. The important peculiarity of the approach developed is that it is recommended to use the hierarchical circuit clustering, obtained by the OCR method at the both stages. For initial partitioning it is possible to use following algorithms: serial, parallel-serial, and dichotomy.

By the serial algorithm on the Reduction Tree T^R the vertex is found, whose number of elements is equal to or greater than the desired value. If the number of elements is what we desire, we create the first partition and move forward to the next partitions. If this number is greater then desired, the problem is to remove the necessary number of elements. The problem recursively continues to final solution.

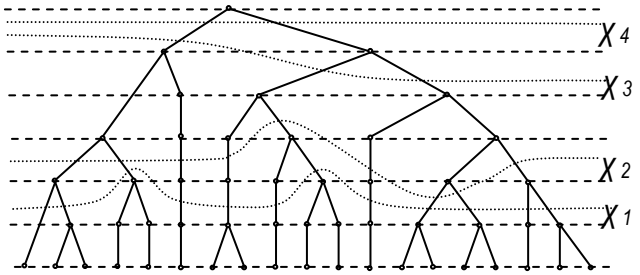


Fig. 5. Cuts in the Reduction Tree

Any cut of the Reduction Tree T^R by an arbitrary line forms subtrees (forest), the set of initial vertices (leaves) of each subtree can be considered as the set of elements of a single partition. For example, cut χ_1 at the Figure 5 can be used when it is necessary to split the circuits into the minimum number of partitions with a number of elements not greater than 2. Seven partitions are formed directly. The five elements remain ungrouped. For their assignment it is possible to construct the Reduction Tree in the subsequent repetition, and so on until the completion the problem. Cut χ_2 creates the three groups for three elements. To form the remaining partitions it is necessary to continue the process as in the previous case. Cut χ_3 gives good initial solution for three partitions (with 6, 5 and 8 elements). The first subcircuit can be directly incorporated into the solution as one partition, and then it would be necessary to transfer one element from the third to the second subcircuit. We obtain the partitions with 6, 6 and 7 elements. Cut χ_4 dived the circuit into two partitions with 6 and 13 elements. For receiving two approximately equal partitions we need to transfer three elements from larger partition into smaller one.

Dichotomy algorithm performs the top-down circuit division with constraint on the number of elements that should be equally divided to the desired number of elements at one partition. In the first step we consider the two highest vertices. This determines the number of possible partitions that can be formed from each vertex and the

number of elements in the remainders. The next step is to transfer the remaining elements from one piece to the other in the optimal way. The problem is reduced to the two new problems of the same type but of lesser size. In both cases, we use identical procedures to transfer the small number of elements from one piece to another, procedures which are performed recursively on the sets that decrease from the step to the step.

Partitioning optimization

For partitioning optimization at the first step we build the separate ORT for two partitions T^{R1}, T^{R2} and use the following procedures:

- *P1.* The exchange: arbitrary element from one partition and arbitrary element (cluster) of other partition.
- *P2*. The exchange: arbitrary clusters between two partitions.
- *P3*. The exchange: arbitrary sets of clusters between two partitions.
- *P4*. The transference: arbitrary element (cluster) from one partition to another.
- *P5*. The transference: arbitrary set of clusters and elements from one partition to another.

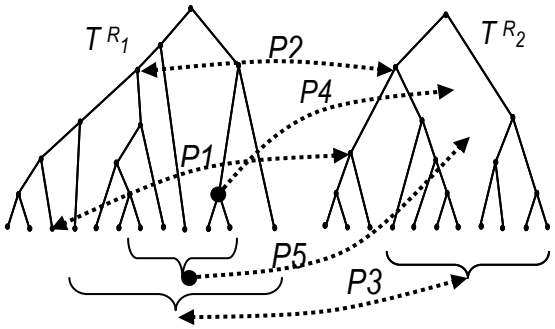


Fig. 6. Procedures for partitioning optimization

Some experimental bipartition results of test-case IBM01 [Alpert, 1998] by using our approach are shown at Table 1. For 100 randomly generated initial solutions we perform optimization. For escaping from the local extreme we use the perturbations by replacing clusters with the smallest value of the solution's worsening. The number of such perturbation is presented at the first column. For all initial solution we received the cut with180 nets that we think it is an optimal result (our conjecture), as was received from other investigators by using another approaches [Karypis, 1997], [Saab, 2000]. Forth column shows the number of the best solutions that were received for all randomly generated initial solutions. The fifth-eighth columns show the numbers of solutions that have 1, 2, 5 and 10 % deviation from the best solution. Last columns show the average solution, average number of iteration and average runtime.

Table 1. Partitioning results for IBM 01

# steps of perturbations	Maximal solution	Minimal solution	# of optimal solutions	Deviation from the best solution				Average solution	Average # of iterations	Average runtime, s
				<= 1 %	<= 2 %	<= 5 %	<= 10 %			
0	707	180	1	6	27	38	41	307	22	174
1	699	180	1	25	57	61	61	236	33	204
2	699	180	3	43	65	67	67	225	43	233
3	699	180	4	52	68	68	68	222	53	261
5	699	180	8	60	68	68	68	222	73	318

Packaging

It is necessary to obtain the partitioning $P^* = \{P_1, \ldots, P_k\}$ for the set of elements $P = \{p_1, \ldots, p_n\}$ so that the total number of partitions is minimized:

$$k \rightarrow min,$$

while satisfying the given constraints:

$$(\forall P_i \in P^*)\,[(n_i \le n_{i\,max}) \,\&\, (m_i^{ex} \le m_{i\,max}^{ex})].$$

Here n_i and m_i^{ex} are the numbers of elements and external nets (IO terminals) in each partition P_i that can not exceed the upper bounds $n_{i\,max}$ and $m_{i\,max}^{e}$. It should also satisfy the same additional constraints as for partitioning.

By the OCR method we also recommend to solve the packaging problem in two stages: initial packaging and packaging optimization.

Initial packaging

The algorithm begins to operate on the cluster of the Reduction Tree T^R, which appears first in violation of the constraint on the number of elements. From this cluster we form the first partition with as many as possible elements without violation on constraint on the number of external nets. Two strategies are used: to remove the minimal number of elements and to identify the best cluster without violation on constraints. The next step consists of the addition of the maximum number of elements. The experiments reveal the advantage of simultaneous combination of both strategies that perform iterative removal and addition of elements and clusters. The partitions separated first have a good density; but the final ones – bad. This is caused first of all by the "greedy" partitioning by serial strategy. As a result, the number of partitions can be greater than the optimal.

Table 2: Packaging Results for FPGAs

Circuit	# of CLBs	# of Nets	Numbers of FPGAs				
			[Kuznar, 1993]	[Nan-Chi Chou,1994	[Bazylevych, 2000]		Theoretic optimum
					Initial	Opt.	
a) with 64 CLB and 58 IO (Xilinx XC2064)							
C499	74	123	2	-	3	2	2
C1355	74	123	2	-	2	2	2
C1908	147	238	3	-	4	3	3
C2670	210	450	6	-	6	6	4
C3540	373	569	6	6	8	6	6
C5315	531	936	11	12	12	10	9
C7552	611	1057	11	11	12	10	10
C6288	833	1472	14	14	14	14	14
b) with 320 CLB and 144 IO (Xilinx XC3090)							
s15850	842	1265	4	3	4	3	3
s13207	915	1377	7	6	6	4	3
s38417	2221	3216	12	10	10	8	7
s38584	2904	3884	17	14	14	10	10

Packaging optimization

The partitions with the number of elements lesser than the constraint merge into one or several without violating it. The next step is the optimization on the set of all partitions that allows to increase the number of elements, but not to exceed of constraints on the first group of partitions, which were not subject for merging. Often such optimization substantially decreases the number of external nets of final partitions up to the desired value. If this is impossible to obtain, then the new final partition is divided into two smaller ones. The first partition should be without violations on constraints; the second may exhibit the violation on the number of external nets, if it is not possible to create it without violation, and so forth, up to the completion of the problem.

The experiments confirm the high efficiency of this approach on the set of some well-known test-cases. We merged only 20% of the better independent pairs at the every level of the Reduction Tree T^R generation. The test results (#FPGAs) are shown at the Tables 2. We used the 64 CLBs and 58 IOs constraints (FPGA Xilinx XC2064) for the tests with Table 2 and 320 CLBs and 144 IOs (FPGA Xilinx XC3090) for the tests with Table 3. As one could see from the tables the obtained results are not worse, and in the 5 cases from 12 they are the best among the known and are optimal. If our results are not being theoretically optimal, they are close to the optimal solutions and differ from them minimally, i.e. only by one partition (circuits c5315, s13207, and s38417) or two partitions (circuit c2670).

Floorplanning and placement

Combined hierarchical clustering and decomposition can be used for floor planning and placement. Such an approach is especially effective for large and very large-scale problems. The problem is solved in several stages:

- the bottom-up free hierarchical circuit clustering;
- the mathematical description of clusters by macro models at every level of decomposition;
- top-down multilevel placement with global and local optimization at every level of decomposition by using macromedels.

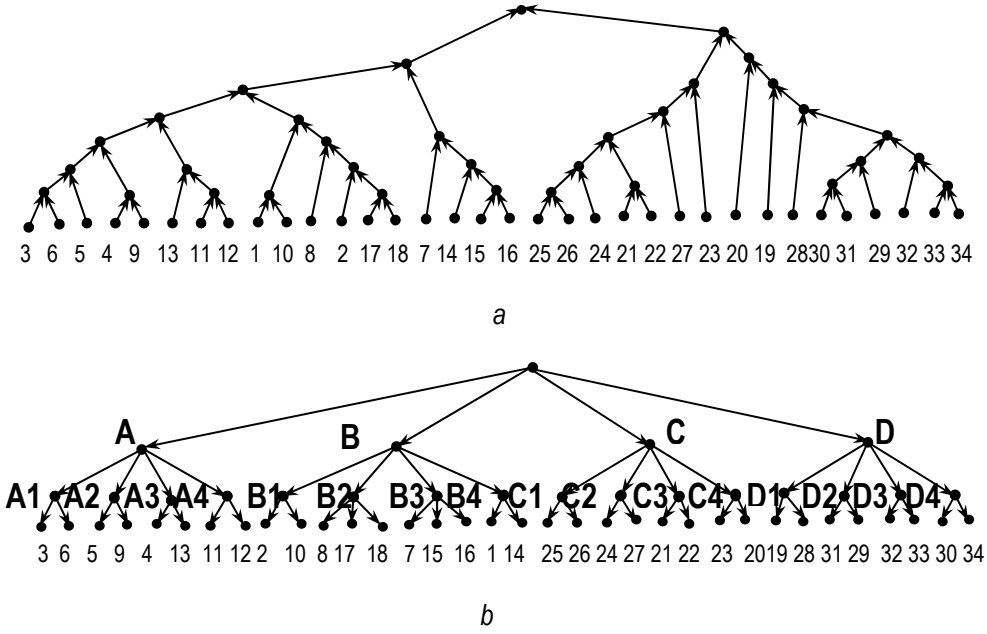


Fig. 7. Bottom-up hierarchical circuit clustering (a) and top-down 3-level of decomposition for Steinberg test-case

Figures 7 and 8 show the results of exploiting the developed approach for the Steinberg placement test-case [Steinberg, 1961]. For the first step (*a*) we build the ORT, for the next (*b*) – the tree level of decomposition (with 4, 16 macro models and 34 initial elements at the lowest level). At every level of decomposition we received some initial solution and performed it optimization using macro models (Figure 8 *a, b* and *c*) by Scanning-area method [Bazylevych, 1981, 1997]. We got the results of L_e = 4119,7 (the summary length of all connections with the Euclidian metric), which is the best comparatively with the other known solutions.

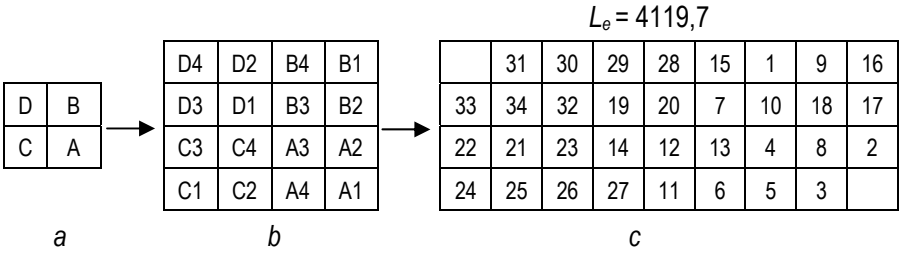


Fig. 8. Multilevel placement for Steinberg test-case

Conclusions

Hierarchical circuit clustering is a good precondition for solving the physical design problems of large and very large-scale electronic devices - VLSI, SOC and for PCB. For hierarchical clustering we developed the OCR method. Basic algorithms were proposed for partitioning, packaging, floorplanning and placement problems. They were used to obtain the initial solutions with not very large number of macromodels, as well as for their optimization. The proposed algorithms have some new properties, for example, they can be efficient in choosing the most appropriate number of partitions into which it is necessary to divide the circuit; arbitrary division coefficient can be chosen for partitioning; the same procedures can be used for initial solution and their optimization. The suggested algorithms have near linear computational complexity and provide good quality of results. For all test-cases investigated, the results are not worse, and in many cases they are better comparatively with obtained by other known methods. For some cases, the optimal results were received for the first time.

Bibliography

[Bazylevych, 1975] R.P.Bazylevych, S.P. Tkachenko. The partitioning problem solving by the parallel reduction method. In: Vychuslitelnaia technika: Materialy konferencii po razvitiiu technicheskich nauk po avtomatizirovannomu proektirovaniiu, V. 7, Kaunas, 1975, pp. 295-298 (In Russian).

[Bazylevych, 1981] R.P. Bazylevych. Decomposition and topological methods for electronic devices Physical Design Automation, Lviv: Vyshcha shkola, 1981, 168 P, (In Russian).

[Bazylevych, 1997] Roman P. Bazylevych, Taras M. Telyuk. VLSI and PCB elements placement optimizing using hierarchical scanning area method. 42 Internationales Wissenschaftliches Kolloquium. Technische Universitat Ilmenau. Ilmenau. 1997, pp. 594-599.

[Bazylevych, 2000] R.P. Bazylevych, R.A. Melnyk, O.G. Rybak. Circuit partitioning for FPGAs by the optimal circuit reduction method. In: VLSI Design, Vol. 11, No 3, pp. 237-248, 2000.

[Bazylevych, 2002] Bazylevych R.P. The optimal circuit reduction method as an effective tool to solve large and very large size intractable combinatorial VLSI physical design problems. In: 10-th NASA Symp. on VLSI Design, March 20-21, 2002, Albuquerque, NM, USA, pp. 6.1.1-6.1.14.

[Bazylevych, 2002] Bazylevych R. P., Podolskyy I.V., Bazylevych P.R. Hierarchical clustering – efficient tool to solve nonpolinomial combinatorial problems of high sizes. In: Shtuchnyy intelekt, Ukraine NAN, No. 3, 2002, pp. 447-483, (In Ukrainian).

[Bazylevych, 2007] R. Bazylevych, I. Podolskyy and L.Bazylevych. Partitioning optimization by recursive moves of hierarchically built clusters. In: Proc. of 2007 IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems. April, 2007, Krakow, Poland, pp. 235-238.

[Garbers, 1990] Jorn Garbers, Hans Jurgen Promel, Angelika Steger. Finding Clusters in VLSI Circuits. In: Proc. of IEEE/ACM Intern. Conf. on Computer-Aided Design, Santa Clara, 1990, pp. 520-523.

[Cong, 1993] Jason Cong and M'Lissa Smith. A Parallel Bottom-up Clustering Algorithm with Applications to Circuit Partitioning in VLSI Design. In: Proc. 30th ACM/IEEE DAC, 1993, pp. 755-760.

[Dutt, 1996] S.Dutt, W.Deng. VLSI circuit partitioning by cluster-removal using iterative improvement techniques. In: ICCAD'96, pp. 194-200.

[Karypis, 1997] G.Karypis, R.Aggarwal, V.Kumar, and S.Shekhar, Multilevel hypergraph partitioning: Application in VLSI domain. In: DAC 97, pp. 526-529.

[Saab, 2000] Yousseb Saab. A new multy-level partitioning algorithm. In: VLSI Design, Vol 11, No 3, pp. 301-310, 2000.

[Alpert, 1998] C.J. Alpert. The ISPD-98 Circuit Beanchmark Suit. In: Proc. ACM/IEEE Intern. Symposium on Physical Design, April 1998, pp. 80-85.

[Kuznar, 1993] Kuznar, R., Brglez, F. and Kozminski, K. Cost minimizationof partitions into multiple devices. Proc. Of IEEE/ACM 30th Design Automation Conference, 1993, pp. 315 - 320.

[Nan-Chi Chou, 1994] Nan-Chi Chou, Lung-Tien Liu, Chung-Kuan Cheng, Wei-Jin Dai and Rodney Lindelof. Circuit partitioning for huge logic emulation systems. Proc. 31 ACM/IEEE Design Automation Conference, 1994, pp. 244-249.

[Steinberg,1961] Steinberg L. The backboard wiring problem a placement algorithm. SIAM Review, 1961, v.3, ¹1, pp.37-50.

Authors' Information

Roman Bazylevych *– Prof. dr.hab., University of Information Technology and Management in Rzeszow; 1/42 Dovbusha str., Lviv, 79008, Ukraine;*

email: rbazylevych@wsiz.rzeszow.pl;

Major Fields of Scientific Research: combinatorial optimization, computer-aided design, algorithms

Marek Pałasiński *– Prof. nadzw. dr hab., Chair of Mathematics and Computer Science Foundations, University of Information Technology and Management in Rzeszow;*

email: mpalasinski@wsiz.rzeszow.pl;

Major Fields of Scientific Research: theoretical computer science, theory of algorithms, graph theory, data mining and algebraic logic

Lubov Bazylevych *– senior scientist, Institute of Mechanical and Mathematical Applied Problems of the Ukrainian National Academy of Sciences;*

email:lbaz@iapmm.lviv.ua;

Major Fields of Scientific Research: applied mathematics and mechanics, combinatorial optimization, computer science problems.

ARTIFICIAL INTELLIGENCE IN MONITORING SYSTEM

Lucjan Pelc, Artur Smaroń, Justyna Stasieńko

***Abstract**: The article presents the neural network constructed in order to use it for monitoring. Its role is to recognize the events in alarm situations (theft, burglary etc.). The film presenting a real break-in into the car was used while testing this network. The main task of monitoring system based on the neural network is to compile such a network which shows the alarm situations as soon as possible using the available equipment.*

***Keywords**: neutral network, monitoring system, neuron.*

***ACM Classification Keywords**: I.2 Artificial Intelligence – I.2.6. Learning*

Introduction

The monitoring programs are often used in order to gather the information about the amount and the quality of the observed object. The gathered information makes it easier to make right decisions, especially if this state is hazardous for human being and the surrounding. It allows also to improve or remove the results in the existing situation. The article aims at showing one of these solutions. There are many monitoring systems available on his market. They are different as far as the quality, functionality and price are concerned. If we take into consideration the functionality the following types can be distinguished: systems which record on the continuous basis, systems which record if the movement is detected and those which record before and after the incident. The current monitoring systems are able to: record on the continuous basis, detect the movement and record the image, detect the disappearance of the image, follow and count the objects, control the cameras, one or more visual channels, and inform the operator by an e-mail or sms. The cost of such systems amounts to 300 – 10000zl. Unfortunately, there are not many systems which can define the situation before the burglary or theft etc.

The majority of systems posses an insignificant or do not posses any functionality as far as the process of supporting the system operator is concerned especially in case of recognizing important facts and events. The monitoring systems are used rather for recording and gathering the courses of events, even undesirable ones. It requires large disk storage for recording an image. In some systems (e.g. Taiwanese ACTI- APP-2000-32) an intelligent management of memory was used by means of recording the events before they occur which is often called “buffering the image". If the system posses such ability, it buffers the image all the time and in the moment of detection the alarm it has also the images recorded before the accident, for example 5sec. Supporting the system operator while recognizing undesirable situations on the recorded image and responding automatically requires from the system to solve problems connected for example with pattern recognition. One of the authors of this article used probabilistic timed automata which if connected with particular spheres on the observed image allowed to define important actions in reference to the given problem [Pelc, 2008]. This article is a kind of a supplement and expansion of that approach. Instead of using probabilistic timed automata the neural network was used here.

While creating the monitoring system three rules should be taken into consideration: the periodicity of measurement, the unification of the equipment and methodology used for the measurement and observation as well as the unification of the results interpretation. This case study uses a film from the Internet showing the real break-in into the car [http://www.youtube.com/watch?v=pLKjm2uGrU4&feature=related]. It resulted

in the idea of creating the monitoring system based on one of the Artificial Intelligence applications "neural network"

The aim of article

The aim of this article is to create a low-cost monitoring system which can be compiled on the basis of computers designed for the general use as well as typical cameras (e.g. network cameras)

The system assumptions:

- one or two cameras
- limited computability of the equipment and disk
- better functionality than in typical monitoring systems
- algorithm of artificial intelligence – the neural network whose computability is not very complex
- the ability to recognize particular actions and reactions depending on their character:
  - ✓ neutral actions – normal work of the system
  - ✓ suspicious actions – generating the warnings by the system, starting recording the sequence of frames
  - ✓ prohibited actions – alarming, calling the operator and recording the sequence of frames

The main feature which distinguishes the monitoring system from many others is its ability to take decisions in order to prevent the prohibited actions and not only to record them. The following method of solving the problem was accepted:

- action was defined as the occurrence of the given movement trajectory in particular period of time
- the movement trajectory is determined on the basis of recognition of the object being observed in the successive spheres of observation
- resignation of the constant recording of the images in aid of the recording which is caused by events
- the event causing the recording is the occurrence of the beginning of the action
- introducing the spheres of observation on the image recorded by the camera limiting at the same time the size of the in-put data.

The method of defining the spheres of observation

The image coming from the camera is covered by the network of spheres of observation consisting of 125 elements. The size of the network is determined automatically depending on the picture definition. The network density can change adjusting the image from the camera. The network is related to the scene or objects being observed whose position may change. However, the main stress is not put here on fluent change of the position and following the object.

The example of using the network related to the scene can be the observation of the gateway. However, in case of the network connected with the object, it can be presented by means of observation of a car parked in different ways and positions. To sum up, the steps of defining the spheres are as follows:

(1) determining if the spheres should take the form of a uniform or condensed network
(2) choosing the object or the scene
(3) indicating the essential points of the image for which the system should condense the network
(4) if needed, defining the size of grid's fields which is different than the implicit one (implicitly there are three types of the size and density of grid's fields: large, medium and small)

(5) in case of the network related to the particular object there is a need of creating the definition of this object. The system defines the object on the basis of its simplified shape, colour or its texture.

(6) If the system is going to be used in applications whose shape does not change or the changes are small than in (5) the size of the object is also taken into consideration. It is worth remembering that in the systems with cameras there is an apparent change of the object size depending on its distance from the camera.

Fig.1 shows the example of the image seen by means of a camera with a network characterized by reduced density. The biggest density is intended for the observed vehicles and it is the situation when the network is related to the observed objects.

Fig 1. An example of the definition of the sphere of observation. Source: Own elaboration.

The logic of the created system

As it has been mentioned earlier, the system should inform about the potential risk of the occurrence of the prohibited action, which as a result, allows to take remedial. In consequence, the logic of the system should provide the proper sensitivity as far as the recognition of the action is concerned and it should also be characterized by the ability of limited prediction. It should be remembered that one of the assumptions at the beginning of this article was the system's ability to recognize the neutral, suspicious and prohibited actions. Detection of the prohibited action is connected with alarming and calling the operator and that is why the system cannot do it precipitously. In other words, the system cannot be too sensitive and its reactions should not be exaggerated. It is especially important in case of determining the prediction. Taking into consideration all of these assumptions, the logic of the system should demand the occurrence of one of these three sorts of action on the basis of the information in the defined spheres of observation. The classic neural networks were used to support this solution. It was necessary to decide about the number of layers in the network. It was assumed that the exit layer will consist of at least three neurons indicating the occurrence of a particular kind of action. In the entrance layer, there should be as many neutrons as the spheres of observation, however, this solution would be enough only in case of recognizing the statistical actions. In the presented example the time lag also should be mentioned. Time is measured with the occurrence of successive frames. That is why, it is necessary to add one more neuron informing about the number of frames which have just appeared. It allows for taking into

consideration their movement dynamics, and not only the statistical trajectory. Therefore, it can be assumed that in the "minimal" version the number of neurons in the entrance layer amounts to the number of spheres of observation enlarged by one.

Classic neural network consists of one or few hidden layers. In the experiment the number of the hidden layers and the number of the neurons in each of them were the object of the simulation analysis whose results will be soon discussed in this article. Taking into consideration the previous assumptions referring to computability, two and more hidden layers were analyzed.

The simulation

The aim of the simulation was to determine the construction of neural network which is the most suitable from the viewpoint of the given problem. Such a network should for example:

- recognize defined actions accurately
- predict the occurrence probability almost without any mistakes
- include the least neurons
- learn quickly

According to the assumptions accepted previously the network with one and two hidden layers were analyzed paying particular attention to the accuracy of recognizing the actions as well as fast learning.

Fig.2, Fig.3, and Fig.4 present hypothetical and simplified movement trajectories referring to three types of actions discussed previously. Fig.2 deals with the forbidden action, Fig.3 is related to the suspicious action and Fig.4 presents the neutral action. The protected sphere is marked with "x" and darkened. The black squares represent the recorded trajectory of movement. The number under each of the pictures informs about the number of frames falling on the particular trajectory.

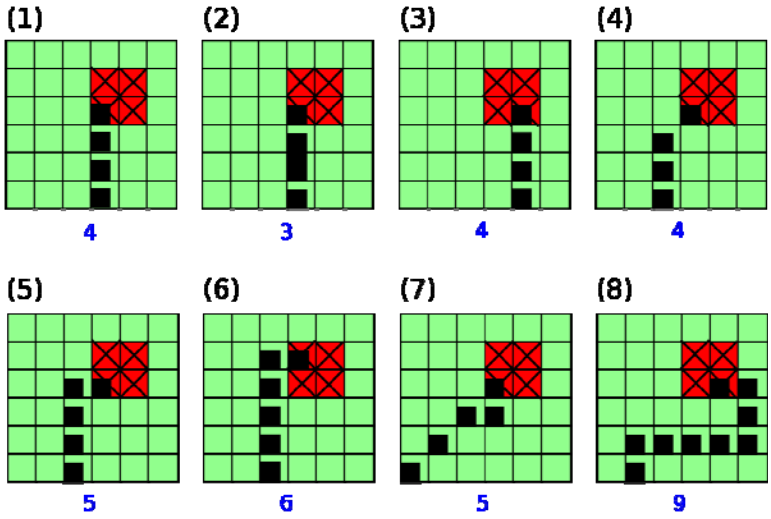


Fig. 2. Examples of movement trajectory for the forbidden-alarm action. Source: Own elaboration.

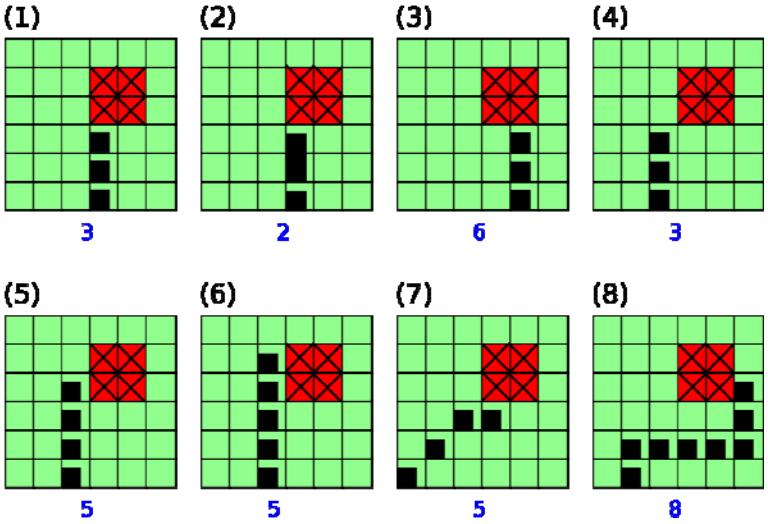


Fig. 3. Examples of movement trajectory for the suspicious-warning action. Source: Own elaboration.

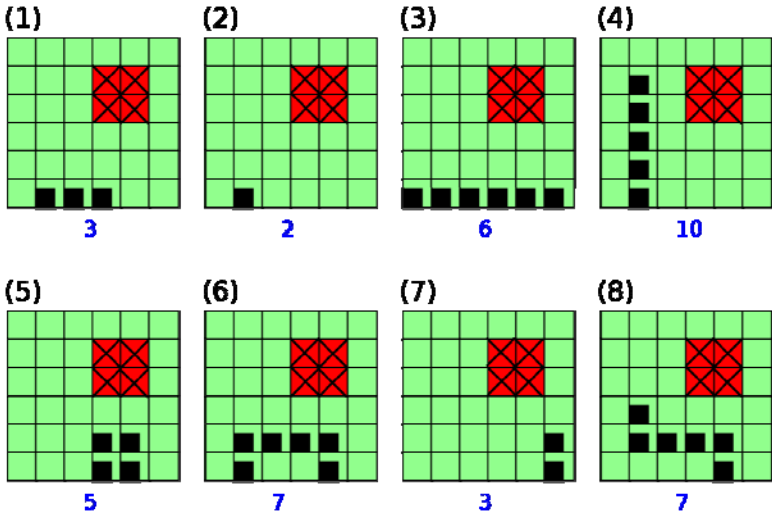


Fig.4. Examples of movement trajectory for the neutral-normal action. Source: Own elaboration.

The elementary picture (1) presented by Fig.2 represents the trajectory going in four steps from the bottom part directly to the marked sphere. Below there are some situations, which could occur for the discussed example of the scene:

- The prohibited action – if the trajectory ends or goes through the marked sphere (Fig.2)
- The prohibited action which is very probable – if the trajectory with great dynamics goes towards the marked sphere (compare the elements (1) and (3) from Fig.2)
- The suspicious action – when the trajectory goes towards the marked sphere but does not reach it, the movement dynamics is not important here
- The suspicious action which is very probable - when the trajectory is near the marked sphere regardless of the dynamics (compare elements (5), (6), (7), (8) from Fig.3)
- The neutral action – the movement trajectory is in a short distance away from the marked sphere regardless the dynamics (Fig.4).

Even after analysing Fig.2 roughly it can be seen that if the object moves quickly on the trajectory with the marked sphere, the system should inform as quickly as possible the occurrence of the prohibited action before

the trajectory reaches the marked sphere. In practice, it would allow to prevent the prohibited action to reach the marked sphere in the given example. The permanent persistence of the suspicious action may indicate the potential occurrence of the prohibited action. The situations presented on Fig.3 and Fig.4 were chosen in such a way that the situations which at the beginning looked like forbidden after the occurrence of the successive frames turned out to be suspicious for example Fig.3 (1) and Fig.2 (1), Fig.3 (2) and Fig.2 (2) etc. Taking into consideration the assumptions that the system should be characterized by the ability to predict the situation, but it should not react precipitously, such kind of situations are interesting.

The simulation results

During the simulation it was checked how quickly the network with one or two hidden layers is able to learn as well as the accuracy of recognition if the action is forbidden, suspicious or neutral.

- **The comparison of the capacity of learning for the network with one or two hidden layers**

Below, there are the tests results of the capacity of learning for the network with one hidden layer depending on the number of the neurons in the hidden layer. Fig.5 presents the reduction of mistakes in the process of learning depending on the number of epochs. Other progresses refer to various numbers of neurons in the hidden layer. It is important to emphasize that at the beginning the number of neurons in the hidden layer improves the parameters of learning but after exceeding the certain number of the neurons the situation changes the other way round. On the basis of the conducted analysis the formula was defined. It counts the number of neurons in the single hidden layer as a quotient of the product and the sum of the number of neurons in the input and output layer in the network (1).

$$N_h \approx \frac{N_{in} * N_{out}}{N_{in} + N_{out}} \tag{1}$$

where

N_{in} – the number of neurons in the entrance layer

N_{out} – the number of neurons in the exit layer

N_h – the number of neurons in the hidden layer.

If the number of neutrons in the entrance layer is much bigger than the number of neurons in the exit layer (at least of an order of magnitude), the radical dependence, which combines the number of neurons of the hidden layer with these from the entrance and exit layer (2), seems to be more accurate.

$$N_h \approx 1.2 * \sqrt{N_{in} * N_{out}} \tag{2}$$

The analogous simulations were carried out with two hidden layers. The number of the neurons in the first hidden layer is chosen on the basis of relations resulting from formula (2).

The comparison of progresses for various number of neurons in particular hidden layers presented by means of Fig.6 indicates the validity of accepting the correlation that the number of neurons in the second layer should be smaller by half than the number of neurons in the first layer (according to the rule of pyramids) (3).

$$N_{h2} \approx 0.5 * N_{h1} \tag{3}$$

Having analysed the networks which learned faster with one or two hidden layers it was included that the network with only one layer is able to learn faster.

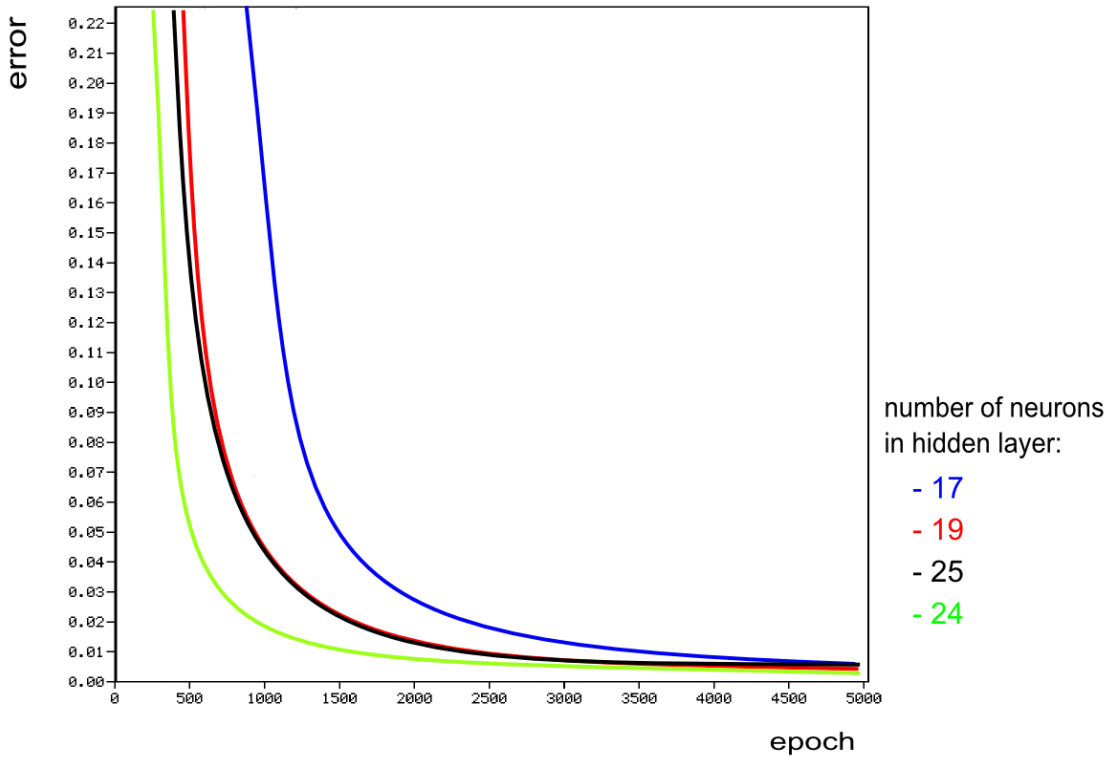


Fig.5. The capacity of learning depending on the number of epochs. Source: Own elaboration

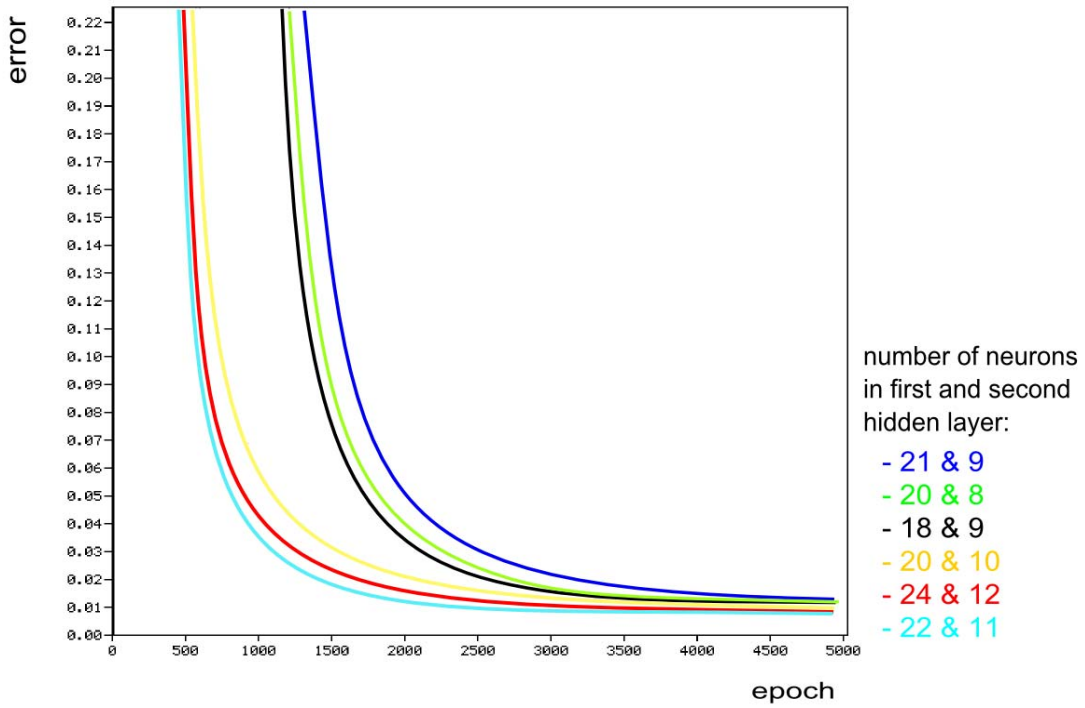


Fig.6. The capacity of learning depending on the number of neurons for two hidden layers. Source: Own elaboration.

- **The comparison of the accuracy in actions recognition**

It turned out that the accuracy of the recognition of the action is bigger for the network with two hidden layers than the network with just one hidden layer. However, in this case the recognition of the forbidden action during the trajectory analysis was precipitate. In consequence, it was the trajectory mostly connected with the suspicious

action. That is why, at the dynamic recognition of the action or rather the prediction of the action, the network with one layer turned out to be more accurate solution. It is illustrated by Table 1 which shows how the neuron network recognized the forbidden-alarm action, and suspicious-warning action in the particular period of time, described in Table as steps. Analyzing for example case, it is easy to notice that the network with two hidden layers reacts too rapidly and notices in the developing warning situation the alarm situation. It can be easily seen in steps 4 and 5 where the network guesses the alarm situation wrongly with the probability bigger than 99%. In the successive steps the network recognizes the warning correctly, but taking into consideration the previous assumptions such behavior of the network is delayed. For the same steps (4 and 5) the network with one layer recognized the alarm with the probability of 1,5% and the proper warning situation with more than 92%. Although the network with one hidden layer recognizes the finished actions less accurately than the network with two hidden layers, for the investigated application the network with one hidden layer is more suitable.

Table 1 Example situations recorded by monitoring system

Example situations recorded by monitoring system		Response of the system in percentage					
		One hidden layer – 19 neurons			Two hidden layers – 22 and 11 neurons		
		alarm	warning	neutral	alarm	warning	neutral
Case 1: warning	Step 1	1,72%	0,19%	98,09%	28,06%	0,08%	71,86%
	Step 2	0,82%	0,41%	98,76%	22,72%	0,08%	77,20%
	Step 3	2,85%	1,93%	95,22%	39,02%	0,08%	60,90%
	Step 4	1,55%	92,43%	6,02%	99,13%	0,79%	0,08%
	Step 5	1,49%	94,62%	3,89%	99,12%	0,80%	0,08%
	Step 6	0,67%	98,83%	0,50%	1,77%	98,23%	0,00%
	Step 7	0,60%	99,03%	0,37%	0,58%	99,42%	0,00%
	Step 8	0,55%	99,08%	0,36%	0,48%	99,52%	0,00%
	Step 9	0,73%	98,97%	0,30%	0,34%	99,66%	0,00%
Case 2: alarm	Step 1	6,73%	0,00%	93,26%	5,16%	0,09%	94,76%
	Step 2	1,61%	0,24%	98,15%	56,15%	0,08%	43,77%
	Step 3	0,87%	0,95%	98,18%	73,90%	0,10%	26,00%
	Step 4	1,44%	7,81%	90,74%	94,39%	0,15%	5,46%
	Step 5	97,20%	0,64%	2,16%	98,70%	0,26%	1,04%
	Step 6	76,49%	23,40%	0,10%	2,76%	97,24%	0,00%
	Step 7	98,65%	1,28%	0,07%	12,84%	87,16%	0,00%

Source: Own elaboration.

The applied algorithm of learning

In the elaborated program was used the algorithm of error backpropagation based on the example found in the literature (4).

$$\delta_{i,j} = O_{i,j}*(A_j - O_{i,j})*(1 - O_{i,j})$$

$$\delta^{t}_{i,j} = \delta^{t-1}_{i,j} + O_{i,j}*(1 - O_{i,j})*\delta^{t-1}_{i+1,k}*Wi_{+1,k,j}$$

$$W^{\pi}_{i,j,k} = W^{\pi-1}_{i,j,k} + \eta*\delta_{i,j}*O_{i-1,k} + \alpha*(W^{\pi-1}_{i,j,k} - W^{\pi-2}_{i,j,k})$$

$$O_{i,j} = \frac{1}{1+e^{\beta(-(I_{i,j}+bias_{i,j}))}} \qquad (4)$$

where

i,j,k	– index of: layer, neuron and weight
t	– point of time
α,β,η	– momentum, beta and learning rate
I,O,A	– input, output and expected value
δ	– error
W	– weight

Paying attention to the given problem the adjustment of the algorithm and coefficients was conducted. The results of the adjustment of coefficients α and η is shown by Fig.7. The most proper choice turned out to be α = 0.6 and η = 0.2.

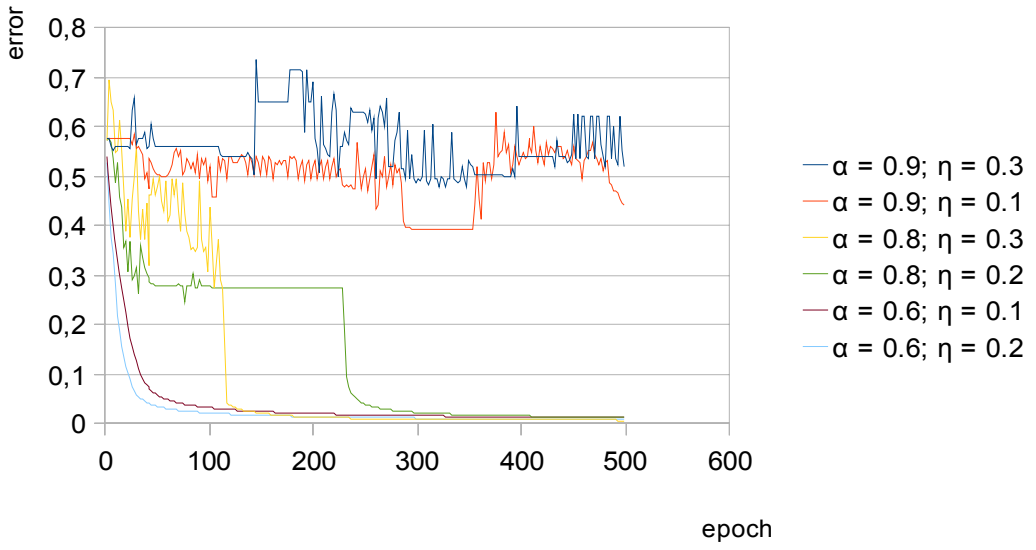


Fig.7. The selection of α and η. Source: Own elaboration

After choosing α and η it is possible to estimate the coefficient β. Fig.8 shows the results of experiments for various β coefficients.

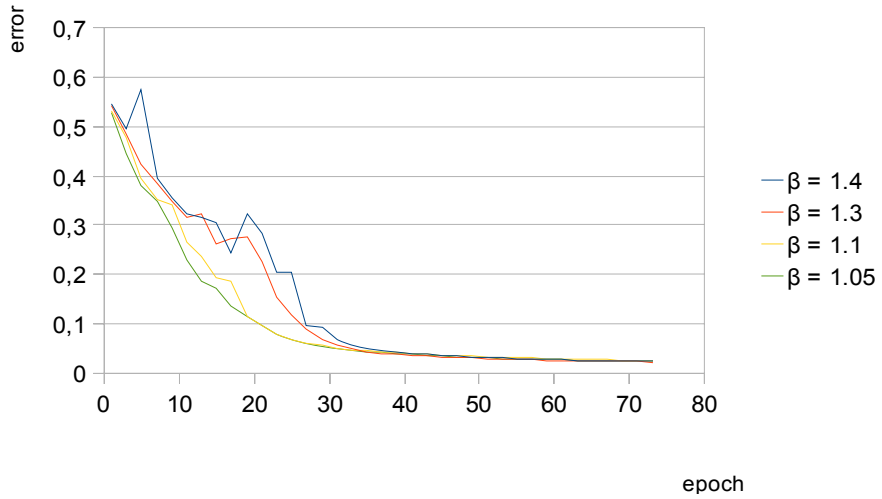


Fig.8. The selection of β. Source: Own elaboration

Analysing the mistake dependence upon the number of epochs it can be noticed from Fig.8 that the most satisfactory results occur for β which is close to 1.

The elaborated program is equipped with the mechanism of measuring out the time needed for learning and recognizing. According to the assumptions from the beginning of this project the computers intended for the general usage were used with Windows 7 Home Premium 64 bit, Dual Core AMD Athlon 64X2 L310 processor and 1.2 GHz RAM 2 GB memory. Satisfactory results of learning were achieved after less than 500 epochs and the time of learning was 5 seconds. The time of recognition fluctuated in single milliseconds. Those two results confirmed the validity of the proposed attitude on the computer equipment intended for the general use which is not very expensive.

The conducted experiments

The given monitoring conception of the system was implemented according to the policy and used for the image analysis coming from the real monitoring. The image referred to the recorded burglaries into the parked cars and thefts (film). Fig.9a) presents the network connected with the real car (first from the left) whereas Fig.9b) shows the network connected with the silver car- on the right. The conception accepted here related to the network with the variable density.

a) b)

Fig.9. Image from the camera with the network put by the monitoring system. Source: Own elaboration

Two examples of the neuron network with one hidden layer were considered. The number of neurons in the hidden layer was established according to pattern (2) while in the adjustment algorithm the accepted coefficients were α = 0.6, η = 0.2 and β =1.05. The network was taught on the basis of the patterns worked out according to the human intuition patterning on the trajectories analysed on Fig.2, Fig.3 and Fig.4. The expert's knowledge was used for working out these patterns.

The practical method of adjusting the system

The simulation analysis and the experiments conducted in real situations let to formulate some practical rules of configuring of the elaborated system fir other applications.

- **Determining the network**

  The network can be chosen as homogeneous or with changeable density. Then, the system operator has to indicate the crucial sphere in the image seen by the camera and establish if the network should be connected with the scene being observed or if it should be the object in that scene.

- **Teaching the network**

  In the situation where the current application is very similar to the previous one, the predefined network without teaching can be used. For example, if take into consideration the car park similar to the one presented by Fig.9 etc. If the current situation is totally different from the ones available in the previous study, the proper patterns should be prepared.

  It is important that they can be prepared in two ways:

  1) Recording the mock situation or
  2) Determine on the basis of the expert knowledge the trajectories by indicating the proper spheres of the network.

Generally, there are even several dozen and that is why it is the task available for the human perception and possible to be released in short time. The system operator has to prepare the trajectories analogous for those presented by Fig.10 b) and c) in the form of the marked spheres of the network.

a)

b)

c)

Fig 10. Course of events in the burglary: a) The network put by the system, b) Suspicious action from the point of view of the in-put neurons, c) The prohibited action. Source: Own elaboration

Conclusions

The presented monitoring system fulfils all the aims established at the beginning of this article which referred to creating low-cost monitoring system which can be combined on the basis of computers intended for general use as well as typical cameras such as network cameras. In addition, the main characteristic feature distinguishing this particular monitoring system from others available on the market was discovered. This feature refers to the

ability of take to prevent the prohibited actions and not only to record them. The use of typical neuron network with the standard configuration allowed to receive satisfactory results.

Bibliography

[Pelc, 2008] L.Pelc, B.Kwolek. Activity Recognition Using Probabilistic Timed Automata. In Peng-Yeng Yin: Pattern Recognition. Techniques, Technology and Application. In-Tech, November 2008, pp. 345-360.

[Rutkowski, 2009] L.Rutkowski. Metodyi techniki sztucznej inteligencji. PWN, Warszawa, 2009.

[Tadeusiewicz, 1993] R.Tadeusiewicz. Sieci Neuronowe. Wydawnictwo RM, Warszawa, 1993.

[Wszołek, 2006] W.Wszołek. Metody sztucznej inteligencji w systemach monitoringu hałasu lotniczego. Zeszyt 3/2006 *patenty.bg.agh.edu.pl/graf/Newsletter_AGH_Czerwiec_2011.pdf* [06.2011]

[Zurada, 1996] J.Żurada, M.Barski, W.Jędruch. Sztuczne sieci neuronowe. PWN, Warszawa, 1996

Netography:

http://www.youtube.com/watch?v=pLKjm2uGrU4&feature=related

http://www.ai.c-labtech.net/sn

http://kik.pcz.pl/nn/index.php

http://www.neuron.kylos.pl/pliki/start.html

http://www.willamette.edu/~gorr/classes/cs449/Backprop/backprop.html

Authors' Information

Lucjan Pelc *– lecturer, Rzeszow University of Technology; W. Pola Street 2, 35-959 Rzeszów and The Institute of Technical Engineering, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail: lpelc@prz-rzeszow.pl*

Major Fields of Scientific Research: Real'time protocols and systems, Formal description techniques, industrial vision systems

Artur Smaroń *– computer science's student, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail: arturo.s@op.pl*

Justyna Stasieńko – *lecturer, The Institute of Technical Engineering, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail:* justyna.stasienko@pwste.edu.pl

Major Fields of Scientific Research: Management Information Systems, Business information technology

DATABASE AND KNOWLEDGE BASE AS INTEGRAL PART OF THE INTELLIGENT DECISION SUPPORT SYSTEM, CREATED FOR MANUFACTURING COMPANIES

Monika Piróg-Mazur, Galina Setlak

Abstract: *The paper presents the structure of a database and a knowledge base which are the integral part of the Intelligent Decision Support System, being developed for a manufacturing company operating in the glass industry. Both modules will be used in the advisory system, whose functions will be classification of defects of products (here: glass packaging, e.g. bottles, jars) and selection of an appropriate (the most beneficial) method of elimination of defects arising in the manufacturing process.*

Keywords*: intelligent decision support systems, knowledge base, knowledge representation, reasoning process.*

ACM Classification Keywords*: I. Computing Methodologies, I.2.1 Applications and Expert Systems, J. Computer Applications,*

Introduction

In a manufacturing company making decisions in the scope of production preparation processes is a basic and key element of the whole manufacturing process.

It is essential for manufacturing companies to obtain information necessary to make an appropriate decision as soon as possible. They employ a lot of experts and specialists to react appropriately at every moment and to make suitable decisions. Decision-making processes are also supported by integrated computer systems, which continuously gather data and analyse different areas of the manufacturing process.

Decision support systems using artificial intelligence techniques are the systems which combine the potential for gathering and processing enormous amounts of data, using increasingly diversified models and intelligent utilizing data and knowledge gathered. The main purpose of development of an intelligent decision support system is to reflect experts' knowledge and experience, which are indispensable for solving problems by the system. Integration of intelligent methods allows to create better and more precise methods which can be applied in this field. In intelligent systems reasoning plays the most important role. The indicator of a system's "intelligence" is the ability to make decisions (through the reasoning process) and the ability to learn and acquire knowledge [Rojek, 2010].

Intelligent Decision Support Systems (IDSS) are the systems which have the potential for gathering and processing huge amounts of decision-making information, conducting analyses of the information and using diversified models, data and knowledge gathered to solve complex decision-making problems. The essential parts of an intelligent decision support system are a database and a knowledge base [Zieliński, 2000].

A database is one of the most important sources of decision-making information for a knowledge base in an IDSS. A database plays a basic role when developing a knowledge base in order to support the technological process, including the quality control process. The purpose of this paper is to present the issues concerning the development of these two component parts of the advisory system being developed.

Characteristics of the selected object of research.

The design of the intelligent decision support system, which is presented in this paper, is being realized for the Glassworks, a company operating in the sector of large companies. In total, the glassworks has 14 production lines, which work in a three-shift system; the capacity of one production line per one shift is 200,000 items of

finished product. To illustrate it better – an automatic machine works with the speed of 275 drops (gobs of molten glass) per minute.

The concept to develop an intelligent decision support system arose after analysing the literature on the subject and numerous visits in the Glassworks. It was found that there was no algorithm to be applied in the case of discovering a defect of a product (here: glass packaging, e.g. bottles, jars) and for selection of an appropriate (the most beneficial) method of elimination of defects which occur during the production process.

The intelligent decision support system being designed should allow to classify product defects and to select an appropriate (the most beneficial) method for their elimination. The system being developed should support a line operator and a production line manager to a degree which is comparable with the support provided by a specialist (an expert) with high qualifications. Effects of the operation of the advisory system operation will allow operators and line managers to make appropriate decisions to eliminate production defects, and in consequence, to improve the technological process [Piróg-Mazur, 2010].

Before setting about working on the expert system being discussed initial assumptions and a method of the system development have been defined:

- the system should suggest solutions within the defined range – supporting a user in solving decision-making problems in the process of finished product quality control, i.e. classification of defects of products (bottles) as well as analysis and selection of an appropriate method of defect elimination, which will also allow to improve the technological process,
- the system should be user-friendly, a user is not required to be an expert in the field, the user interface will be based on questions and answers in the natural language,
- the system should provide texts, drawings and possibly simulations - databases in the form of text and graphic files, which contain additional or more complete information,
- the system can be developed in any programming language.

The technological process in the Glassworks precisely defines the process of converting a raw material (semi-finished product) into a finished product, which is compliant with requirements specified in a project. Development of technological processes is a very important phase of production preparation. However, its automation is very difficult due to large contribution of the experience of process engineers to the designing process. Traditional designing of technological processes is dominated by the activities which to a large degree are based on the experience, skills and intuition of a process engineer. Technological processes and their costs are dependent on a process engineer's experience.

When designing a technological process information from different sources is used. Technological processes are influenced by different kinds of information and limitations: information on a product, limitations related to technological capabilities of a manufacturing company and output, requirements concerning a product manufacturing, competences of a process engineer (professional experience, creativity), methods and resources used in technological planning and data gathered previously (technological databases and knowledge bases) [Rojek, 2010].

According to the definition, a technological process is a quantitatively and qualitatively structured set of actions that change physical properties (shape, size), the form or chemical properties of a specific substance (material). Technological process, together with support actions (transfer of material), constitutes a production process that results in a final product. The process of glass production comprises 9 main actions, connected with transforming raw materials and materials into ready products (for the purpose of an external recipient).

- Glass batch preparation – accurately weighed out and mixed raw materials constitute the so called batch. Glass cullet is a very important raw material. Even 80% of natural resources can be replaced with it.

- Melting – the batch is transported to the foundry furnace, namely the glass melting furnace and melts in the temperature of 1500*C. Such a high temperature is provided by gas fire burners, situated on both sides of the glass melting furnace. Molten glass is pushed by a new portion of batch.
- Forming process – a stream of molten glass is cut into sections – drops of glass (known as gobs) of a weight that corresponds to the weight of the container being formed. Gobs are transferred to forming machines. The compressed air shapes, in the initial phase, a glass bubble that falls into the moulds where it takes on its final shape.
- Hot refinement – bottles or jars are transferred to a tunnel-chamber where the compound of tin is sprayed. It penetrates the glass surface and results in higher mechanical resistance of wares and gives them shine.
- Tempering – Wares are moved slowly on a conveyor belt inside a tunnel-lehr and solidify under control. It protects bottles against future cracking.
- Cold refinement – cooled glass wares are subject to a process that makes them still more shining and flexible.
- Quality check and sorting – it is automatically checked whether wares have flaws. If yes, the machine immediately eliminates defective wares.
- Packaging an lamination – once the wares have bee checked, they are conveyed to an automatic machine – palletizes, which arranges them in layers on pallets and protects them with a heat-shrinkable film. Packed wares are transferred to warehouses.
- Storage and shipping – prepared to client's order wares wait in the warehouse to be shipped [Stowarzyszenie Opakowań Szklanych, 2010].

In the case of such an extended production process the probability of occurrence of defects is very high. Optimization of the technological process is essential for the company.

Characteristics of database

A database development process comes down to defining objects in individual objects/tables and their attributes.

When designing the database the following questions have been asked:

- what data are we interested in?
- what format will they have?
- how are they related to each other?

Large production lines, which consist of several dozen machines linked to each other, have measurement points. Currently, in the glassworks data from measurement points are collected by PIC - Production Information Computer. Controlling the software (setting parameters which should be checked: glass wall thickness, profiled body, bottle neck, bottom, setting sensitivity – permissible norm and ideal norm) is indispensable for maintaining desired parameter values on a constant level. These parameters are adjusted every time when a product range (a product) is being changed. Control and measurement apparatus adjusts sensitivity. Sensitivity to critical defects is set to 100%. The higher sensitivity (expressed numbers) the larger number of rejects.

PIC software provides the following information:

- summary of losses on a specific production line,
- summary of losses in the whole glassworks,
- summary of rejects on a measurement point (cold end),
- losses on a selected production line,

- losses on a selected production line (detailed report),
- rejects per specific defects (percentage value),
- rejects per specific defects (number of items),
- machine downtime report,
- summary of results on all production lines,
- switching to another production line.

Dynamic data are related to the defects being monitored - a number of defects in time intervals: in 10 min., in an hour, in one shift and in 24 hours. Table 1 contains real data from one production line. There are 5 measurement points (FP1, ..., FP5) situated along the line, which record quality and quantity of defect occurrences. The table presents numbers of defects, their abbreviated names, percentage values in individual measurement points (FP) and their totals.

Table 1 . Data extracted from measurement points one production line. Source:: System PIC

DETECTOR ID	FP1 %	FP2 %	FP3 %	FP4 %	FP5 %	TOTAL %
L1 SPEK.101	0.05	0.11	0.08	0.19	0.16	0.11
L2 SPEK.101	0.09	0.05	0.19	0.35	0.00	0.15
L3 SPEK.119	0.00	0.00	0.00	0.00	0.00	0.00
L4 PECH. W GL.	0.34	0.43	0.24	0.44	0.92	0.39
L5 SPEK.102	0.03	0.01	0.03	0.30	0.05	0.08
L6 SPEK.121	0.40	0.25	0.59	0.27	0.48	0.39
L7 NIEROWNLOGLY	0.83	0.56	0.38	0.34	0.98	0.56
L9 KRZYWY	0.26	0.20	0.44	0.37	0.59	0.33
L10 OOR-CMG	0.02	0.15	0.02	0.14	0.04	0.07
L11 SPEK.119	0.43	0.25	0.15	0.56	0.17	0.32
L12 SWA	0.57	0.48	0.43	0.40	0.58	0.48
L13 OOR-IPS	0.01	0.00	0.00	0.00	0.00	0.00
L14 SPEK.DNO	0.01	0.05	0.20	0.16	0.03	0.10
L16 FTA	1.81	1.98	1.72	1.94	1.71	1.85
L17 CID	0.45	0.52	0.36	0.36	0.67	0.44
L18 ROZDMUCHANA	0.00	0.00	0.01	0.00	0.00	0.00
L19 CIENKI GORA	0.09	0.14	0.08	0.03	0.27	0.10
L20 CIENKI DOL	1.26	0.89	0.95	0.57	0.82	0.93
L21 SSG1	0.00	0.00	0.00	0.00	0.00	0.00
L22 SSG2	0.00	0.00	0.00	0.00	0.00	0.00
L23 SSG3	0.00	0.00	0.00	0.00	0.00	0.00
L24 BHA	0.73	0.58	0.61	0.97	0.73	0.71
% REJECTED	6.22	5.58	5.42	5.86	6.84	5.82
INSPECTED	152638	144079	150891	115036	30026	592670

One of the first phases of the database designing process is development of a conceptual data model, which is of key importance for usefulness and quality of a database being designed. It is created independently of solutions characteristic for any logical models and database management systems. The conceptual model will allow to present the technological process described above in a formalized way. The main purpose of database conceptual modelling is to create a design which reflects the fragment of reality being analyzed, which is free of details that would locate it among models of a specific class (object, relational or others) and which is independent of a programming platform. The final effect of the conceptual designing process is a design containing three kinds of elements [Put, 2009]:

- facts, i.e. objects and events, which are to be stored in a database,
- attributes which describe individual facts,
- types of relationships between facts.

A design is typically presented graphically in a form of an entity-relationship diagram, which is supplemented with a detailed text description of information which it contains. In the diagram, facts are usually denoted with rectangles, attributes are denoted with ellipses and relationships between facts are denoted with lines linking rectangles and with symbols near lines which describe a type of relationship (Fig. 1).

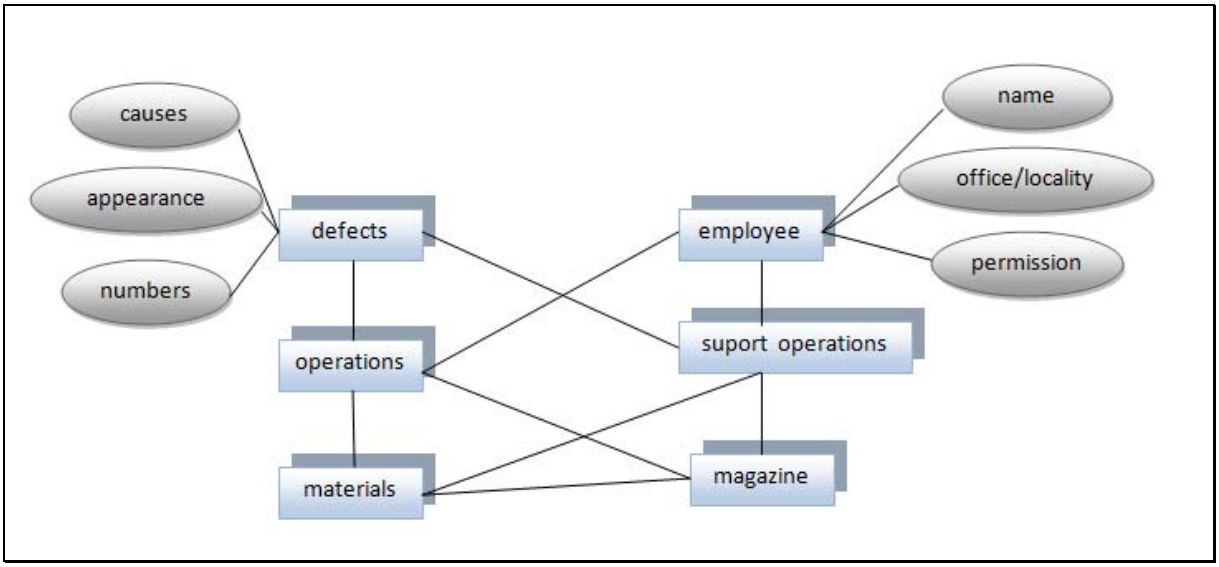


Fig. 1 Example of conceptual design - *graphic form (fragment/episode). Source: own work*

Selection of the relational model as a data storage method means that it is necessary to translate a universal conceptual design into a design in which data about facts is stored in tables, and attributes - according to the assumptions of the relational model – are atomic, which sometimes means the need to create additional tables and relationships between them. The logical design, presented in the form of a diagram, in the further phase of the process will be a basis for creation of a physical design and its implementation in a selected relational DBMS [Put, 2009].

Fig. 2 presents the relational database design developed on the basis of the conceptual design.

A database conceptual design, which is the final effect of the designing process participated by a future user, is the basis for creation of a logical design, which takes into account the specifics of a system in which the database will be implemented. The universal character of the conceptual model and its independence of the logical model allow to design databases not taking into account details of a particular model type.

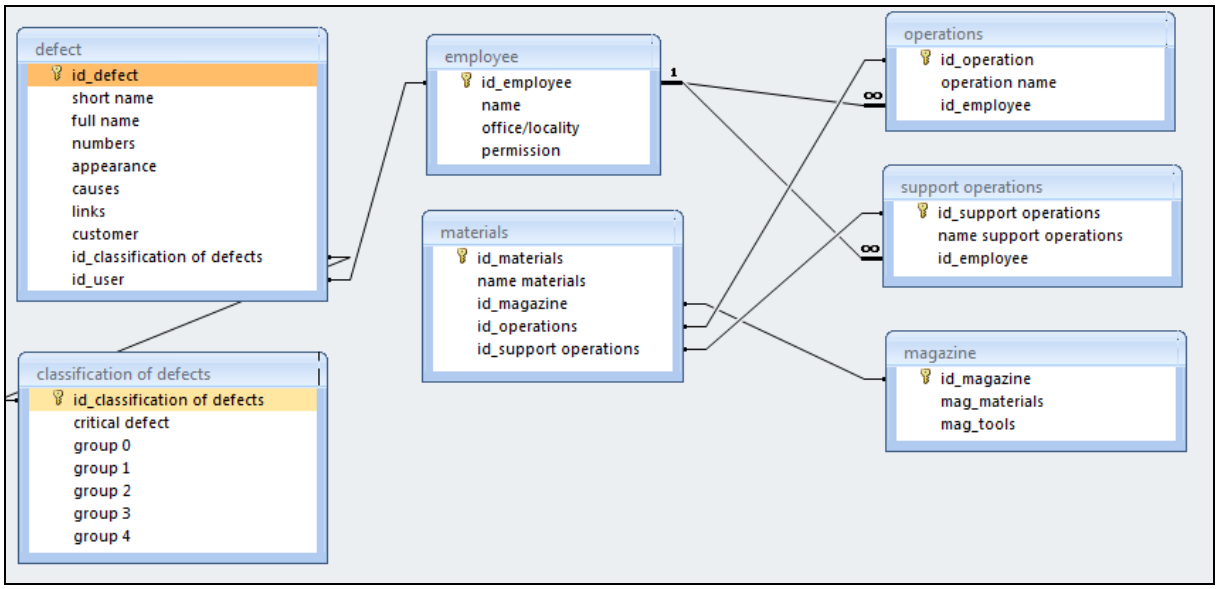


*Fig. 2. R*elationship database *Source: own work*

The following tables have been included in the relational database model:

- Products table - contains information on a product manufactured (product card, finished product specification),
- Materials table - contains information on materials (semi-finished products) used,
- Operations table - contains a list of operations carried out subsequently (technological operations),
- Auxiliary Operations table - contains a list of auxiliary operations to be carried out in the case of specific products (additional operations),
- Defects table - contains detailed information on individual defects.

The database contains vocabulary data and dynamic data and operates in real time. Vocabulary data concerns subject-matter objects and it changes much slower; however, it is also updated if objects are changed. As far as dynamic data is concerned, the database is archived periodically due to fast growth of data and a huge number of records. Dynamic data is related to gathering measurements from measurement points (FP), which read instantaneous measurements. This data will be used in the knowledge acquisition system as teaching files. The database structure is still being developed.

In the current phase, the largest collection of data is data containing classification of defects and complementary data containing a list of defects, defect descriptions, photos, reasons for occurrence of defects and methods of their elimination.

Measurement data indicates correlations between defects. A few correlations between defects are presented below:

- If the defect "melted bottom" occurs, the defect "deformed bottle lip" in a similar quantity can also be expected; in this case there is a lack of glass needed to form a bottle lip,
- uneven bottle – uneven bottle lip,
- thin bottom – thin glass walls – thin product.

Defects which are classified as critical are defects which may cause hazardous conditions for a product user (every defect which may result in glass inside a bottle). Products with such defects should not get into the annealing furnace. If they get into the annealing furnace, this fact should be reported to the cold end and these products should be rejected before they reach the end of the annealing furnace. It should be remembered that any of these defects may cause injury to a customer or a consumer.

Technological knowledge base in a company

Technological knowledge plays a very important role in a manufacturing company. Systems supporting the design of technological processes, currently being developed, allow to make different methods of data presentation available, transform and exchange data.

Knowledge acquisition is a process of defining knowledge, on the basis of which an expert system will provide answers in the form of an expertise. Defining knowledge consists in acquiring knowledge from an expert in a form which allows formalizing it. An expert in a given field is responsible for the content of knowledge, whereas a knowledge engineer is responsible for its form.

Technological knowledge is a collection of information on a technological process realized in specific conditions of a given enterprise. The contribution of an expert and a knowledge engineer are described in the reference titles [Rojek, 2007]. Technological knowledge is a dynamic collection, i.e. it changes in time along with changes of a technological process. Additionally, it is assumed that technological knowledge may be processed in the way specific to the phases of an advisory system development. The following phases of the process are distinguished:

- acquisition of technological knowledge,
- development of technological knowledge representation models,
- recording knowledge in a system's technological knowledge base.

A knowledge engineer uses the following information for assessment of knowledge sources:

- information necessary to carry out work (materials from non-serial and serial publications),
- information concerning all processes realised in the production system (materials collected in the Glassworks, consultations with the plant manager of O-I Produkcja Polska, consultations with specialists having expertise knowledge on different phases of the technological process - an expert's knowledge),
- methods of finished product quality assessment (ideal norm and permissible norm),
- possible variants of modernisation (purchase of new machines, modernisation of existing machines, new technologies, new materials etc.),
- criteria for assessment of variants of the system development.

As a result of a dialogue, on the basis of data entered by a user and data from measurement points the expert system will perform a process consisting of:

- recognition of a defect of a product (here: bottles) and its classification into one of the groups (e.g. Group 0 - critical defect - leaky bottle lip, overblown bottle lip/collar, scratches in a bottle lip/collar),
- recognition of the cause of a defect (a mechanical defect, a defect of a form etc.),
- determining ways or methods of elimination of a defect,
- selection of an optimal solution out of previously determined methods.

On the basis of the process presented above the system suggests a method of elimination of product defects occurring on the production line.

Artificial intelligence package SPHINX by AITECH will be used for development of the intelligent decision support system. The following software tools will be used for the implementation of the system: PC-Shell – expert system shell – for development of basic modules of the system, CAKE – for presentation of knowledge elements and explanation how they are used and DeTreex – to acquire knowledge, decision-making rules from the database.

PC-Shell shell system is a hybrid system with the blackboard architecture, so it may use different sources of knowledge for solving problems. PC-Shell 4.5 supports the following sources of knowledge: expert knowledge bases, applications based on neural networks and databases with text explanations.

A knowledge base in the PC-Shell system is divided into five blocks: the block of knowledge sources description, blocks of facets, rules, facts description and the block of control. A knowledge base in PC-Shell may contain the following elements:

- descriptions, or in other words, facts, which are indicative sentences. A fact may be represented in the form of a relationship between certain objects and have different features (attributes),
- rules, which are indispensable for solving a problem in a given field,
- relationships,
- procedures.

The general format of description of rules in PC-Shell is presented below:

[number_of_rule :] conclusion1 if

condition_1 & condition_2 &...& condition_n.

Example:

[Rule No.1:] <misadjustment of plunger and guide ring> if

<uncentred plunger cylinder>&<plunger cylinder is not aligned with invert>

& <too low plunger cylinder> & <glass reaches plunger>

All rules are numbered and express logical associations between elements of knowledge in a given field or they contain a description of certain actions. Facts, expressed in the form of indicative sentences, represent elements of knowledge and they are treated as statements or conclusions. There are clear semantic associations between rules and facts [Piróg-Mazur, 2011, Buchalski, 2005].

Conclusion

Manufacturing companies currently operating in the market collect more and more data on production processes, delivery processes, customers and their requirements, products' susceptibility to failure and control processes. The decision-making process is a process consisting in processing information. Classical methods of acquiring and analyzing information often fail, and additionally, they often refer to legacy data.

The paper presents the characteristics of the selected object of research, the conceptual data model, formalization of knowledge and the information collected during visits in the manufacturing company for which the intelligent decision support system is being developed.

The need for development of an intelligent decision support system arose from the practice and numerous meetings with a production line manager. It was realized that there was a lack of algorithms of action in the case of finding a defect of a product (here: glass packaging, e.g. bottles, jars) and for selection of an appropriate (the most beneficial) method of elimination of defects.

The system being designed, which is based on integration of selected tools of artificial intelligence and a knowledge base will allow to solve complex problems occurring in the production system faster and more effectively, using experience and intuition of a manager as an expert [Piróg-Mazur, 2011].

Bibliography

[Buchalski, 2006] Buchalski Z., „Zarządzanie wiedzą w podejmowaniu decyzji przy wykorzystaniu systemu ekspertowego", Bazy danych: Struktury, algorytmy, metody, WKiŁ, Warszawa, 2006.

[Buchalski, 2005] Buchalski Z., „Akwizycja wiedzy w systemie ekspertowym wspomagającym podejmowanie decyzji", Bazy danych: Modele, Technologie, Narzędzia, Kozielski S., Małysiak B., Kasprowski P., Mrozek D. (red), WKŁ, 2005

[Piróg-Mazur, 2010] Piróg-Mazur, „Wykorzystanie systemu ekspertowego do wspomagania podejmowania decyzji w przemyśle szklarskim", w Bazy Danych: Aplikacje i Systemy, Studia Informatica Volume 31, Number 2B (90) 2010.

[Piróg-Mazur, 2010] Piróg-Mazur, „Przykład postaci systemu doradczego do zastosowania w przemyśle szklarskim", w: Technologie informatyczne i ich zastosowania. pod red. Aleksandra Jastriebowa, i inni, Politechnika Radomska, Radom, 2010

[Piróg-Mazur, Setlak 2011] Piróg-Mazur M., Setlak G., „Budowa bazy danych oraz bazy wiedzy dla przedsiębiorstwa produkcyjnego w przemyśle szklarskim", w Bazy Danych: Aplikacje i Systemy, Studia Informatica Volume 32, Number 2B (97) 2011.

[Piróg-Mazur, Setlak 2010] Piróg-Mazur M., Setlak G., „Conceptual model of decision support system in a manufacturing enterprise", 2nd International Conference on Intelligent Information and Engineering Systems, INFOS 2010, Rzeszów-Krynica, Poland

[Pondel, 2003] Pondel M, „Wybrane narzędzia informatyczne pozyskiwania wiedzy i zarządzania wiedzą", Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 975. Pozyskiwanie wiedzy i zarządzanie wiedzą, Wrocław, 2003

[Put, 2009] Put D., „Model konceptualny jako wspólna platforma niejednorodnych modeli logicznych", Akademia Ekonomiczna w Krakowie, 2009.

[Rojek, 2009] Rojek I., „ Wspomaganie procesów podejmowania decyzji i sterowania w systemach o różnej skali złożoności z udziałem metod sztucznej inteligencji", Wydawnictwo Uniwersytetu Kazimierza Wielkiego, Bydgoszcz, 2010

[Rojek, 2009] Rojek I., „Inteligentny system wspomagania decyzji dla sterowania siecią wodociągową", projekt badawczy Nr R11 001 01.

[Rojek, 2009] Rojek I., „Baz danych i baza wiedzy dla miejskiego systemu wodno – ściekowego", w: Bazy danych: Nowe Technologie, Kozielski S., Małysiak B., Kasprowski P., Mrozek D. (red), WKŁ, 2007

[Weiss , 2002] Weiss Z., „Techniki komputerowe w przedsiębiorstwie", Wydawnictwo Politechniki Poznańskiej, Poznań, 2002

[Zieliński, 2000] Zieliński S., „Inteligentne systemy w zarządzaniu", Teoria i praktyka, PWN, Warszawa, 2000.

[Źródło] Stowarzyszenie forum opakowań szklanych, 2010

Authors' Information

Monika Piróg-Mazur *– Editor in chief, Institute of Technical Engineering, The Bronislaw Markiewicz State School of Technology and Economics in Jaroslaw, Czarniecki Street 16, 37-500 Jaroslaw, Poland; e-mail: m_pirog@pwste.edu.pl*

Major Fields of Scientific Research: knowledge representation, hybrid intelligent system.

Galina Setlak *– D.Sc, Ph.D., Associate Professor, Rzeszow University of Technology, Department of Computer Science, W. Pola 2 Rzeszow 35-959, Poland, and The State Professional High School, Czarnieckiego 16, Jarosław, Poland, e-mail: gsetlak@prz.edu.pl*

Major Fields of Scientific Research: decision-making in intelligent manufacturing systems, knowledge and process modeling, artificial Intelligence, neural networks, fuzzy logic, evolutionary computing, soft computing.

STUDY OF INTEGRATION ALGORITHM AND TIME STEP ON MOLECULAR DYNAMIC SIMULATION

Janusz Bytnar, Anna Kucaba-Piętal

***Abstract**: A simulation is reliable when the simulation time is much longer than the relaxation time of the quantities in question. The aim of this work is to address the question when Molecular Dynamics (MD) simulation is reliable and how it depends on the integration algorithms and optimal time step. There were certain problems related to the choice of integration algorithms on Molecular Dynamics simulations. The effect of time step on convergence to equilibrium in Molecular Dynamics simulation has been studied.*

***Keywords**: Molecular Dynamics, computer simulations, integration algorithms*

***ACM Classification Keywords**: A.0 General Literature - Conference proceedings*

Introduction

After obtaining the results of a research, each scientist needs to consider verification and validation of those results. Computer simulation of molecular systems is playing an ever growing role in academic and industrial research. In areas ranging from materials science and chemistry to pharmacy and molecular biology, computer simulation is already a part of daily practice. The behavior of a variety of molecular systems can be studied by using the Molecular Dynamics (MD) simulation method. These include liquids, solutions, electrolytes, polymers such as proteins, DNA, and polysaccharides, as well as membranes, liquid crystals, crystals, and zeolites [Allen, 1987], [Bicout , 1996].

Computer simulation of molecular systems requires software to calculate the interatomic interactions and to integrate the equations of motion [Griebel, 2007].

Many models, for example in materials science or in astrophysics, contain large number of interacting bodies (called particles), as for example stars and galaxies or atoms and molecules. In many cases the number of particles can reach several millions or more. For instance, every cubic meter of gas under normal conditions (i.e., at temperature of 273.15 Kelvin and pressure of 101.325 kilopascal) contains 2.68678 x 10^25 atoms (Loschmidt constant). 12 grams of the carbon isotope C12 contain 6.02214 x 10^23 atoms (Avogadro constant).

These are some of the reasons why computer simulation has recently emerged as a third method in science besides experimental and theoretical approaches. Over the past years, computer simulation has become an indispensable tool for the investigation and prediction of physical and chemical processes. In this context, computer simulation means the mathematical prediction of technical or physical processes on modern computer systems [Griebel, 2007].

The deterministic method of Molecular Dynamics (MD) simulation, although theoretically valid for the whole range of densities, is employed mainly for liquids and solids [Allen, 1987]. The long flight paths between collisions of gas molecules make the method of Molecular Dynamics prohibitively expensive, while other methods, like e.g. Direct Monte-Carlo Simulation, can give satisfactory results at much lower computational cost. Molecules in liquids are densely packed and remain in constant contact with the neighbours. Under such conditions Molecular Dynamics seems to be the most accurate and, at the same time, the most efficient simulation method.

Molecular Dynamics requires the description of the molecules and the forces acting between them. Perhaps the most often, to describe the Van-der-Waals forces, the Lennard-Jones potential is used. It assumes that the molecules are spherically symmetric, repelling one another at close and attracting at far distances.

Procedure of Molecular Dynamic

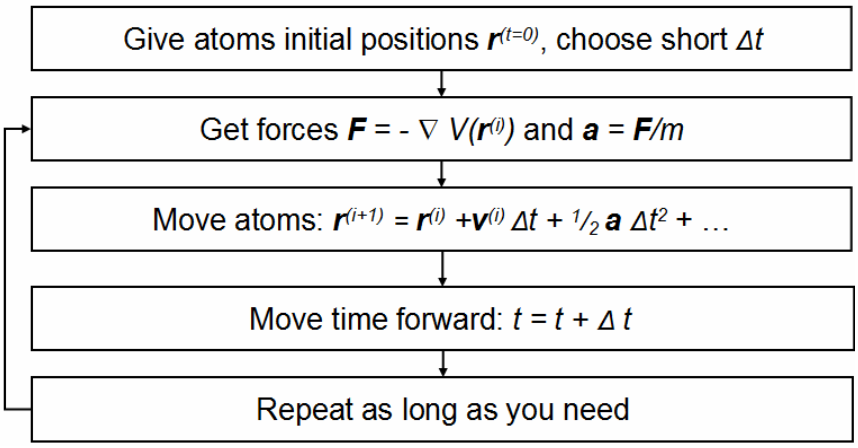
Molecular Dynamics (MD) is a computer simulation technique:

- the time evolution of interacting atoms is followed by integrating their equations of motion.
- the laws of classical mechanics are followed, and most notably Newton's law:

$$F_i = m_i a_i \tag{1}$$

$$a_i = d^2 r_i = dt^2 \tag{2}$$

- The Molecular Dynamics (MD) procedure can be written as follows:



- Given the initial set of coordinates and velocities, the subsequent time evolution is in principle completely determined.
- Atoms and molecules will 'move' in the computer, bumping into each other, vibrating about a mean position (if constrained), or wandering around (if the system is fluid), in a way similar to what real atoms and molecules do.
- The computer calculates a trajectory of the system
- 6N-dimensional phase space (3N positions and 3N moments).

The orientation of the molecules can be represented in several ways, however the use of quaternions [Refson, 2001] seems to be the most advisable. The most important advantage of quaternions is the fact, that they lead to equations of motion free of singularities (which is not the case for e.g. Euler angles). This, in turn, leads to good numerical stability of the simulation.

Integration algorithms used in Molecular Dynamics simulation are based on finite difference methods, with discretized time and the time step equal to Δt. Knowing the positions and some of their time derivatives at time t (the exact details depend on the type of algorithm), the integration scheme gives the same quantities at a later time ($t + \Delta t$). With such procedure the evolution of the system can be followed for long times [Allen, 1987] .

Stages of simulation:

Initiation: placing the molecules of water and the copper atoms in the knots of crystalline mesh. After that the velocities of the molecules are initialized. Their values are sampled at random from the Maxwell – Boltzmann distribution for the assumed temperature.

Balancing: after initiation the positions of molecules are far from equilibrium. The whole ensemble is allowed to move freely for some time to attain equilibrium positions. This is always connected with decreasing the potential and increasing the kinetic energy of the molecules, i.e. increasing the temperature of the medium. This excess temperature must be removed by a suitable "thermostat".

Actual simulation: after attaining equilibrium, the simulation starts. The required data (specified in advance) are accumulated in "dump-files" in preselected time intervals. Any property of interest.

Integration Algorithms

The engine of a Molecular Dynamics program is the time integration algorithm, required to integrate the equation of motion of the interacting particles and follow their trajectory.

The integration scheme gives the possibility to find particle position at a later time t + Δt. By iterating the procedure, the time evolution of the system can be followed for long times.

Of course, these schemes are approximate and there are errors associated with them. In particular, one can distinguish between:

- Truncation errors, related to the accuracy of the finite difference method with respect to the true solution. Finite difference methods are usually based on a Taylor expansion truncated at some term, hence the name. These errors do not depend on the implementation: they are intrinsic to the algorithm.
- Round – off errors, related to errors associated to a particular implementation of the algorithm. For instance, to the finite number of digits used in computer arithmetic.

Both errors can be reduced by decreasing Δt. For large Δt, truncation errors dominate, but they decrease quickly as Δt is decreased. For instance, the Verlet algorithm has a truncation error proportional to Δt^4 for each integration time step. Round – off errors decrease more slowly with decreasing Δt, and dominate in the small Δt limit [Ercolessi, 1997].

Using time integration techniques, it is possible to determine the velocity and position of a particle from its acceleration. There is a variety of different numerical methods available, however the nature of Molecular Dynamics simulations has narrowed down the field to a handful of methods. Methods which require more than one force calculation per time step are considered wasteful and can only be considered if the time step can be proportionally increased, while still maintaining the same accuracy.

Similarly, adaptive methods that change the time step dynamically are useless due to the rapidly changing neighbourhood of each atom. As a result, only two methods have become mainstream in Molecular Dynamics field, that is, the Verlet method and predictor-corrector method [Rapaport , 2004].

Both methods are based on finite difference techniques, derived from the Taylor expansion of the r(t).

Basic Verlet Method

In Molecular Dynamics, time integration algorithm that is used very common is Verlet algorithm [Verlet, 1967]. The basic idea is to write two third-order Taylor expansions for the positions r(t), one forward and one backward in time. The basic form of the Verlet method is defined by the equation:

$$r(t+\Delta t) = 2r(t) - r(t-\Delta t) + (\Delta t)^2 a(t) + O(\Delta t^4) \quad (3)$$

where a(t) is the acceleration. Via the combination of the force calculation with Newton's second law of motion, the acceleration is defined as

$$a(t) = -(1/m)\nabla U(r(t)) \tag{4}$$

While not required for computation, the velocity variable can be found by using the equation

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} \tag{5}$$

The Verlet algorithm uses positions and accelerations at time t and positions from time t-Δt to calculate new positions at time t+Δt. The Verlet algorithm uses no explicit velocities. The advantages of the Verlet algorithm are:

- It is straightforward
- The storage requirements are modest

The disadvantage is that the algorithm is of moderate precision.

A problem with this version of the Verlet algorithm is that velocities are not directly generated. While they are not needed for the time evolution, their knowledge is sometimes necessary. Moreover, they are required to compute the kinetic energy E_K , whose evaluation is necessary to test conservation of the total energy $E = E_K+E_P$. We can also calculate temperature of the simulated molecular system from kinetic energy.

$$E_K = \frac{1}{2} k_B T \tag{6}$$

Where k_B is the Boltzman constant, T is the temperature.

This is one of the most important tests to verify that a Molecular Dynamics simulations of real processes is proceeding correctly.

However, the error associated to this expression is of order Δt^2 rather than Δt^4. To overcome this difficulty, some variants of the Verlet algorithm have been development.

Velocity Verlet Method

However more common algorithm is a related one, Velocity Verlet algorithm. Here the velocity, position and accelerations at time t+Δt are obtained from the same quantities at time t [Verlet, 1967]. This uses a similar approach but explicitly incorporates velocity, solving the first-time step problem in the Basic Verlet algorithm:

$$r(t+\Delta t) = r(t) + v(t)\Delta t + 1/2\, a(t)\Delta t^2 \tag{7}$$

$$v(t+\Delta t) = v(t) + 1/2\,[a(t) + a(t+\Delta t)]\Delta t^2 \tag{8}$$

Beeman Method

This algorithm is also closely related to the Verlet algorithm

$$r(t+\Delta t) = r(t) + v(t)\Delta t + 2/3\, a(t)\Delta t^2 - 1/6\, a(t-\Delta t)\Delta t^2 \tag{9}$$

$$v(t+\Delta t) = v(t) + v(t)\Delta t + 1/3\, a(t+\Delta t) + 5/6\, a(t)\Delta t - 1/6\, a(t-\Delta t)\Delta t \tag{10}$$

The advantage of this algorithm is that it provides a more accurate expression for the velocities and better energy conservation. The disadvantage is that the more complex expressions make the calculation more expensive [Beeman, 1976].

A variant of the Verlet method, called the velocity-Verlet method, addresses this problem by directly including the velocity in computation. As a result, particle velocities are known at the same time step as coordinates, and the high-order accuracy of the method is maintained. Additionally, particle velocities are necessary for kinetic energy calculations, which play a critical role in most Molecular Dynamics simulations [Haile, 1997].

Time Step

Lennard – Jones potential is the most popular interaction potential used in Molecular Dynamics (MD) simulations to describe Van-der-Waal forces [Karniadakis, 2005]. The form of the Lennard – Jones potential is as follows:

$$V(r) = 4\varepsilon\left[\left(\frac{\delta}{r}\right)^{12} - \left(\frac{\delta}{r}\right)^{6}\right] \tag{11}$$

where □ and □ are the Lennard – Jones parameters that depend on the atoms involved in the interaction. Note that:

- □ is related to the interaction strength, and a higher □ corresponds to a higher interaction energy between the atoms
- □ corresponds to the distance at which the potential between the two atoms goes to zero, which can be approximately taken as the diameter of a fluid atom.

The term ~$1/r^{12}$, dominating at shorter distance, models the repulsion between atoms when they are brought very close to each other.

The term ~$1/r^6$, dominating at large distance, constitute the attractive part. This is the term which gives cohesion to the system. A $1/r^6$ attraction is originated by van der Waals dispersion forces, originated by dipole – dipole interactions in turn due to fluctuating dipoles. These are rather weak interactions, which however dominate the bonding character of closed – shell systems, that is, rare gases such as Argon. Therefore, these are the materials that Lennard – Jones potential could mimic fairly well [Ercolessi, 1997]. The parameters □ and □ are chosen to fit the physical properties of the material.

In the Molecular Dynamics (MD) simulation with Lennard – Jones interaction potentials, the time and the other physical quantities are represented and typically computed using reduced units. Table 1 summarizes the units for various quantities used in simulations for instance, length, temperature, and density. In the Table 1, symbols □ and □ denote constants as defined in equation (10), k_B is the Boltzman constant, and m is the mass of a atom.

Table 1. Units for various quantities in Lennard – Jones fluids [Griebel, 2007], [Karniadakis, 2005]

Length	δ	Velocity	$(\varepsilon/m)^{1/2}$
Mass	m	Shear rate	$(\varepsilon/m\delta^2)^{1/2}$
Energy	ε	Stress	ε/δ^3
Time	$(m\delta^2)/\varepsilon)^{1/2}$	Viscosity	$(m\varepsilon)^{1/2}/\delta^2$
Number density	δ^{-3}	Diffusivity	$\delta(\varepsilon/m)^{1/2}$
Temperature	ε/k_B		

In many publications [Griebel, 2007], [Karniadakis, 2005] authors calculate time step from the formula:

$$\Delta t = 0{,}001 * \sqrt{\frac{m\delta^2}{\varepsilon}} \tag{12}$$

We intended to use as large a time step as possible so that we can explore more of the phase space of the system. However, since we truncate the Taylor's series expansions, the time step needs to be small enough so that the expansion can provide a reliable estimate of the atomic positions and velocities at the end of the time step (see Fig. 1). For typical algorithms with a time accuracy of order three, one uses a time step that is a fraction of the period of the highest-frequency motion in the system [Karniadakis, 2005]. For a typical simulation of water transport, where the O – H bond length is fixed, a time step size of 1.0 to 2.0 fs is commonly used.

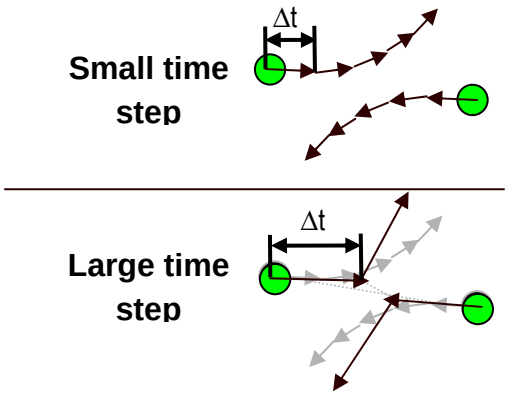


Fig. 1. The effect of the small and large temporary step

Results

The aim of the research was to analyze the influence of integration algorithm and time step on the computational time and memory complexity. Also influence of the size of molecular systems on efficency integration algorithms was studied.

The simulations were carried out for one molecular model of water TIP4P and three integration algorithms were applied: Velocity Verlet, Beeman and Beeman algorithm with Predictor – Corrector modifications.

We use for all algorithms different time steps Δt = 0,00001, Δt = 0,00002, Δt = 0,00005, Δt = 0,0001 and Δt = 0,0005 picosecond long. The calculations were carried out over 100 000 time steps.

The program MOLDY [Refson, 2001], suitably modified, was used for this purpose. Moldy is free software; which may redistribute it and/or modify it under the terms of the GNU.

The physical properties of materials and their electrostatic interactions were taken into account. The number of water molecules was equal to 500 and 20000. The periodic boundary conditions were applied. The Lennard-Jones potential was assumed for interactions between water molecules [Kucaba – Pietal, 2004], [Bytnar, 2008].

Molecular Dynamics is always an approximate science approach, the longer the time step, the less accurate the results. In the worst case scenario, the time step will allow atoms to move too far between single iterations, allowing atoms to get closer together than they ever could in a real liquid. This usually causes an incorrect "chain reaction", whereby two close particles repel at a much faster speed than normal causing them to bump even closer other atoms, which are repelled at an even greater velocity. This effect compounds until all atoms are moving at unrealistic speeds and eventually arithmetic overflows will occur. When we attempted to use larger time step (1 fs) the program crashed.

If we only consider figures 2-5 then we can deduct that ideally time step should be as small as possible.

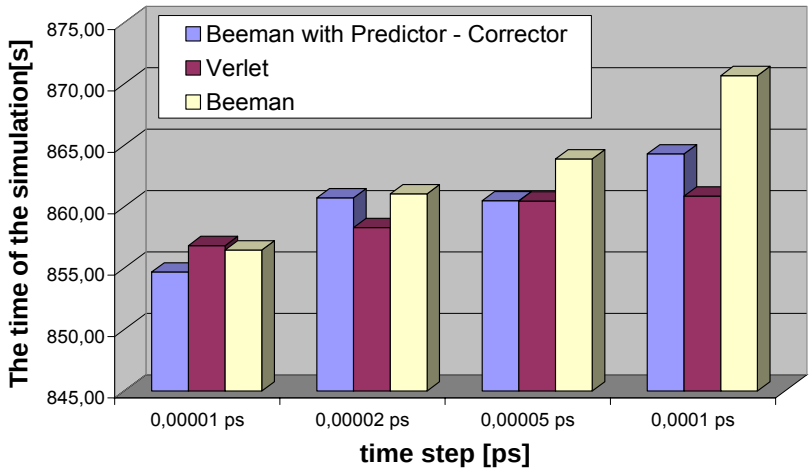


Fig. 2. The time of the whole simulation (500 molecules of water) – Integration algorithms with various time steps

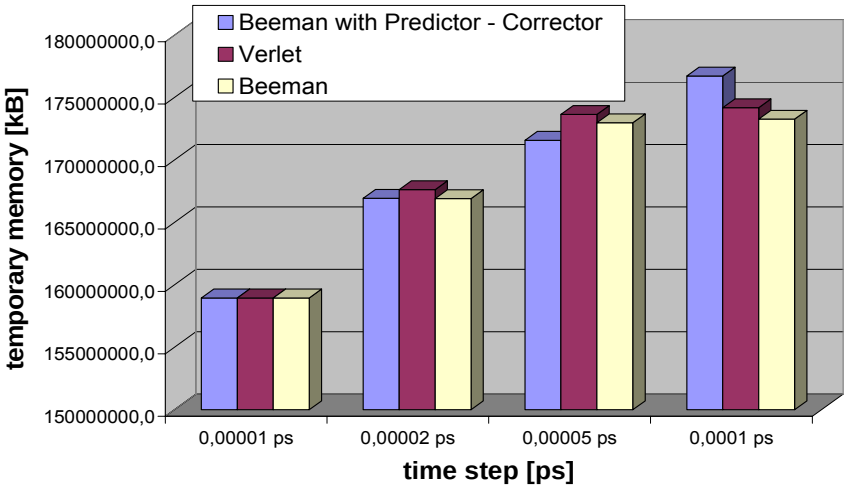


Fig. 3. The reservation of the temporary memory (500 molecules of water) – Integration algorithms with various time steps

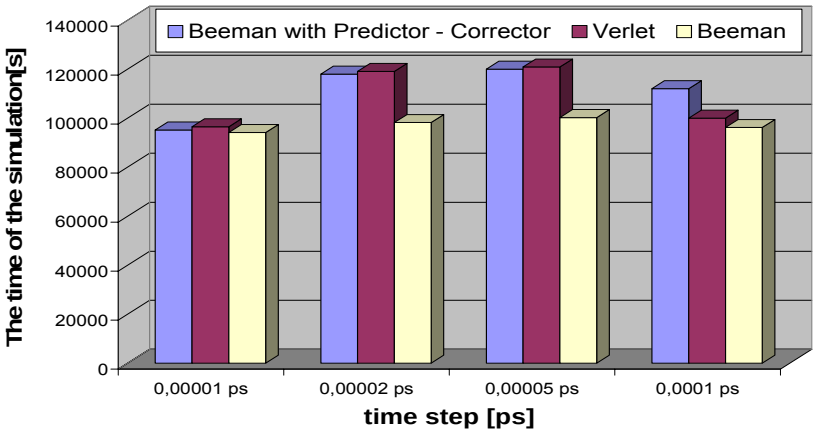


Fig. 4. The time of the whole simulation (20000 molecules of water) – Integration algorithms with various time steps

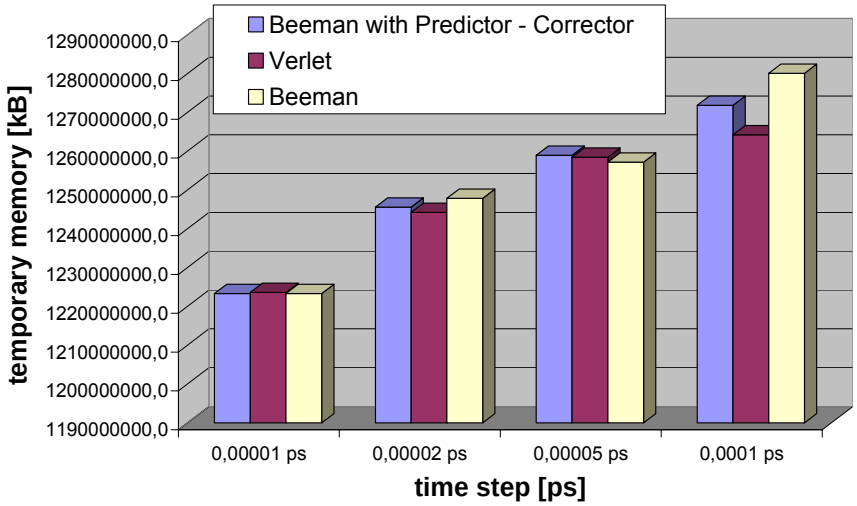


Fig. 5. The reservation of the temporal memory (20000 molecules of water) – Integration algorithms with various time steps

In figures 2-5, several different time step simulations were run in different integration algorithm. Using a small time step of 0,00001 ps gives good results connected with the reservation of the memory and the time of the whole simulation. The increment of the time step forces the use of the greater spaces of the memory and time of the whole simulations.

Conclusions

In Molecular Dynamics simulation, there exist several algorithms to realize equations of motion of molecular systems.

From the presented diagrams we can see that the choice of these algorithms with correct time steps decides about the time of the whole simulations and also about the necessary reservation of the memory.

The choice of the time step in integration algorithm from equation (11) can be good for abstract molecular models. However, this type of simulations does not take into account many factors, which are very important for simulations of real materials. In Molecular Dynamics simulations of real processes not only speed of calculations and reservation of the memory are important. In this case what is very important is the thermodynamics of the simulated molecular configuration during the whole simulation as also the obtained results of physical properties simulated materials which in confrontation with the results of these materials in the larger scale will be satisfactory.

The influence of the correct choice of the time step was also present in paper [Bytnar, 2010]. From the presented diagrams (considered in paper [Bytnar, 2010]) it is clear, that for the problems considered, i.e. real flows in nanochannels, the Gaussian thermostat is much more efficient if the time step Δt is well chosen.

If we consider only time of the whole simulations and required memory (see Figures 2-5) we can say that the best performance and accuracy is if the time step is as small as possible also for small and large molecular systems.

Additionally we can see that the performance of all integration algorithms (Velocity Verlet, Beeman and Beeman with Predictor – Corrector modifications) is very similar also for small (500 molecules of water) and large (20000 molecules of water) molecular systems. This means that the choice of the integration algorithm does not have large signification for Molecular Dynamics simulations.

The smaller time step effects enlargement of the time of the whole simulation and the required memory, however the larger temporary step can effect generating the inaccurate trajectory of motion. Clearly, the larger time step, the less accurately our solution will follow the correct classical (see Fig 1).

From the presented figures in this paper it can be deduced that a good way of checking whether the time step is satisfactory is to run an equilibrium simulation of small molecular system because results for small (500 molecules of water) and large (20000 molecules of water) molecular system are wery similar.

In summary, the choice of time step has a big impact on accuracy of simulations. It is recommended that, to avoid the incorrect “chain reaction” phenomena, if two atoms get unrealistically close, the user should be warned that the time step should be decreased and be given the option to terminate the program since the results are already effectively useless. A more advanced program might provide warnings if the user enters an unrealistically large time step before the simulation is allowed to start.

The physical explanation of other peculiarities of the presented diagrams, particularly the thermodynamics properties molecular systems, requires further investigation.

The study of Verification and Validation methods gives the possibility of the construction of the correct computer model to the description of the studied phenomenon, how also receipt of the exact and authentic results of computer simulations.

Influence on the results of the simulation in the Molecular Dynamics method has e.g. Integration algorithms and time step, molecular model of water or another material, mechanisms to control the temperature of the system (thermostats). Therefore, the study of the methods of Verification and Validation will be closely connected with the above mentioned factors.

Acknowledgment

Calculations were made in the Interdisciplinary Centre for Mathematical and Computational Modeling (ICM) University of Warsaw (grant no. G44-9):

https://granty.icm.edu.pl/IcmGrants/displayGrant/showGrants.jsp

We also want to thank for the possibility of the realization of calculations in the Institute of Fundamental Technological Research PAN, Department of Mechanics and Physics of Fluids in Warsaw.

Bibliography

[Allen, 1987] M. P. Allen and D. J. Tildesley: Computer Simulation of Liquids, Clarendon Press, Oxford, 1987,

[Bicout , 1996] D. Bicout and M. Field: Quantum Mechanical Simulation Methods for Studying Biological Systems, Springer, Berlin, 1996,

[Griebel, 2007] M. Griebel S. Knapek, G. Zumbusch: Numerical Simulation in Molecular Dynamics, Numerics, Algorithms, Parallelization, Applications, Springer-Verlag Berlin Heidelberg, 2007,

[Ercolessi, 1997] F. Ercolessi: A molecular dynamics primer, Spring College in Computational Physics, ICTP, Trieste, Italy, 1997,

[Rapaport , 2004] D. C. Rapaport: The Art of Molecular Dynamics Simulation. Cambridge University Press, 2004,

[Verlet, 1967] L. Verlet: Computer `experiments' on classical fluids. I. thermodynamical properties of Lennard-Jones molecules, Phys. Rev. 165, 201-214, 1967,

[Beeman, 1976] D. Beeman: Some multistep methods for use in molecular dynamics calculations, J. Comp. Phys. 20 130-139, 1976,

[Haile, 1997] J. Haile: Molecular Dynamics Simulation: Elementary Methods. New York: John Wiley and Sons Inc., 1997,

[Karniadakis, 2005] G. Karniadakis, A. Beskok, N. Aluru: Microflows and Nanoflows – Fundamentals and Simulation, Interdisciplinary Applied Mathematics, Springer, 2005,

[Refson, 2001] K. Refson: Moldy User's Manual. Chapter II, ftp://ftp.earth.ox.ac.uk/pub,

[Bytnar, 2010] J. Bytnar, A. Kucaba-Piętal, Z. Walenta: Veryfication and Validation of Molecular Dynamics Simulation, Publications S.H.F. Second European Conference in Microfluidics, ISBN 978-2-906831-85-8, Session P4-3-Liquid Microflows, Toulouze, France, 2010,

[Bytnar, 2008] J. Bytnar, A. Kucaba-Piętal, Z. Walenta: Influence of molecular models of water on computer simulations of water nanoflows - Proceedings of International Conference on Computer Science and Information Technology, Volume 3, ISSN 1896-7094 pages 269 – 275, IEEE CS Press, Los Alamitos, CA, 2008

[Kucaba – Pietal, 2004] A. Kucaba-Pietal: Microflows modelling by use micropolar fluid model, OW RUT, Rzeszow, 2004;

Authors' Information

Janusz Bytnar *–Technical and Economical State School of Higher Education in Jaroslaw, Institute of Technical Engineering, ul. Czarnieckiego 16, 37-500 Jaroslaw, Poland; e-mail: janusz.bytnar@pwste.edu.pl*

Major Fields of Scientific Research: Software technologies, Molecular Dynamics, Computer simulations, Integration algorithms research

Anna Kucaba - Pietal *– Technical and Economical State School of Higher Education in Jaroslaw, Institute of Technical Engineering, ul. Czarnieckiego 16, 37-500 Jaroslaw, Poland;*

Rzeszow University of Technology, The Faculty of Mechanical Engineering and Aeronautics Powstancow Warszawy 8, 35-959 Rzeszow, Poland; e-mail: anpietal@prz.edu.pl

Major Fields of Scientific Research: Scientific Calculations, Computattional Mechanics,, Molecular Dynamics, Computer simulations, Nano and Micromechanics

INFORMATION SYSTEMS FOR METROLOGY

Roman A. Tabisz, Łukasz Walus

***Abstract**: An important application of the information technology (IT), which is the creation and improvement of Information Systems for Metrology (ISM) is discussed. These systems initially had a form of Metrological Database (MD) and Metrological Knowledge Bases (MKB). At present, the advanced versions of these systems have a form of the Metrological Expert Systems (MES) using artificial intelligence methods. Two selected examples of Information Systems for Metrology and their practical use are described. An actual state of the ISM ("CAMPV System") designed and developed in the Department of Metrology and Diagnostic Systems, Faculty of Electrical and Computer Engineering at Rzeszow University of Technology, is presented. The system consist of the portal for communication with the system via the Internet and the Metrological Database (MD) collecting the data on the specification of measurement equipment as well as the data obtained in the process of equipment calibration. It has been planned that the system will be expanded by adding the Analytical Metrological Database (AMD) and Metrological Knowledge Base (MKB). A full version of "CAMPV-Expert System" will be dedicated to the computer-aided design of electrical and electronic measurement channels as well as the computer-aided validation of measurement processes, which include such channels.*

***Keywords**: information systems for metrology, metrological databases, metrological knowledge bases, metrological expert systems.*

***ACM Classification Keywords**: A.0 General Literature - Conference proceedings*

***Conference topic**: Industrial Control and Monitoring*

Introduction

The development of microelectronics and related development of information technology (IT) has enabled the creation of measurement information systems (MIS), which allow for measuring and collecting large amounts of measurement data. A vital part of the measurement equipment is the software, which carries out important functions related to the performance of measurement activities and therefore determines the accuracy of final results obtained in the measurement process. Such kind of software is called measurement systems software (MSS). The measurement results are now the primary source of our knowledge about the objects and the phenomena of the real world. The metrology therefore has become the basis for the development of many different fields of science, industry and trade. Measurement results are also the foundation on which the scientific claims are created. In the industrial applications, the measurement results are used for the assessment of products' conformity with their technical specification, the evaluation of the quality of manufacturing processes as well as the control and monitoring of their condition. In the trade, the mutual financial settlements are often based on the measurement results. In each of these areas the highest possible quality of measurement results is desired. The basic condition of reaching this goal is the proper design and appropriate execution of the measurement process. Moreover, the quality of measurement results largely depends on the quality of the hardware and software as well as the level of competence of the operators supervising the measurements processes. The management of the measurement equipment and the effective control of the measurement processes require some additional metrological actions such as calibration, repeatability and reproducibility analysis (R&R) as well as inter-laboratory comparison (IC). This increases the need to collect the data concerning the properties of the measurement equipment and the data obtained in the calibration processes. For this reason, the Metrological Databases (MDB) are created with the use of appropriate IT.

The general model of the measurement process, which incorporates the main factors determining the numerical values of the final measurement results is presented on the Fig.1., on which the measuring equipment is specified by a rectangle. The measurement equipment includes both hardware and software. The circle inscribed in the rectangle marks software as an integral part of the measurement equipment. The software is now incorporated in almost all kinds of measuring equipment as it outscored other technologies when it comes to the implementation of important measurement functions such as: measurements process control, correction of the numerical values of the measured quantity, saving the measurement results as well as reading, decoding and displaying them.

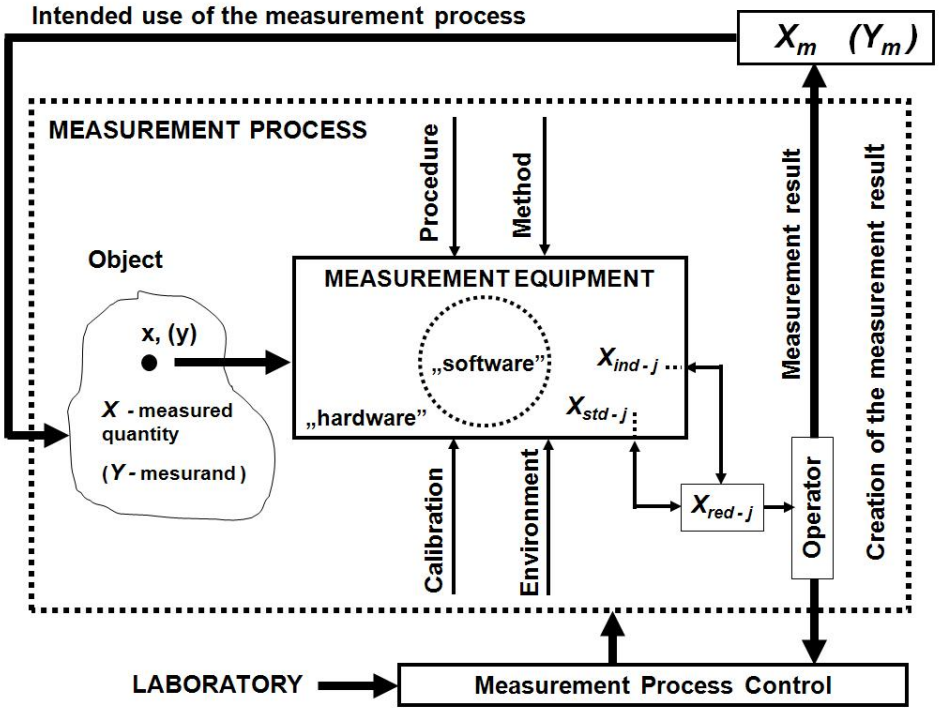


Fig.1. The general model of the measurement process.

x *- numerical value of the measured quantity*

y *- numerical value of the measurand*

X_{ind-j} *- numerical value indicated on the display of measurement equipment*

X_{std-j} *- numerical value stored in the memory of the measurement equipment*

X_{red-j} *- read the value, on which measurement result (* X_m *), is created*

Y_m *- measurand, calculated based on several measured quantities (* X_i *)*

The general model of the measurement process, as shown in Fig.1., has been developed taking into account the original definition of the measure proposed in [1], and models of Information Measurement Systems (IMS) proposed in [2]. These models take into account the types of measurement scales. With such approach, this model is universal and can be applied to all situations in which measurements are carried out. It is independent from the measurement scale adopted to implement the measurement procedure as well as from the composition of the hardware and the software.

The types of the measurement scales have been defined in [3]. In every area of possible application - scientific research, industry or trade - the measurement processes are performed in the specific order. The ultimate goal is always to describe the properties of the real object according to its true state. It allows for making right decisions

when it comes to the objects' management and helps the scientists in formulating the accurate and consistent descriptions of the reality of the phenomena. Relevance assessment of the actual state of measured properties of an object, event or phenomenon, depends largely on the accuracy and proper interpretation of obtained results, which, in turn, is strictly associated with the type of the measurement scale. If above is not taken into account, not only the measurements become ruseless but their false results may lead to wrong decisions.

A definition of the measurement proposed in [1] is as follows:

„The measurement is a process of empirical, objective assignment of symbols to attributes of objects and events in the real world, in order to describe them". [1]

However, the proposal formulated in [2] differentiates the MIS due to the measurement scales used and has been developed in the following order:

"The models proposed in this paper can be used for the development of concepts, principles, and guidelines, which can support decision making in the arrangement of design and application of the Measurement Information Systems (MIS)". [2]

„In order to apply these scientific and technical achievements, the designer or the user of a MIS has a possibility to choose between different measurement procedures and also between different constructive modules of the MIS". [2]

Such approach - taking into account the definition proposed in [1] and the proposal of different types of MISs presented in [2] – allows for the correct interpretation of the measurement results and collected information about the measured object, event or phenomenon. Consequently, it makes possible to gain credible knowledge about the real world, manufactured products and also about the measurement science. In order to ensure the desired quality of the measurement results, one should design, manufacture and exploit the MIS using the appropriate hardware and software composition as well as taking into account the type of measurement scale applied in the measurement process. Only properly designed, implemented and applied MIS can be used to effectively collect large amounts of credible data. Credible measurement data can and should be collected in the Operational Metrological Databases (OMD). These databases currently represent the most basic form of Information Systems for Metrology (ISM). The more complex form of the ISM are systems in which the Analytical Metrological Database (AMD), also called data warehouse, is added to the OMD. The extraction of data from OMD to the AMD made through suitably designed ETL operation (Extraction Transaction Loading) provides a grouping of the data according to a well-thought-out strategies and allows for their storage in a useful and inviolable form. This is necessary to carry out various kinds of analysis, which require extracting desired information and knowledge. The concept of extracting information from the measurements and acquiring knowledge from obtained information is described in [4]. It justifies the usefulness of the definition of measurement proposed in [1].

The most complex form of the ISM are Metrological Expert Systems (MES), in which the properly prepared Metrological Knowledge Bases (MKB) are supplemented by automatic inference systems using artificial intelligence methods. Information Systems for Metrology (ISM), including OMD, AMD, MKB and Inference Systems (IS) can be used in metrology for the computer-aided design of measurement processes and measurement equipment (hardware and software). They can also be used to improve the competence of the operators, who supervise the measurement processes and improve the metrological characteristics of MISs, by adjusting the measured numerical values in order to ensure the accuracy of the final results.

In order to present the specific applications of ISM, two selected examples will be described as well as the current state of the third one, which currently has been in the process of design, implementation and development. The first example is the Metrological Knowledge Base - one of major components of the education system operators of Coordinate Measurement Machines (CMM) [5]. The second example is the metrological expert system using

neural networks to predict corrections, calculate the Polish Universal Time Coordinate - UTC (PL) and propagate the Polish Official Time - OT (PL) [6]. The third example is a computer system supporting the validation of the measurement process, designed, implemented and developed in the Department of Metrology and Diagnostic Systems, Faculty of Electrical Engineering and Computer Science, Rzeszow University of Technology. Each of described examples implements the concept of knowledge acquisition from information extracted from the measurements data collected in the appropriate metrological databases and then stored in the metrological knowledge bases.

The process of extracting knowledge from data described in [4] goes in the order, which can be simply described as: *"measurement-information-knowledge".* This order has been used in all described examples, but in each one for different reason. In the first example, the objective is to improve the competence of the operators of CMM. The operator is the key person responsible for the creation of the final result in the model of measurement process shown in Fig.1. Therefore, the more complex is the measurement equipment, the more qualified the operator should be. In the second example, the metrological expert system analyzes the data published by the Time Laboratory of the International Bureau of Weights and Measures in Paris (BIMP) [7] and predicts the corrections, which allow for calculation of the exact value of the Polish Universal Time Coordinated - UTC (PL). The difference between these corrections and the corrections published by the BIMP should not exceed 10 ns. In the third example, the main goal is to use the *"CAMPV System" (Computer Aided Measurement Process Validation System)* for validation of the measurement process, conducted on the basis of historical calibration data of measurement equipment, stored in the metrological database. This system also allows testing and calibration laboratories for cooperating on-line.

The IT System for Metrology used in the European Education and Training Programme for operators of CMMs

The European education and training programme for operators of CMMs was developed in 2001-2005 within the European Research Project EUKOM [5]. This project was funded by "LEONARDO DA VINCI" programme of European Commission, DG Education and Training. The main objective of the project "European Training for Coordinate Metrology" was to create a common, Europe-wide approach to the training in the field of coordinate measuring technology, which would meet the requirements of continuing education. The results of the project were implemented by the association CMTrain e.V. [8], independent of CMM producers, which currently is being developed and tested of training by a mixed method. This method combines the capabilities of distance learning (e-learning) with verification of acquired knowledge during the practical handling of CMMs. In the original version the distance learning module was created with the ILIAS system [9], which is the open source software. The effectiveness of learning was achieved through a carefully prepared Metrology Knowledge Base developed by experts from the six centers from different countries. The project was coordinated by the Chair of Quality Management and Manufacturing Metrology at University of Erlangen-Nuremberg. [10]. The structure of Metrological Knowledge Base EUKOM system has been adapted into the learning material coherent and common for the whole of Europe.

The starting point for the study was to establish three levels of competence of the operators: level 1 - "CMM-User", level 2 - "CMM-Operator" and level 3 - "CMM-Expert." For each level of competence approx. 15 training modules was developed and made available online. Respective modules include the learning content appropriate for a given level of competence and concerning areas where knowledge is required and needed for the CMMs' operators. The areas include: metrology, geometry, statistics, computer science, quality management, standardization, measurement equipment, technology of production and technical drawing systems used in computer-aided design systems (CADs). Thanks to the cooperation of specialists in the field of construction and

exploitation of Coordinate Measuring Machines, a three-levels metrological knowledge base structure was developed, which is shown in Table.1.

Table 1. Three-levels structure of the Metrological Knowledge Base of the EUKOM system for education of CMMs' operators

Types of qualification	Levels	Modules of the Metrological Knowledge Base of the EUKOM System.
CMM- User	I	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
CMM- Operator	II	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
CMM- Expert	III	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

The structure of the Metrological Knowledge Base of the EUKOM system, shown in Table.1., was developed by the eminent experts from seven European universities [10] engaged in the development of CMM metrology and its applications in the industry. The 3-levels structure of education, including about 15 basic training modules at every level of competence, was established after the adoption of the basic assumption that the industries using the CMMs need three categories of operators. Three categories of operators are needed so that the very expensive and complex measurement equipment, such as CMM, could be properly used and the desired accuracy of measurement results was ensured. It was found that the lowest level of competence (1-level) requires an employee (called "CMM-User") able to perform actions such as a simple measurement tasks involving: mounting the measured parts on a test bench and preparing the measurement procedure to start as well as operating the machine-controlled software. An employee called "CMM-Operator", whose qualifications were classified as "Level II" competence, should be able to define the measuring task on the basis of technical drawings, create software that controls the measurements, assess the accuracy of obtained measurements results and perform the correction of measurement results, taking into account deviations caused by various factors. The highest, "Level-3" of qualifications is intended for an employee who has been called "CMM-Expert." He should be able to plan, program and optimize the measurement for any established measurement tasks as well as estimate the uncertainty of the measurement results. Additionally, he should know and use quality management methods.

The training concept described above is presented in detail in [11]. This concept has been used to develop the Metrological Knowledge Base, intended for training of CMMs' operators near their workplace, but through the "e-learning" form. Rich experience and valuable results obtained during the execution and implementation of the EUKOM project are good case study showing how the Information System for Metrology should be designed for distance learning of the operators who supervise measurement processes. The first action should be to define the target group (or groups) of operators, supervising a particular type of measurement process. As in each "e-learning" system an essential part is the knowledge base, the structure of such a base should follow the basic assumption concerning the competence of the operator obtained after completion of training. The levels of these competencies should be determined by the professional research and educational centers, which work closely with the industry that uses a particular type of measurement equipment. The structure of a knowledge base, designed in such a particular way, and integrated into appropriately selected training materials, is the primary factor determining the quality of education systems of operators, who supervise measurement processes.

The information technology used in e-learning system for preparing and making available the knowledge base is a secondary matter. Currently, there are many platforms (open source software) for creating “e-learning” systems, for example: [9], [12], [13]. The choice between one of these or one of commercially offered platforms depends on such factors as an expected number of people simultaneously using the “e-learning” system, a kind of supplementary teaching materials that are to be made available to learners. Among other factors taken into consideration are: the ability of teachers to prepare individual learning modules, the possibility of easy modification of the content of the various education modules and the ability to verify the acquired knowledge. In any case, the decisive factor in the quality of Information System for Metrology (ISM) designed for distance learning of operators will be the structure of the metrological knowledge base. It should be developed by the experts in the field of measurements on the basis of agreed level of competence, which the person benefiting from the training and intended to play certain role in the measurement process should reach.

The IT System for Metrology designed to predict the corrections needed to calculate of the Polish Universal Time Coordinated UTC(PL)

Another type of Information System for Metrology (ISM) is an expert system designed for predicting corrections necessary to calculate the Polish Universal Time Coordinated - UTC (PL). This predictive information system for metrology has been continuously improved. Prediction of corrections values is implemented in this system by using various types of neural networks. [14]. This system allows for determining the Polish Official Time OT(PL) introduced in [15], [16] and valid in the Republic of Poland since 2004. This time is calculated as increased by one or two hours referring to the UTC (PL) [17], [18], [19]. It should be determined and maintained with an accuracy not exceeding 10 ns in relation to the universal time coordinated UTC determined by the Time Laboratory of the International Bureau of Weights and Measures in Paris (BIMP). BIMP creates UTC by calculating a weighted average based on systematic comparisons of 300 most accurate frequency and time atomic standards appearing in many countries around the world. BIMP also calculates corrections (PNMI) for each of National Metrology Institutes (NMI) including the corrections to the UTC (PL). In Poland, the Central Office of Measures (GUM) in Warsaw responsible for creation of the values of UTC (PL). The role of GUM in Poland is relevant to NMIs in other countries. The corrections for Poland (PPL) are determined by the BIMP and are published in a special bulletin "Circular T" [7]. These corrections are calculated for the day as:

$$P_{PL(BIMP)} = UTC - UTC\,(PL) \tag{1}$$

The problem is that the values of these corrections are announced only a month after their calculation in Paris. Therefore NMIs around world are forced to use the appropriate method to determine the predicted value of the corrections (PPRED) for the day. Predicting of the values of corrections is based on the historical collection of the corrections for each country published in the “Circular T" bulletin [7]. In different countries various methods of prediction are used. The most common is the analytic prediction method based on the linear regression method extended for stochastic differential equations [20]. Regardless of what prediction method will be applied, it is assumed that the error of such prediction for the day, for the NMI of the each country, should not exceed 10 ns, referring to the value of the corrections that will be published for this country in the bulletin "Circular T" [7]. Hence for Poland this condition can be formulated as follows:

$$\Delta = P_{PRED(PL)} - P_{PL(BIMP)} \ll 10\ ns \tag{2}$$

Although the analytic prediction method described in [20] is widely applied by NMIs and brings good results, it is certainly time consuming and quite difficult. For this reason, the studies [6], [14], [21] have been undertaken to

develop and implement an IT system for metrology, enabling the automatic prediction of the corrections for Poland PPRED (PL). This system has been implemented and tested in the Time and Frequency Laboratory of GUM, responsible for designating and propagation of Polish Official Time OT (PL). This time is established for winter and summer and is calculated as:

$$OT(PL) = UCT(PL) + 1\ hour. \quad or \quad OT(PL) = UCT(PL) + 2\ hours. \tag{3}$$

UTC (PL) is determined with use of the Polish Atomic Time Scale - TA (PL) created on the basis of mutual comparisons of atomic time standards used by over 20 Polish laboratories under and two foreign (Lithuania and Latvia) [19] collaborating under the agreement [17], [18]. The leading laboratory is the laboratory of the Polish Academy of Sciences, Space Research Centre (AOS) in Borowiec [22], for which BIMP separately calculates corrections PAOS (BIMP) and publishes them in the "Circular T" bulletin. AOS also participates in the creation of the Galileo-European Satellite Navigation System [24]. Studies [21], [23] showed that thanks to TA (PL) it is possible to meet the requirements of a specific inequality (2). Using the TA (PL) and the ISM, which automatically predicts the corrections PPRED (PL), GUM creates UTC (PL), which is one of the most accurate national UTCs in the world. Fig.2. shows the block diagram of the ISM using artificial neural networks [21]. This system combined with GUM's selected atomic time standard ensures the creation of UTC (PL) with an accuracy not exceeding 10 ns with respect to UTC time set by the BIMP in Paris.

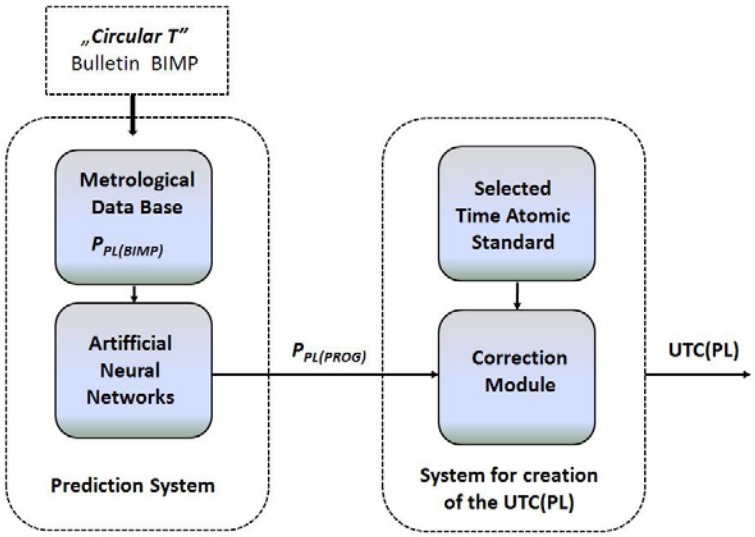


Fig.2. Block diagram of the ISM intended for automatic predicting of corrections PPRED(PL and for creation of the UTC(PL)

The IT System for Metrology intended for Computer-Aided Measurement Processes Validation - "CAMPV System"

Validation is an activity aimed at experimental verification of whether the properties of the method, object, process or software fulfill the requirements for their intended use. The summary of validation must provide the appropriate certificate definitely confirming that the outcome of the validation is positive. In case of the designing process, the goal of validation is to check and confirm that the prototype of the product meets the requirements of the application, for which the product is designed. In case of the measurement process, validation is an experimental verification and confirmation that the measurement characteristics of the measurement process

meet the requirements of its intended use. The full validation cycle of the measurement process - developed and described in [25] - is a set of action, which should result in the issuing of a certificate confirming that the metrological characteristics of the measurement process meet the requirements of its intended use. The full validation cycle consists of 7 specific steps:

1. Identification and characterization of the intended use of the measurement process (IUMP)

2. Determination of metrological requirements for the intended use of measurements (MRIUMP)

3. Choosing the right measurement process (SMP)

4. Determination of metrological characteristics of the selected measurement process (MCMP)

5. Comparison of the metrological requirements of the intended use of measurement process (MRIUMP) with the metrological characteristics of the measurement process(MCMP)

6. Determining result of the comparison and say whether it is positive or negative

7. Execution of one of alternative actions:

7.1. If the result of the step 6 is positive: preparation and generation of validation certificate

7.2. If the result of the step 6 is negative: selection of a different measurement process followed by determination of the measurement characteristics of the new measurement process (step 4) leading steps 5 and step 6.

Fig. 3. shows the flow diagram of the full validation cycle of the measurement process [25].

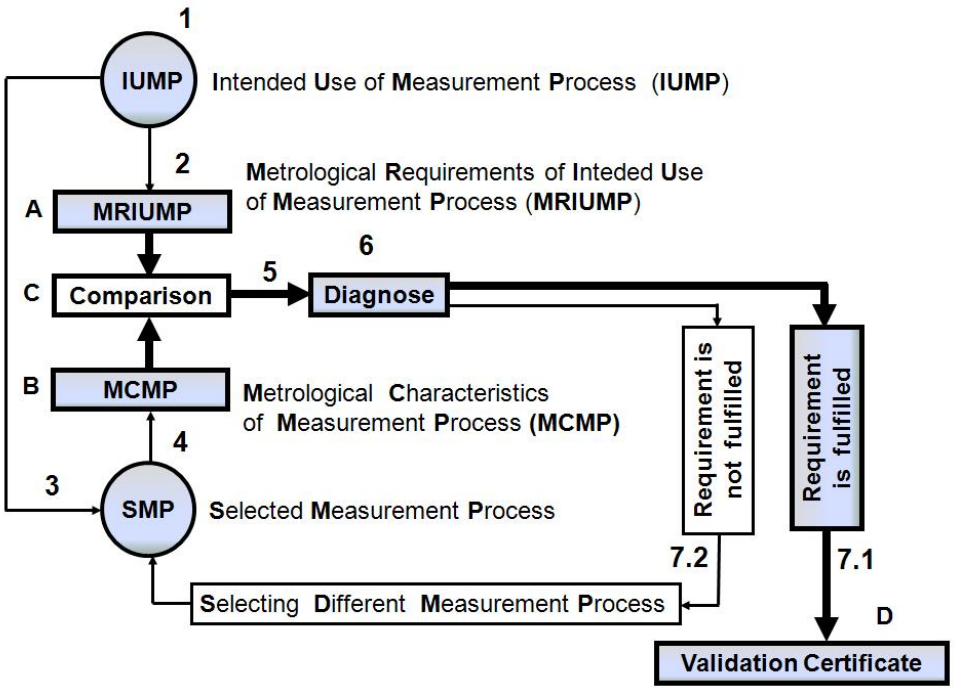


Fig.3. The flow diagram of full validation cycle of measurement processes consisting of 4 phases (A, B, C, D) and 7 specific actions

It should be emphasized, though, that the loop of steps: (7.2) - (4) - (5) - (6) can be repeated as many times as necessary to achieve the positive result of the step 6.

All seven steps are described in details in [25]. They can be grouped into 4 phases:

A – determination of the Metrological Requirements of the Intended Use of the Measurement Processes (MRIUMP),

B – determination of the Metrological Characteristics of the Measurement Process (MCMP),

C – comparison of MRIUMP with the MPMP according to assumed criterion of the validation,

D – generation of the Validation Certificate (VC) confirming that the selected measurement process (SMP) meets the requirements of its intended use.

The measurement process can be accepted and implemented for intended use only if the result of validation is positive. Such rule particularly refers to production quality control, conformity assessment of products or health and environmental protection. In order to execute all steps of the full validation cycle of the measurement process, relevant metrological data must be collected and stored, such as technical specifications of measurement equipment (hardware and software), results of its calibration and results of their statistical analysis, which are the metrological characteristics of evaluated measurement processes.

Having regard to above, the Department of Metrology and Diagnostic Systems, Faculty of Electrical Engineering and Computer Science in Rzeszow University of Technology has been undertaking systematic efforts [25], [26], [27] aiming in development of the most adequate methodology for validation of the measurement processes and creation of the information system, which would support the implementation of this methodology. The outcome of this work is the information system for metrology named "CAMPV-system" (the first part of the name is an abbreviation derived from the full name: Computer Aided Measurement Processes Validation).

Fig.4 shows the structure of the ISM of "CAMPV-system", including website portal providing the online access to the system and the metrological operations database (MOD) designed to collect technical data of the measurement equipment and the results of its calibration. Already designed and implemented modules of the "CAMPV system" are marked by dark background in Fig.4.

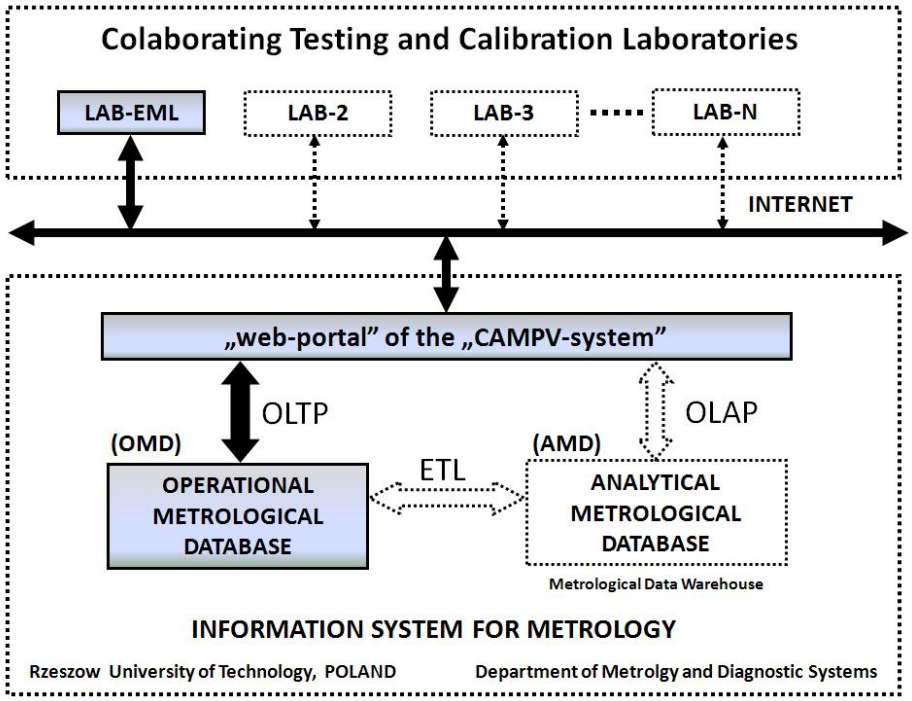


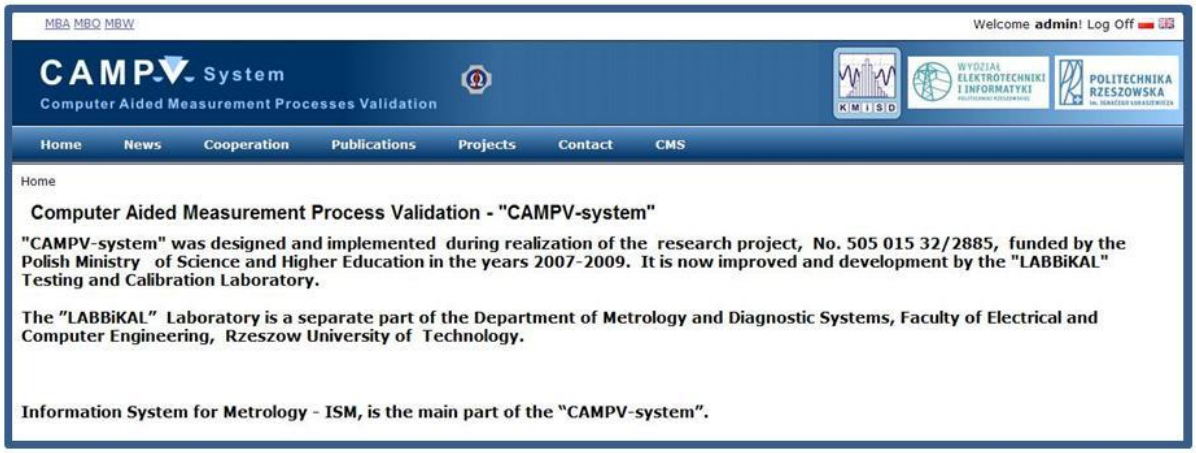
Fig.4. Structure of the CAMPV-system including already developed modules (dark background) and the planned module: Analytical Metrological Database (AMD).

The choice of information technology for respective modules depended on the progress of system's development and the experience gained during the process. Regardless chosen technology, the supreme and indisputable assumption has been as follows:

"ensuring authorized online access to the "CAMPV system" for collaborating laboratories and enabling the extension of system's functionality by adding new modules".

The first prototype of the website portal for the "CAMPV-system" was based on PHP technology, whilst the metrological database used mySQL [28]. The next version of the website portal was based on MS ASP.NET technology and metrological database used MS SQL 2005 technology. The choice of these technologies is explained in [29]. The most important reason for choosing the Microsoft's software environment was to ensure the continuation of the development of the "CAMPV-system". Moreover, the Microsoft Technologies can help in creating of more complex models of metrological databases in the future. The last but not least, Department of Metrology and Diagnostic Systems has had an access to Microsoft's MSDN license, including the license for use of the MS Windows Server. This software ensures compatibility between the created information system for metrology (ISM) and the operating system, which manages the work of the servers.

In the near future the CAMPV-system is expected to be completed by adding the Analytical Metrological Database (AMD), indicated by dotted line on the diagram shown in Fig.4. Such component will ensure an appropriate grouping of the collected data and separating the AMD users from On-Line Transaction Process (OLTP) operations. The AMD also allow for using the Extraction Transaction Loading (ETL) operations to group data in the most suitable form for their future analysis. In order to find this form, the On-Line Analytical Processing (OLAP) has to be determined. Having above in mind, the future version of "CAMPV-system" will have an adequate structure, as presented in Fig. 4. CAMPV system has recently been improved by using newer versions of Microsoft technologies. The website portal has been rebuilt with use of Model View Controller technology (MVC v.3.), with the RAZOR engine and MS SQL 2008 environment for creating databases. This technology allows for dynamic changes in the structure of the menu names of the website portal as well as tabs and content published through Content Management System (CMS). It also allows for creating different language versions. Following this improvements, the website portal has become a module that not only can be quickly rebuilt but also easily expanded with use of the Microsoft Silverlight technology and WCF RIA Services. Fig.5. shows one of the portal screens appearing after selecting the "home" button.



Rys.5. The view of the website "home" of the website portal of the "CAMPV-system"

Operational Metrological Database (OMD) is currently being reconstructed with use of MS Silverlight technology. The advantage of this solution is that the User Interface (UI) of the webpage is loaded only once, and the data used to fill the forms are loaded later during work with the server. Other innovations are planned to be implemented in the future, including drag&drop technology applied for creating single measurement channels consisting of elementary modules or creating a complex measurement process with use of several single measurement channels.

Fig.6. shows the design of the future OMD website. It provides access to one of the modules of the OMD, which collects technical data of measurement equipment and loads files with calibration results of measurement equipment. This module - named Applied Measurement Processes (AMP) - will also be equipped with features such as creating measurement processes, collecting the results of the repeatability and reproducibility analysis as well as collecting the results of interlaboratory comparisons. The results of calibration loaded into the OMD are saved in the Excel 2007 files. In the future other formats are expected such as .csv and .xml.

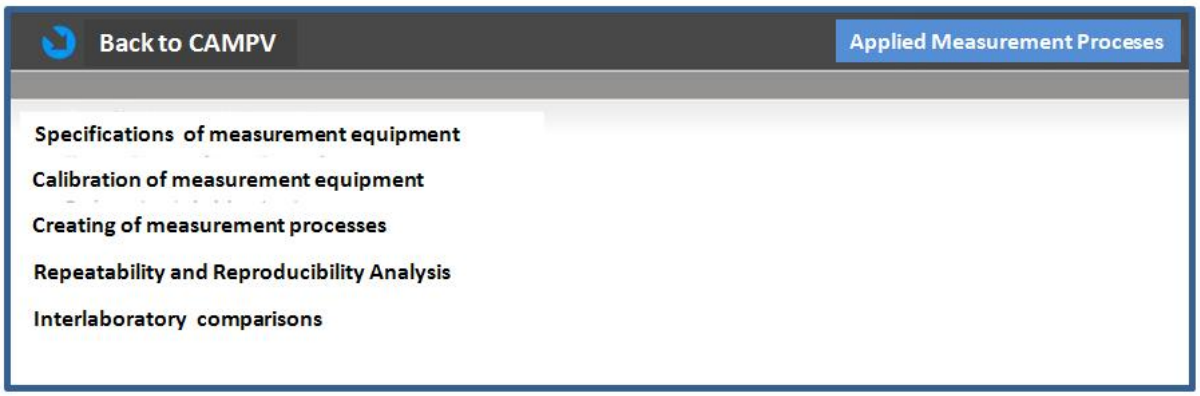


Fig.6. The design of the future Operational Metrological Database (OMD) website accessing Applied Measurement Processes (AMP) module of the "CAMPV-system"

Conclusion

Three examples of different Information Systems for Metrology (ISM) are described. The first one has been designed to train operators of coordinate measuring machines (CMM) with use of the "e-learning" system. A Metrological Knowledge Base (MKB) structure plays an important role in this type of ISM. Such MKB contains elementary teaching modules tailored to the level of competence of the operators, who improve their measurement skills. The second system has been designed to predict the corrections, which allow for creation of the Polish Universal Time Coordinated UTC (PL). The key module in this type of ISM is the structure of the artificial neural network predicting the corrections values. The third system has been designed for the computer-aided measurement processes validation. In this type of ISM the major role plays an open and modular structure, accessible through the website portal and ready for a continuous expansion. This ISM called "CAMPV-system" has been tested and developed in the Department of Metrology and Diagnostic Systems, Faculty of Electrical and Computer Engineering of the Rzeszow University of Technology. The target version of "CAMPV-system" will include two databases and a knowledge base. This form may become the basis for creation of the metrological expert system "CAMPV-EXPERT-system" supporting the validation of the measurement processes using artificial intelligence methods.

Based on above-described examples, one can conclude that the information technology contribute significantly to the development of the metrology and its applications. A key condition of successful creating of the high quality IT systems for the metrology is a close cooperation between scientists and practitioners in the field of computer science and metrology.

Bibliography

[1] [Finkelstein-1982] Finkelstein L.: Theory and philosophy of measurement, In: Sydenham P.H. (ed.) Handbook of Measurement Science, Chichester: Wiley, 1982, pp. 1-30.

[2] [Muravyov-1997] Muravyov S.V., Savolainen v.: Towards describing semantic aspect of measurement. Proceedings of the XIV IMEKO World Congress, Tampere, Finland, 1997.

[3] [Stevens-1946] Stevens S.S.: On the theory of scales of measurements. SCIENCE.

[4] [Finkelstein-2003] Finkelstein L.: Analysis of the concept of measurement, information and knowledge. Proceedings of the XVII IMEKO World Congress. Dubrovnik. Croatia.2003. pp. 1043-1047.

[5] [Keferstein-2007] Keferstein C.P., Marxer M.: EUKOM-European Training for Coordinate Metrology. Proceedings of the 13 International Metrology Congress. Lille. France. 2007.

[6] [Cepowski-2009] Cepowski M, Miczulski W.: Metody prognozowania państwowej skali czasu. Materiały VII Konferencji Naukowo-Technicznej Podstawowe Problemy Metrologii. Sucha Beskidzka. 2009. ss.12-15

[7] Circular T". Bulletin. (ftp://ftp2.bipm.fr/pub/tai/publication/cirt/)

[8] www.cm-train.org

[9] http://www.ilias.de

[10] Research project: „European Training for Metrology". University Erlangen-Nuremberg. Chair Quality Management and Manufacturing Metrology. (in polish version: www.lm.ath.bielsko.pl)

[11] [Weckenmann-2004] Weckenmann A., Jakubiec W., Płowucha W.: Training concept EUKOM. Leonardo da Vinci pilot Project. European Training for Coordinate Metrology. D/02/B/F/PP 112 662 EUKOM. Erlangen. 2004.

[12] http://moodle.org

[13] http://www.openelms.org

[14] [Miczulski-2011] Miczulski W., Sobolewski Ł.: Wpływ sposobu przygotowania danych na wynik prognozowania poprawek dla UTC(PL) z zastosowaniem sieci neuronowej GMDH. Materiały z VIII Konferencji Naukowo-Technicznej. Podstawowe Problemy Metrologii. Krynica. 2011.

[15] Ustawa z dnia 10 grudnia 2003. O czasie urzędowym na obszarze Rzeczypospolitej Polskiej (Dz. U. z 2004. Nr16, poz.144.)

[16] Rozporządzenie Ministra Gospodarki, Pracy i Polityki Społecznej z dnia 19 marca 2004, w sprawie sposobów rozpowszechniania sygnałów czasu urzędowego i uniwersalnego czasu koordynowanego UTC(PL).

[17] [Czubla-2006] Czubla A.,Konopka J., Nawrocki J.: Realization of atomic SI second definition In context UTC(PL) and TA(PL). Metrology and Measurement System. No 2, 2006, pp. 149-159.

[18] [Nawrocki-2006] Nawrocki J., Rau Z., Lewandowski W., Małkowski M., Marszalec M., Nerkowski D.: Sterring UTC(AOS) and UTC(PL) by TA(PL), In Proceedings of the Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 2006.

[19] [Marszalec-2011] Marszalec M., Lusawa M., Nerkowski D.: Wyniki badań algorytmów zespołowych skal czasu z wykorzystaniem bazy danych dla TA(PL). Materiały z VIII Konferencji Naukowo-Technicznej. Podstawowe Problemy Metrologii. Krynica. 2011.

[20] [Panfilo-2008] Panfilo G., Tavella P.: Atomic clock prediction based on stochastic differentia equations. Metrologia 45, 2008, pp.108-116.

[21] [Cepowski-2009]] Cepowski M., Miczulski W.: Zastosowanie sieci neuronowych do prognozowania państwowej skali czasu. Materiały z XVII Sympozjum Modelowanie i Symulacja Systemów Pomiarowych. Krynica 2009.

[22] Polska Akademia Nauk. Centrum Badań Kosmicznych. Obserwatorium Astrogeodynamiczne w Borowcu (http://www.cbk.pl)

[23][Miczulski-2010] Miczulski W., Cepowski M.: Wpływ sieci neuronowej i sposobu przygotowania danych na wynik prognozowania poprawek UTC- UTC(PL). Pomiary Automatyka Kontrola vol.56, nr 11, 2010, ss.1330-1332.

[24] Polski Punkt Informacyjny GALILEO (http://galileo.kosmos.gov.pl)

[25] [Tabisz-2010] Tabisz R.A.: Walidacja procesów pomiarowych. Przegląd Elektrotechniczny (Electrical Review). R86 Nr 11a, 2010, ss. 313-318

[26] [Tabisz-2009] Tabisz R.A.: Validation of Industrial Measurement Processes. Proceedings of the 13th International Metrology Congress. 18-21 June. Lille. France. In the book: Transverse Disciplines in Metrology. ISTE-Wiley. 2009 pp.791-801.

[27] [Tabisz-2009] Tabisz R.A.: Computer Aided measurement Process Validation. Proceedings of the 14th International Metrology Congress, 22-25 June, Paris. 2009.

[28] [Adamczyk-2007] Adamczyk K., Tabisz R.A.: Zastosowanie wybranych technologii informatycznych do tworzenia Metrologicznej Bazy Danych. Pomiary Automatyka Kontrola. Vol.53. nr 12, 2007, ss. 51-54.

[29] [Świerzowicz-2008] Świerzowicz J., Adamczyk K., Tabisz R.A.: Kluczowe etapy tworzenia Informatycznego Systemu Metrologicznej Bazy Danych przeznaczonego dla sieci laboratoriów badawczych i wzorcujących. Pomiary Automatyka Kontrola. Vol.54, nr 12, 2008, ss. 869-873.

Authors' Information

Roman Aleksander Tabisz *– Department of Metrology and Measurement Systems, Faculty of Electrical and Computer Engineering, Rzeszow University of Technology. W. Pola.2. 35-959 Rzeszow, Poland. e-mail: rtabisz@prz.edu.pl*

Major Fields of Scientific Research: Industrial metrology, measurement processes validation

Łukasz Walus *– The owner of the IT enterprise: SolvSoft. Łukasz Walus. 36-230 Domaradz 460A, Poland, e-mail: ukaniow@gmail.com*

Major Fields of Scientific Research: Business applications of the IT Technology.

Business Intelligence Systems

BUSINESS DISCOVERY – A NEW DIMENSION OF BUSINESS INTELLIGENCE

Justyna Stasieńko

Abstract. *This article deals with the issue of Business Intelligence (BI), especially its next generation - Business Discovery(BD). This tool can help to fill the gap between traditional solutions we get from BI and standalone office productivity applications. Its users are able to forget new paths and make new discoveries. Here we want to present BD as being thoroughly complementary to traditional ERP, CRM, BI, and data warehousing systems. BD brings a whole new level of analysis, insight and value to the information stored within these systems. What is more, its users are not burdened with interfaces which are difficult to use and configure.*

Keywords: *Business Intelligence, Business Discovery, information, analysis, Qlickview*

ACM Classification Keywords: *K.6 Management of Computing and Information Systems - K.6.0 General* Economics

"Computers are useless. They can only give you answers."

Pablo Picasso

Introduction

Currently, one can notice a sea-change approach to the role and importance of information. So far, information has often been treated as a by-product or, at best, co-implemented business processes. Now the information is one of the most important organising resources. The information is a factor that increases the knowledge about the reality surrounding a man, or specific intangible asset, which, with economic progress and development of means and forms of social communication, is becoming more and more important, transforming the face of many traditionally organised economies of the world.

The importance of information in the modern world should be analysed in various aspects. One can even attempt to say that in a sense, the information is the engine of progress in every area of human life. Following an increasing pace, technological development to a great extent depends on the speed and the quality of information. Thus, the access to information should be easy, and the way it is used should present its values.

In order to achieve business benefits the strategic decisions are taken. These decisions result from the study of gathered information. For hundreds of years experience had been mainly a source of information. They were mainly mathematics and mathematical models, then the statistical models, econometric, and now the Internet.

A growing number of data, a flow of information, fast communication via the Internet and new market challenges mean that data analysis, broadening knowledge and skilful decision-making in an increasingly complex market are essential to survive and run business. Solutions in an enterprise operate at many levels and play various roles supporting the administration or management, yet they have their own limitations. Especially in the

processing of a large number of diverse data and in the use of information in many fields. In order to meet the demands BI solutions have been created .

Business Intelligence is now the most important and inevitable point of contact between sciences and business. Business Intelligence tools enable an easy access to information, its analysis and sharing across the organisation and its business environment. They give the possibility to integrate data from different sources and their comprehensive analysis in terms of business needs. BI gives a preview of all business organisations. Their goal is to support effective business management and business planning by providing the right information. They support the work of managers in managing key areas of business. Most generally, one can present it as a process of transforming data into information and the information into knowledge that can be used to enhance the competitiveness of enterprises. Among these tools there are both management systems of information resources, reporting and analysis tools, and also solutions enabling to boost managing performance.

Information in the organisation is generated and processed mostly in the transactional software, such as ERP, CRM. These systems have evolved over the years. The first stage of its development was automating tasks and processes, reporting and logging. This has contributed to an increase in the information registration. The second phase was the emphasis on resource planning and the efficiency of business processes. For better reporting and data analysis data warehouses have been introduce– an aggregation of transactional information and multi-dimensional analytical models. Due to the need of a larger amount of information for analysis and its presentation in a friendly manner, tools in the field of decision support proved to be necessary. The solution which came out to meet the problems of modern organisations are Business Intelligence Systems (BI). Traditional BI technology use data warehousing, analytic tools (OLAP and Data Mining), and presentation techniques. They enable to optimise operations, to increase efficiency, to reduce the risk of taking wrong decisions, as well as to reduce costs and to maximise profits.

Analytical tools

Analytical tools existing on the market support a decision maker by providing him/her with the necessary knowledge - in the form of reports based on historical and current data - to make decisions. They allow the standard and advanced reports using statistical analysis, forecasting, relationship between the data search, research trends [Nycz, 2008].

The basic analytical tools include query generation tool and reporting (Query & Report - Q & R), spreadsheets, OLAP mining and data visualisation tools [Dudycz, 2004].

Query and reporting tools are the most basic tools for data analysis, in particular, gathered in the data warehouses. There are two types of reporting: the standard and so called 'ad hoc'. Frequently they answer the questions "what happened?", "What level of sales was as in the previous year?" etc. In the second half of the 1990s, it was noted that the data stored in databases, transactional systems used in companies caused a lot of trouble to analysts who carried out the assessment of a business enterprise. This problem was solved by introducing analytical techniques implemented, inter alia, in spreadsheets [Dudycz, 2004], which enable to create models that generate periodic reports automatically [Sierocki, 2007]. Spreadsheet offers flexibility when it comes to the definition of the conditions of analysis and ease of use. The difficulty arises when the basis for analysis are large volumes of data or high complexity of the model. In order to achieve the desired analytical flexibility a wide enough diagram of processing must be built, often based on large amounts of macro-commands and sheets. This solution cause difficulties of managing the data, and at the same time it is prone to user's errors. Sheets have a limited working capacity for the data, which almost eliminates their usefulness for the analysis of large portions of data, reaching hundreds of thousands of transactions.

Facing new challenges, a concept of multidimensional databases and OLAP technology emerged, which allowed for a dynamic and multidimensional analysis of business data.

OLAP architecture is encountered in data warehouses and tools for analysis such as query languages, data mining, artificial intelligence, as well as report generators. Through proper presentation, visualization and aggregation, it enables to display and view data from different points of view allowing its user to examine them quickly. In addition, it is characterized by the possibility of an interactive reporting without knowledge of programming languages and to obtain answers to complex and often non-standard queries in a current mode. Therefore, OLAP tools are often used to perform analysis of sales trends, financial analysis (data warehouse), or to pre-screen the data set by the analyst in the initial phase of statistical analysis [Sierocki, 2007].

Data visualization tools are designed to increase transparency and legibility of presented information. Most of the analytical tools offer simple dependence images between the data.

Existing tools of data acquisition and processing of analytical reports, such as generators or spreadsheets were not able to fully meet the needs of managers who have to make a relatively fast growing in-depth analysis.

Facing these challenges, a concept of multidimensional databases and OLAP technology was introduced, HOLAP, MOLAP and ROLAP, which allow for a dynamic and multidimensional analysis of business data. Architecture of the OLAP (MOLAP, DOLAP, ROLAP, HOLAP) may be encountered in data warehousing and data analysis tools such as generators, reports, query languages, data mining, artificial intelligence. It enables to display and view the data from different points of view allowing its user to examine them quickly through an appropriate method of presentation, visualization and aggregation.

OLAP systems are characterized by the possibility of:

- perform multidimensional analysis according to complex search criteria,
- interactive reporting without knowledge of programming languages,
- obtaining answers to complex and often non-standard (so called 'ad hoc') queries in a current mode.

Therefore, OLAP tools are often used to perform analysis of sales trends, financial analysis (data warehouse), or to pre-screen the data set by the analyst in the initial phase of statistical analysis.

In order to deepen the analysis and discovery of repetitive behaviours in large data sets through data mining, matching various models and relationships between data analysis, special methods are used for data mining. Data mining is a methodology that refers to a technique derived from mathematical statistics and machine learning algorithms. Information extracted by using these tools can be used in areas such organisations as the support of decision making, forecasting, financial analysis and risk analysis, optimisation.

A frequently used data mining tool is a universal, integrated system for statistical data analysis - STATISTICA. This software not only contains statistical and graphical procedures for general use, but also powerful tools for analysis and visualisation of data, as well as specialised analytical techniques (e.g. social studies, biomedical, or technical) [Dudycz, 2004].

Other tools to cope with the analysis of the large amounts of data processed into information, and then into knowledge are BI systems.

BI systems should improve the management of knowledge in an organisation at the three levels presented in Table 1

Table 1. Tasks of Business Intelligence Systems

Management level	BI tasks
Operating	Analysis carried out ad hoc, information on current operations, finances, sales, collaboration with suppliers, customers, clients, etc.
Tactical	Fundamentals of decision making in marketing, sales, finance, capital management. Optimising future actions and modification of financial factors, technology in the implementation of strategic objectives.
Strategic	Precise setting of goals and tracking their implementation, to perform various comparative statements, conducting simulation development, forecasting future performance under certain assumptions.

Source: Own elaboration.

The new generation of Business Intelligence systems

Business Intelligence turns out to be the new quality in the management conception. BI systems are used in order to create and improve the relationship with a customer, yet at the same time to boost management effectiveness. Unfortunately, traditional BI software seems to have failed, as far as delivering on this vision is concerned, as a result of its complexities, time lags, and expensive professional services requirements.

Forrester Research's definition of BI is "a set of methodologies, processes, architectures, and technologies that transform the raw data into the meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making." This definition of BI covers the whole data-to-insight process (including data preparation). It all appears to be time-consuming, especially to plan and implement - from collecting requirements, to building a data warehouse, to populating a metadata layer. Traditional BI software users have problems to learn and use it, and, in consequence, adoption is limited. Moreover, distribution of information and analytic tools is tightly controlled.

Traditional BI systems are high cost and IT driven. They are chosen, installed, and maintained by IT organisations and in most cases not by the business people themselves who will use them later on. Owing to the complexity of the system, not many people (in an organisation) feel skilled enough to form business insights. When business analysts and IT professionals want to be sure that they deliver the right analysis, the back-and-forth questions and answers with their business constituents make it really difficult for them. Traditional BI usually constitutes more or less centralised, pre-packaged reports or predetermined queries which users can run to get updated numbers. It often happens that the information we get is mostly static. Therefore, users having a question which is outside the standard configuration need to log a ticket with IT and expect their assistance (sometimes weeks or even months). The drawback of BI is the fact that it is centralised, possessed by IT, difficult to change or modify, and slow to deliver results. What is more, it is also expensive and highly complex.

In addition, if one wants to deploy a traditional BI solution, it can take him/her up to a year and a half. By and large,it seems to be due to requirements gathering and data modelling and integration efforts. This is definitely not what the business needs as an organisation can live and die within this time. Furthermore, traditional BI requires lots of services and support in order to keep its various components working smoothly.

BI in the company combines finance, manufacturing, warehousing, logistics, purchasing, sales, HR, planning and strategy - in short, all aspects of a company. Therefore, BI uses a common repository of information - Data Warehouse. All facts come from individual branch systems through ETL processes converted into information and

stored in the DSA - Data Staging Area data warehouse. From this information the system uses the second part of the Business Intelligence which converts this information into knowledge and provides the user through the presentation layer. Thanks to the class Business Intelligence supports managers effectively, and enables, inter alia, building a What-If analysis, budgets and controlling systems.

Traditional BI turns out to be excessively bloated and rigid. Further evolution of these systems will lead to the revival of petrified BI. During this evolution Business Discovery (BD) has emerged. Trying to answer why BI platform displaces BD, one would indicate the four trends that have caused the evolution of the BI software market. The first one is the ability to search the Internet and to obtain a rapid response. The second trend is the community network (social networking) that enables to communicate, to share information and to develop robust, professional and personal networks (with no requirement of technology background). Other trends are the development of mobile and task-specific applications. In conclusion, BD are much faster, open and straightforward at the same time, mobile and addressed for everyone.

Business Discovery platforms (offering new solutions), as opposed to its ancestor, can have a total cost of ownership that is half that of other BI solutions. Business Discovery is a whole new way of doing things for Business Intelligence.

Concept of Business Discovery

If we decide to store data in-memory, it means we no longer have to deal with a database located somewhere else, and we receive no queries and no retrieval. Thanks to it, there is no delay in returning results, whatsoever.

Business Discovery (BD) is a complement to traditional BI, ERP, CRM, and data storage systems. It also introduces
a new level of analysis, knowledge and value of information that fall within these systems. Additionally, it became a response to the unmet needs of users as it is a new way of doing BI. It is a bottoms-up approach that puts the user in control, fulfilling the promise of BI. The main aim of BD is to help users to solve specific business problems in a timely way to get answers to the most critical questions and also to share knowledge and analysis among individuals, groups, and even organisations.

Most essentially, users are capable of gaining insights that address their individual needs at every level of the organisation. It seems that they are not limited to particular paths they must follow, or questions they are obliged to formulate in advance. The key issue is the fact that they can ask about what they need.

Business Discovery gives an opportunity of a whole new level of analysis, insight, and value to existing data stores with user interfaces that are clean, simple, and straightforward. It is complementary to traditional BI software and other enterprise applications. BD is enriched with new opportunities for BI. The most vital one would include an application interface (insight for everyone), the time to provide results of the analysis, mobility of applications, remixability and reassembly, and finally social and collaborative environment.

Importantly, everyone can create an insight by means of Business Discovery. It's the equivalent of open source computing or peer creation. This is definitely intelligence creation — rather than just information consumption. First of all, Business Discovery is not a large collection of centrally-controlled, pre-packaged, and tightly-distributed data. Secondly, rather, it provides data access and analysis to individuals and groups, and allows them to get what they ask for faster and more accurately than ever before.

BD enables instant analysis. The user receives the results at the bedside, where it lasted weeks with the traditional BI. It has a direct access to all necessary data. Technology can ask any questions to which answers are kept on-line.

At all levels in an organisation business decision makers need data at their fingertips, wherever they are. They want to work how and where they like - whether that be in the warehouse, on the customer site, or on the trade show floor. Tablets and other large-form-factor mobile devices promise to make business data ubiquitous. Unlike traditional BI solutions, Business Discovery platforms provide an intuitive interface and an application infrastructure that is tailor-made to exploit the opportunity of a truly mobile, well-informed workforce.

BD supports mobile applications. This allows the provision of data while using mobile devices such as iPhone, Android.

Business Discovery platforms empower anyone to quickly develop and deploy simple, focused, and intuitive applications that can be easily reused. These applications are easy to modify, mash up, and share, allowing innovation to flourish at the edges of the organisation and spread inward. The new opportunity is leveraging a model that lets any user quickly develop and deploy task-specific, purpose-built BI applications. BD platform enables its users to quickly create and implement among others their own applications. These applications are able to quickly solve specific problems.

Nobody can predict what questions business users will have when they start exploring data — not even the users themselves. Traditional BI solutions require IT or power users to get involved whenever new questions arise. In contrast, Business Discovery platforms make it easy for business users to remix and reassemble data in new views and create new visualisations for deeper understanding. With BD, users generate insights like never before.

BD makes it easy to "remix" and reassemble data to the new views (previews) and a fast way to create visualisations.

BD is a social and collaborative environment. It enables its users to share and collaborate on insight and analysis. They can share insights within Business Discovery apps or through the integration with collaboration platforms. Business Discovery is about creating a community of users who engage in wiki-like decision-making to drive knowledge that can cascade across an organisation.

QlikView BI in-memory

QlikView is a modern and innovative approach to Business Intelligence. What adds value to the existing QlikView BI applications include: making the process of assembling, associating, and preparing data for analysis simple and straightforward, allowing users to interact with data in the way they think-associatively. A great advantage of QlikView is that data is collected in memory, which improves analysis and convenience to use.

QlikView was built with a simple architectural premise. All data should be held in memory, and all calculations should be performed when requested and not prior. The goal was to deliver the powerful analytic and reporting solutions in a quarter of the time, at half the cost and with twice the value of competing OLAP cube based product.

QlikView is the world’s first associative, in-memory BI platform. It manages associations among data sets at the engine level, not the application level, by storing individual tables in its in-memory associative engine. Every data in the analytic dataset is associated with every other data point in the dataset.

Associative cheese means finding answers to questions, but also the questions that have not yet been started. What is meant by a simple application is creating questions that do not require knowledge of creating queries in SQL. This associative experience gives decision makers a better overview of their business.

Visualisation is of dual significance. Firstly, it refers to the visual display of summarised forms of information. Secondly, the ability to see those displays change as the selected date is changed. QlikView offers various ways

of data presentation: graphs, charts, tables and others. It also enables its users to create different types of measures that enhance the analysis process. It provides flexible, intuitive and powerful data visualizations.

Using the QlikView application is pleasant and not too complicated (Fig.1)

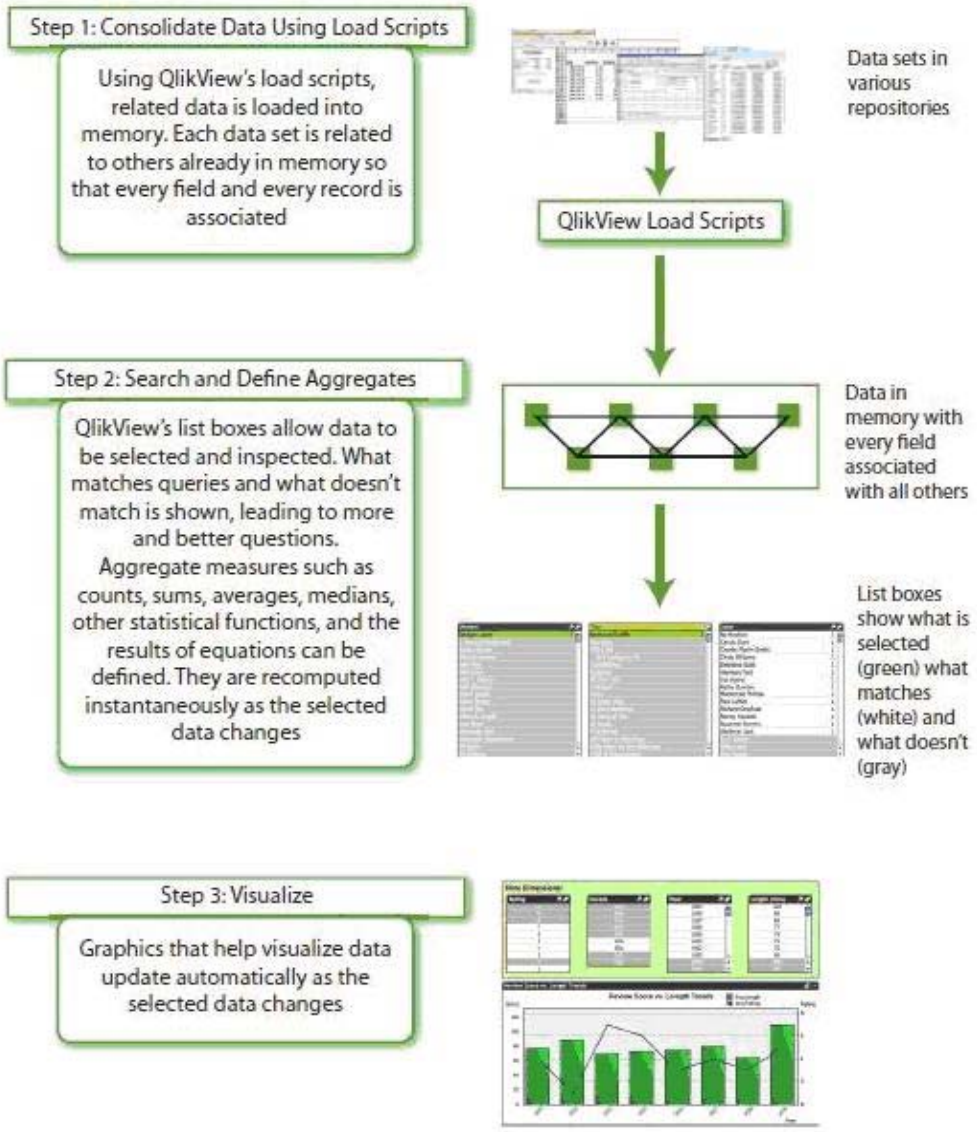


Fig. 1. A Step-by-Step Tour Through QlikView. Source: [The Art of Business Discovery p.8]

QlikView stores data in memory instead of retrieving it from databases and OLAP cubes, it cannot only display what is included in your search, but also what is excluded as well. This is certainly a real departure from traditional BI (i.e. for the first time, unknowns become known).

Compared to other analytical tools or BI based on the OLAP 'cube', an application of QlikView present the analysed data to the user in a faster and simpler way. QlikView gives a chance to impose criteria onto the data. In a straightforward and quick way, QlikView gives the opportunity to return to the previous data or to add further ones upon existing criteria. Due to its simplicity and ease to use, a user-friendly and attractive QlikView interface is a modern and highly efficient application. It has also, not available for other solutions, the time of submission of new studies, computing power and flexibility. Software flexibility leads to the lack of restrictions on the number

of dimensions and measures, and its power – to virtually immediate response to inquiries from the system, even with databases of up to five hundred million records. It also provides the possibility of an immediate transition to a single transaction.

QlikView enables to view data from different perspectives (Fig. 2). Indeed, some analysis can also be done, for example in a spreadsheet, but QlikView provides instant viewing of data by analysing different criteria.

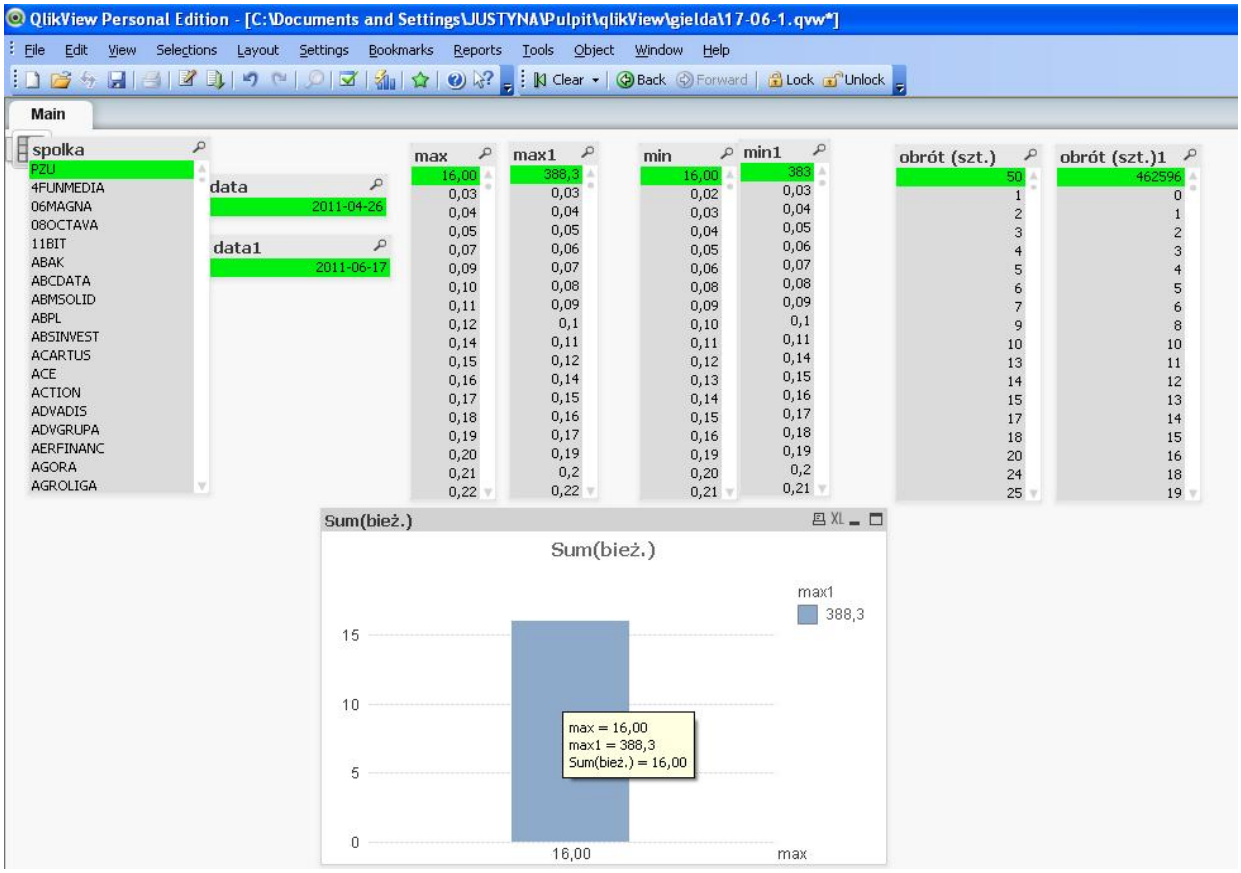


Fig. 2. Comparison of speed, start and end of a listed company PZU session held on 26.04.2011 and 17.06.2011
Source: Own elaboration

It is not necessary to switch from a previous analysis (Fig. 2), all one needs to do is to simply click or select the appropriate criteria and get a completely different analysis (Fig. 3) that no longer applies to the company, PZU, but only to the same level of stock prices such as 0.19 PLN during the trading session on 17.06.2011. This is a great help because another analysis from the beginning is not required. While working on some data one can analyse them in many ways without switching between windows, sheets, etc.

QlikView software enables to transfer analysis results, which remain still just as functional, onto hardware. The application allows printing of the results in the form of reports, exporting them to MS Excel or saving as PDF. It is practically capable of integrating all data formats - from standard relational data into text reports, the data from Excel and XML streams. QlikView is quickly and easily deployed and integrated with existing enterprise systems.

Table 2 shows the advantages of QlikView applications compared to traditional BI systems.

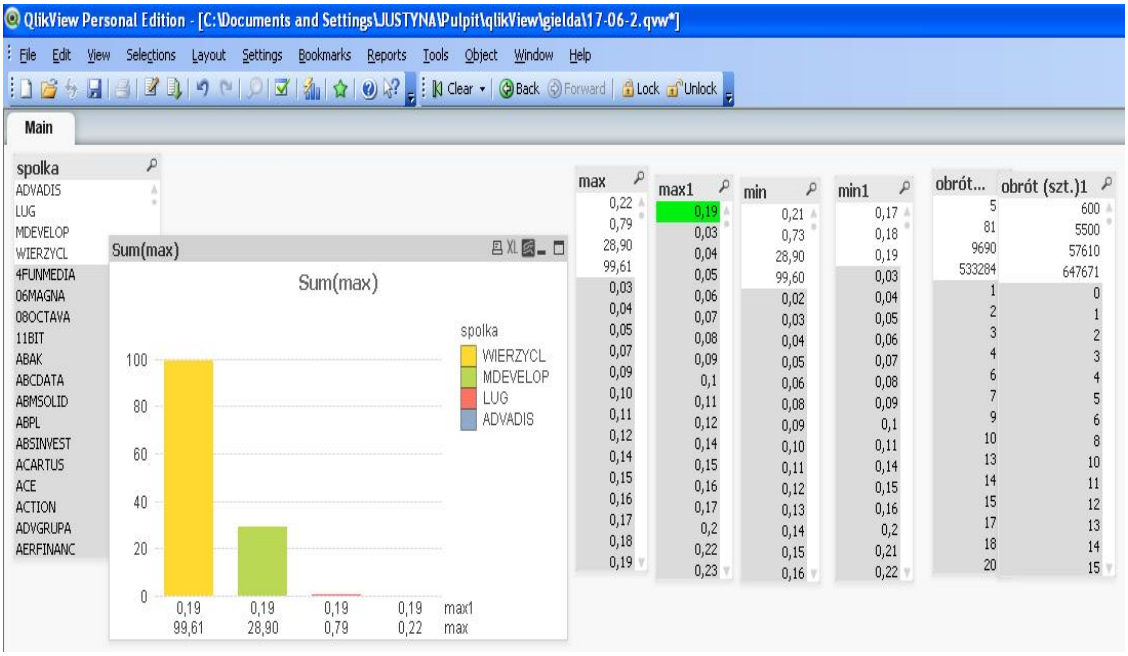


Fig. 3. Maximum price for shares at the 0,19 PLN level on June 17th, 2011. Source: Own elaboration

Table 2. The differences between traditional BI and in-memory BI systems, such as QlikView

Traditional BI	QlikView
Lots of tools: data warehouse and data marts, OLAP, query and reporting tools, data mining	Simple architectural premise - all data should be held in memory
Presentation techniques of analysis (dashboards, scorecards, reports)	Presentation techniques of analysis (charts)
Many users	One user
Longer time of implementation (approximately 18 months)	Short time of implementation (several weeks)
High costs of implementation	Low costs of implementation
Less convenient and flexible for user	Easy to use, flexible
Time-consuming and complex process of information processing	Fast query and on demand calculation engine

Source: Own calculation.

At the end of 2010, there is another version (QlikView 10). A new feature is the ability to deploy the software on any platform: local, cloud computing and mobile devices. The release of the platform in the cloud is via Amazon's Elastic Compute Cloud (EC2). It has also enormous power to allow flexible analytical processing of large data sets while maintaining access to the details. It is not limited by any number of dimensions. Changes in the designed applications can be performed quickly and without possessing advanced knowledge of programming. The application offers new opportunities to present the results of studies using AJAX. New ways to visualise data facilitate the understanding of the data presented. Searching for information through associative search capabilities has been improved as well.

Summarising QlikView in-memory analysis and reporting is simplifying analysis for everyone and has clearly demonstrated its affordability and value to organizations across industries for solving their performance and information challenges.

Conclusion

BI systems have existed / operated in the market for about two hundred years. Transforming data into information and information into knowledge enabled business decisions, allowing users to make effective and informed choices based on data analysis. Technology development in this field has also brought change. It turned out that traditional BI systems are too complicated, delayed and requires professional services, which are costly. Therefore, the direction of change went toward cheaper and faster applications. Thus, there arose the Discovery Business Systems which are a whole new way of doing things for Business Intelligence.

BD bridges the gap between traditional BI solutions and standalone office productivity applications, enabling users to forge new paths and make new discoveries. BD works with what you have and infuses new capabilities into BI: insight for everyone, zero-wait analysis, mobility, an app–like model, remix ability and reassembly, and a social and collaborative experience.

Bibliography

[Dudycz, 2004] Dudycz H. Przetwarzanie analityczne podstawą rozwiązań informatycznych klasy Business Intelligence. Materiały konferencyjne SWO, Katowice, 2004. http://www.swo.ae.katowice.pl/_pdf/138.pdf

[Januszewski, 2008] Januszewski A. Funkcjonalność Informatycznych Systemów Zarządzania, t.2 Systemy Business Intelligence, PWN, Warszawa, 2008.

[Nycz, 2008] Nycz M., Smok B. Busienss Intelligence w zarządzaniu. Materiały konferencyjne SWO, Katowice, 2008. http://www.swo.ae.katowice.pl/_pdf/421.pdf

[Olszak,2005] Olszak C. Wspomaganie decyzji w erze informacji i wiedzy. W: Systemy wspomagania organizacji. Praca zbiorowa pod redakcją H. Sroki i T. Porębskiej-Miąc. AE, Katowice, 2005, s.346-353.

[Sierocki, 2007] Sierocki R.: OLAP to efektywna technologia przetwarzania danych analitycznych. Controlling i Rachunkowość Zarządcza nr 1/2007. http://www.infosynergia.eu/pliki/publikacje%20-%20Robert%20Sierocki%20-20OLAP%20to%20efektywna%20technologia.pdf

[Surma, 2007] Surma J. Wsparcie strategicznych decyzji zarządczych z wykorzystaniem Business Intelligence. Innowacyne Aspekty Strategii Przedsiębiorstwa pod red. Lidii Nowak. Politechnika Częstochowska, Częstochowa, 2007. http://www.surma.edu.pl/wp-content/czestochowa-jerzy-surma-2007.pdf .

[Surma 2009] Surma J.Business Intelligence – Systemy Wspomagania Decyzji Biznesowych, PWN, Warszawa, 2009.

RELATED WHITE PAPERS

The QlikView Product Roadmap, December, 2010

http://www.qlikview.com/us/explore/resources/whitepapers/qlikview-product-roadmap

The Associative Experience: QlikView's Overwhelming Advantage, October, 2010

http://www.qlikview.com/us/explore/resources/whitepapers/the-associative-experience

QlikView Architectural Overview, October, 2010

http://www.qlikview.com/us/explore/resources/whitepapers/qlikview-architectural-overview

http://community.qlikview.com/blogs/theqlikviewblog/archive/2010/09/21/software-should-fit-the-business-not-the-other-way-around.aspx

A QlikView White Paper; *Business Discovery: The Next Generation of BI*; Published January, 2011; http://www.qlikview.com/us/explore/resources/whitepapers/business-discovery-the-next-generation-of-bi

A CITO Research Explainer: *The Art of Business Discovery*; May 2010

http://www.qlikview.com/us/explore/resources/whitepapers/the-art-of-business-discovery [March, 2011]

A QlikView Technology White Paper: *QlikView Architectural Overview* Published October, 2010 http://www.qlikview.com/us/explore/resources/whitepapers/qlikview-architectural-overview [March, 2011]

A White Paper from QlikTech International AB: *QlikView 7. The In Memory Business Intelligence Revolution.* http://www.businessintelligence.bg/files/presentations/Intro_to_QlikView_061206.pdf [February, 2011]

Authors' Information

Justyna Stasieńko – *lecturer, The Institute of Technical Engineering, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail: justyna.stasienko@pwste.edu.pl*

Major Fields of Scientific Research: Management Information Systems, Business information technology

Intelligent Applications: Medical and Diagnostic System

PERFORMANCE OF COMPUTER-AIDED DIAGNOSIS TECHNIQUES IN INTERPRETATION OF BREAST LESION DATA

Anatoli Nachev, Mairead Hogan, Borislav Stoyanov

Abstract*: This study explores and compares predictive abilities of six types of neural networks used as tools for computer-aided breast cancer diagnosis, namely, multilayer perceptron, cascade-correlation neural network, and four ART-based neural networks. Our experimental dataset consists of 803 patterns of 39 BI-RADS, mammographic, sonographic, and other descriptors. Using such a combination of features is not traditional in the field and we find it is better than traditional ones. The study also focuses on exploring how various feature selection techniques influence predictive abilities of the models. We found that certain feature subsets show themselves as top candidates for all the models, but each model performs differently with them. We estimated models performance by ROC analysis and metrics, such as max accuracy, area under the ROC curve, area under the convex hull, partial area under the ROC curve with sensitivity above 90%, and specificity at 98% sensitivity. We paid particular attention to the metrics with higher specificity as it reduces false positive predictions, which would allow decreasing unnecessary benign breast biopsies while minimizing the number of delayed breast cancer diagnoses. In order to validate our experiments we used 5-fold cross validation. In conclusion, out results show that among the neural networks considered here, best overall performer is the Default ARTMAP neural network.*

Keywords*: data mining, neural networks, heterogeneous data; breast cancer diagnosis, computer aided diagnosis.*

ACM Classification Keywords*: I.5.1- Computing Methodologies - Pattern Recognition – Models* - Neural Nets

Introduction

Breast cancer is one of the leading causes of death for women in many countries. Mammography is currently the most widely used screening method for early detection of the disease, but it has a low negative predictive value. Many investigators have found that more than 60% of masses referred for breast biopsy on the basis of mammographic findings are actually benign [Jemal et al., 2005], [Lacey et al., 2002]. One goal of the application of computer-aided diagnosis (CAD) to mammography is to reduce the false-positive rate. Avoiding benign biopsies spares women unnecessary discomfort, anxiety, and expense. The problem is nontrivial and difficult to solve. Breast cancer diagnosis is a typical machine learning problem. It has been dealt with using various data mining techniques and tools such as linear discriminant analysis (LDA), logistic regression analysis (LRA), multilayer perceptions (MLP), support vector machines (SVM), etc. [Chen et al., 2009], [Jesneck et al., 2006].

Current CAD implementations tend to use only one information source, usually mammographic data in the form of data descriptors defined by the Breast Imaging Reporting and Data System (BI-RADS) lexicon [BI-RADS, 2003]. Recently, Jesneck et al. [2007] proposed a novel combination of BI-RADS mammographic and sonographic descriptors and some suggested by Stavros et al. [1995], which in combination with MLP show promising results. The MLP have been largely applied in the data mining tasks, but one of their major drawbacks is unclear optimal architecture, which includes number of hidden nodes, activation functions, and training algorithm to learn to predict. Another major problem is to specify optimal set of descriptors used for data mining, which effectively reduces the training and testing datasets to a dimensionality which provides best performance for the application domain. Our study was motivated by addressing those problems and particularly focusing on how reduction of dimensionality of that new combination of descriptors affects performance of not only MLPs, but also other neural network models, such as cascade-correlation nets and those based on the adaptive resonance theory (ART), introduced by Grossberg [1976].

The paper is organized as follows: Section 2 provides a brief overview of the neural networks used in this study: multilayer perceptron (MLP), cascade-correlation neural networks, fuzzy ARTMAP, distributed ARTMAP, default ARTMAP, and ic ARTMAP; Section 3 introduces the dataset and its preprocessing; Section 4 presents and discuses results from experiments; and Section 5 gives the conclusions.

Neural Networks for Data Mining

A variety of neural network models are used by practitioners and researchers for clustering and classification, ranging from very general architectures applicable to most of the learning problems, to highly specialized networks that address specific problems. Each model has a specific topology that determines the layout of the neurons (nodes) and a specific algorithm to train the network or to recall stored information.

Multilayer Perceptions (MLP)

Among the neural network models, the most common is the multilayer perceptron, which has a feed-forward topology and error-backpropagation learning algorithm [Rumelhart & McClelland, 1986]. Typically, an MLP consists of a set of input nodes that constitute the input layer, an output layer, and one or more layers sandwiched between them, called hidden layers. Nodes between subsequent layers are fully connected by weighted connections so that each signal travelling along a link is multiplied by its weight w_{ij}. Hidden and output nodes receive an extra bias signal with value 1 and weight θ. The input layer, being the first layer, has input nodes that distribute the inputs to nodes in the first hidden layer. Each hidden and output node computes its activation level by

$$s_i = \sum_j w_{ij} x_j + \theta \quad , \tag{1}$$

and then transform it to output by an activation function. The MLP we use in this study has one hidden layer with two hidden nodes and log-sigmoid activation function

$$O_i(s_i) = \frac{1}{1 + e^{-\beta s_i}} \quad , \tag{2}$$

We trained the MLP by adaptive learning rate algorithm developed by Jacob [1988], also called delta-bar-delta, or TurboProp. The adaptive learning rate method proposes more flexibility and a higher speed of convergence, compared to the classic backpropagation algorithm.

Cascade-Correlation Neural Networks (CCNN)

CCNN [Fahlman & Libiere, 1990] are supervised self-organizing networks with structure similar to backpropagation networks. Instead of adjusting the weights in a network of fixed topology, a CCNN begins with a minimal number of nodes, then automatically trains and adds new hidden nodes one by one and do not change them over the time. It creates a multi-layer structure called a 'cascade' because the output from all input and hidden nodes already in the network feed into new nodes.

A CCNN has three layers: input, hidden and output. Initially, the network begins with only input and output nodes. The output layer consists of a single neuron if the network is used for regression problems, or contains several neurons for classification problems, one per class label. The hidden layer is empty in the beginning – every input is connected to every output neuron by a connection with an adjustable weight. Such a simple cascade-correlation network has considerable predictive power and for a number of applications it provides excellent predictions. If not, however, the network adds new hidden nodes one by one as illustrated in Figure 1, until the residual error gets acceptably small or the user interrupts this process.

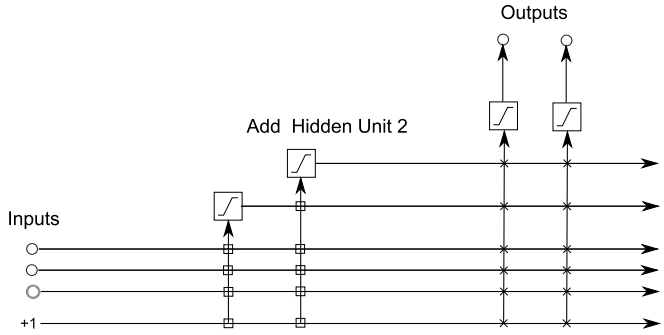


Fig.1. Cascade architecture after adding two hidden nodes (adapted from [Fahlman & Lebier, 1991]). The vertical lines sum all incoming activation. Boxed connections are frozen, 'x' connections are trained repeatedly

The cascade-correlation architecture has several advantages over the traditional backpropagation neural nets. First, as the network is self-organizing and determines its own size and topology during the growth of the hidden layer during training, there is no need to decide how many layers and neurons to use in the network. This is a major problem of the backpropagation networks which implies use of predefined network architecture and designing a near optimal network architecture is a search problem that still remains open. Secondly, cascade-correlation nets learn very quickly (often 100 times as fast as a backpropagation network) and retain the structures they have built even if the training set changes. Finally, they have less chance to get trapped in local minima compared to the backpropagation nets.

Fuzzy, Distributed, Default, and IC ARTMAP Neural Networks

The adaptive resonance theory (ART) introduced by Grossberg [1975] led to the creation of a family of self-organizing neural networks, such as the unsupervised ART1, ART2, ART2-A, ART3, fuzzy ART, distributed ART and the supervised ARTMAP, instance counting ARTMAP, fuzzy ARTMAP (FAM), distributed ARTMAP, and default ARTMAP. ARTMAP is a family of neural network that consists of two unsupervised ART modules, *ARTa* and *ARTb*, and an *inter-ART* module called map-field as shown in Figure 2. An ART module has three layers of nodes: input layer *F0*, comparison layer *F1*, and recognition layer *F2*. A set of real-valued weights W_j is associated with the *F1-to-F2* layer connections between nodes. Each *F2* node represents a recognition category that learns a binary prototype vector w_j. The *F2* layer is connected through weighted associative links to a map field F^{ab}.

The ARTMAP learning can be described by the following algorithm [Carpenter et al., 1991]:

1. *Initialization*: All *F2* nodes are uncommitted, and all weight values and network parameters are initialized.

2. *Input pattern coding*: When a training pattern is presented to the network, a process called complement coding takes place. It transforms the pattern into a form suited to the network. A network parameter called vigilance parameter (ρ) is set to its initial value. This parameter controls the network 'vigilance', that is, the level of details used by the system when it compares the input pattern with the memorized categories.

3. *Prototype selection*. The input pattern activates layer *F1* and propagates to layer *F2*, which produces a binary pattern of activity such that only the *F2* node with the greatest activation value remains active, that is, 'winner-takes-all'. If such a node does not exist, an uncommitted *F2* node becomes active and undergoes learning.

4. *Class prediction*. The class label *t* activates the F^{ab} layer in which the most active node yields the class prediction. If that node constitutes an incorrect class prediction, then another search among *F2* nodes in Step 3 takes place. This search continues until an uncommitted *F2* node becomes active (and learning directly ensues in Step 5), or a node that has previously learned the correct class prediction becomes active.

5. *Learning*. The neural network gradually updates its adaptive weights towards the presented training patterns until a convergence occur. The learning dynamic can be described by a system of ordinary differential equations.

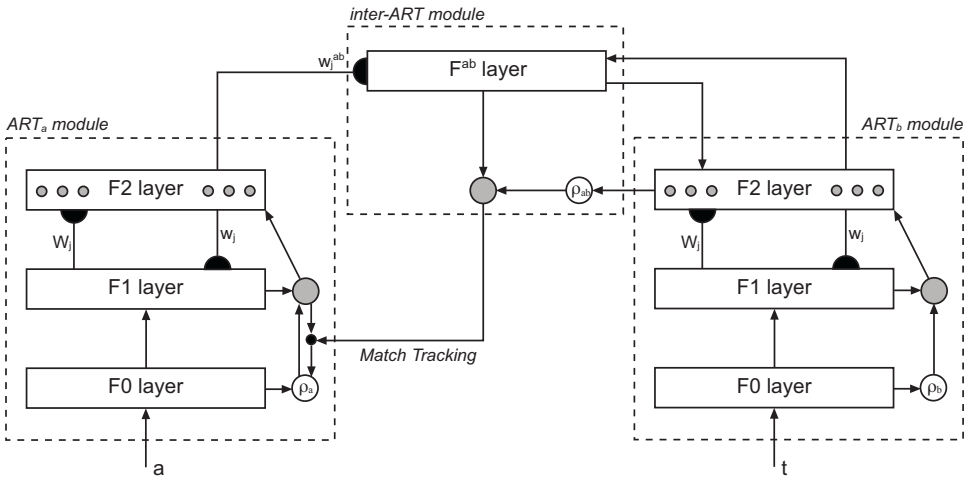


Fig. 2. Block diagram of an ARTMAP neural network adapted from [Carpenter et al, 1991]

Fuzzy ARTMAP was developed as a natural extension to the ARTMAP architecture. This is accomplished by using fuzzy ART modules instead of ART1, which in fact replaces the crisp (binary) logic embedded in the ART1 module with a fuzzy one. In fact, the intersection operator ($\cap$) that describes the ART1 dynamics is replaced by the fuzzy AND operator ($\wedge$) from the fuzzy set theory ($(p \wedge q)_I \equiv \min(p_i, q_i)$) [Carpenter et al., 1992]. This allows the fuzzy ARTMAP to learn stable categories in response to either analog or binary patterns in contrast with the basic ARTMAP, which operates with binary patterns only.

ART1 modules in an ARTMAP net map the categories into *F2* nodes according to the winner-takes-all rule, as discussed above, but this way of functioning can cause category proliferation in a noisy input environment. An explanation of that is that the system adds more and more *F2* category nodes to meet the demands of predictive accuracy, believing that the noisy patterns are samples of new categories. To address this drawback, a new distributed ART module was introduced. If the ART1 module of the basic ARTMAP is replaced by a distributed ART module, the resulting network is called **Distributed ARTMAP** [Carpenter, 1997].

Instance Counting (IC) ARTMAP adds to the basic fuzzy ARTMAP system new capabilities designed to solve computational problems that frequently arise in prediction. One such problem is inconsistent cases, where identical input vectors correspond to cases with different outcomes. A small modification of the fuzzy ARTMAP

match-tracking search algorithm allows the IC ARTMAP to encode inconsistent cases and make distributed probability estimates during testing even when training employs fast learning [Carpenter & Markuzon, 1998].

A comparative analysis of the ARTMAP modifications, including Fuzzy ARTMAP, IC ARTMAP, and Distributed ARTMAP, has led to the identification of the **Default ARTMAP** network, which combines the winner-takes-all category node activation during training, distributed activation during testing, and a set of default network parameter values that define a ready-to-use, general-purpose neural network for supervised learning and recognition [Carpenter, 2003]. The Default ARTMAP features simplicity of design and robust performance in many application domains.

Data and Preprocessing

Our tests used a dataset that contains data from physical examination of patients, including mammographic and sonographic examinations, family history of breast cancer, and personal history of breast malignancy, all collected from 2000 to 2005 at Duke University Medical Centre [Jesneck et al., 2007]. Samples included in the dataset are those selected for biopsy only if the lesions corresponded to solid masses on sonograms and if both mammographic and sonographic images taken before the biopsy were available for review. Data contain 803 samples, 296 of which are malignant and 507 benign. Out of 39 descriptors, 13 are mammographic BI-RADS, 13 sonographic BI-RADS, 6 sonographic suggested by Stavros et al. [1995], 4 sonographic mass descriptors, and 3 patient history features [BI-RADS, 2003], [Jesneck et al., 2007], [Nachev & Stoyanov, 2010]. There are also class label that indicates if a sample is malignant or benign.

We preprocessed the dataset in order to addresses the problem of large amplitude of variable values caused by their different nature and different units of measurements. Consistency we achieved by mapping all data values into the unit hypercube (i.e. all values between 0 and 1), using a linear transformation

$$x_i^{new} = \frac{x_i^{old} - \min_i}{\max_i - \min_i} \tag{3}$$

applied to each variable (data column) separately. This scaling down of values is essential requirement for certain types of neural networks, and particularly for the ARTMAP models we used in our study.

Another preprocessing step was feature selection. In many cases and application domains removing redundant features from the data can help to alleviate effect of curse of dimensionality, avoid overfitting, and speed up learning process. Exhaustive search approach among all possible subsets of features is not applicable in our case as the dataset has cardinality 39. Alternative approaches could be using subset selection algorithms or feature ranking techniques. The former one is preferable as it usually provides good results. We tested genetic search, best first search, subset size forward selection, race search, and scatter search. Our tests showed that the subset size forward selection, proposed by Guetlien et al. [2009] gives good results with all types neural networks we experimented with. This method output a set of 17 descriptors (s17): patient age, indication for sonography, mass margin, calcification number of particles, architectural distortion, anteroposterior diameter, mass shape, mass orientation, lesion boundary, special cases, mass shape, mass margin, thin echo pseudocapsule, mass echogenicity, edge shadow, cystic component, and mass margin. Two of these are general descriptors; three - mammographic BI-RADS; five - sonographic BI-RADS; four - Stavros'; and three - sonographic mass descriptors. The feature set is relatively balanced in representing different categories of data. In our experiments we also used a set of 14 descriptors (s14) proposed by Jesneck et al. [2007] and obtained by stepwise feature selection. We also used the original full set of 39 descriptors (s39).

Empirical Results and Discussion

We used simulators of the neural network models explored here. Series of test showed that best architecture of the multilayer perceptron is one hidden layer with two nodes. We also used a cascade-correlation neural network with Turboprop2 learning based on the Fahlman's work [Fahlman & Libiere, 1990]. Each of the four types ARTMAP neural networks were tested with 41 vigilance parameter values from 0 to 1 and step of increment 0.025. In order to avoid bias in training due to the specific order of training samples, we applied 5-fold cross validation and summarized results in four categories: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Prediction accuracy was calculated by Acc=(TP+TN)/(TP+TN+FP+FN). For the purposes of ROC analysis, we also calculated true positive rate TPR=TP/(TP+FN), and false positive rate FPR=FP/(TP+FN).

No doubts, accuracy is the most common performance estimator of a model, which is used in a vast amount of studies and applications, but in many cases and problem domains it is not sufficient, even can be misleading where important classes are underrepresented in datasets (i.e. class distribution is skewed), or if errors of type I and type II can produce different consequences and have different cost. Secondly, the accuracy depends on the classifier's operating threshold, such as threshold values of MLP or vigilance parameter of ARTMAP NN, and choosing optimal threshold can be challenging. The deficiencies of accuracy can be addressed by the Receiver Operating Characteristics (ROC) analysis [Fawcett, 2006], which plots curves between two indices: TPR and FPR.

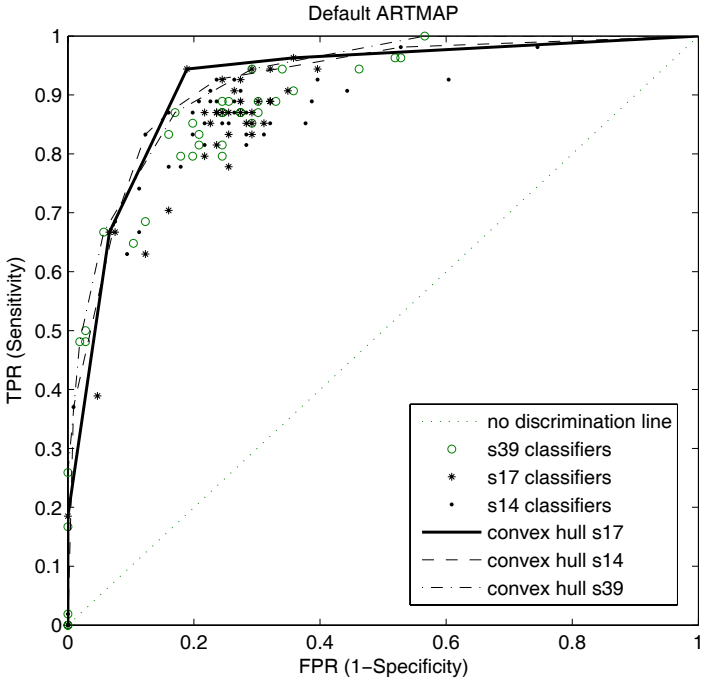


Fig. 3. ROC analysis of Default ARTMAP neural network tested with three sets of descriptors: all available (s39), a selection proposed by Jesneck et al. [2007] (s14), and a set of 17 descriptors (s17) proposed by authors. The model was tested with 41 vigilance parameter values from 0 to 1 with step of increment 0.025

ROC curves are step functions which can be used to select the optimal decision threshold by maximizing any pre-selected measure of efficacy. In general, the best possible prediction method would yield a point in the upper left corner or coordinates (0, 1) of the ROC space, representing 100% sensitivity (all true positives are found) and 100% specificity (no false positives are found). A completely random guess would give a point along a diagonal

line, also known as line of no-discrimination, which is from the left bottom to the top right corners. The optimal classifier would be represented by the most 'northwest' point on the curve, which is the most distant one from the no-discrimination line. Plotting discrete classifiers, such as ARTMAP, require additional processing. Since they do not use operational threshold to be varied in order to generate the ROC curve, the vigilance parameter variance plays the same role. Each parameter value plots one point on the ROC space, and the curve that connects the most northwest points along with the two trivial classifiers (0,0) and (1,1), called also ROC Convex Hull (ROCCH) represents the model as a whole. The ROCCH lines that link points of 'real' classifiers define a continuum of possible classifiers that can be obtained by linear combinations of the plotted ones. ROC analysis also provides additional metrics for estimation of model performance, such as Area Under the ROC curve (AUC) and partial Area Under the ROC curve (pAUC) where sensitivity is above certain value (p). The bigger the AUC / pAUC, the better the model is.

Table 1. Performance of MLP, CCNN, Fuzzy ARTMAP, Distributed ARTMAP, Default ARTMAP, and IC ARTMAP. Metrics for comparison include: area under the ROC curve (AUC), partial AUC at sensitivity above 90% ($_{0.90}$AUC), specificity at 98% sensitivity, and maximal accuracy (ACC_{max}). Models have been tested with three variable selections: s39, s17, and s14. Typical radiologist assessment values are also included.

MLP	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.89	0.91	0.86	0.92
$_{0.90}$*AUC*	0.62	0.68	0.55	0.52
Spec /98% sens	0.37	0.49	0.27	0.52
ACC_{max}	0.89	0.91	0.87	n/a

CCNN	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.896	0.911	0.907	0.92
$_{0.90}$*AUC*	0.648	0.68	0.731	0.52
Spec /98% sens	0.427	0.5	0.49	0.52
ACC_{max}	0.828	0.848	0.838	n/a

Fuzzy ARTMAP	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.851	0.815	0.838	0.92
$_{0.90}$*AUC*	0.586	0.393	0.413	0.52
Spec /98% sens	0.099	0.082	0.091	0.52
ACC_{max}	0.838	0.813	0.819	n/a

Distributed ARTMAP	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.786	0.819	0.744	0.92
$_{0.90}$*AUC*	0.226	0.272	0.187	0.52
Spec /98% sens	0.047	0.057	0.039	0.52
ACC_{max}	0.831	0.856	0.819	n/a

Default ARTMAP	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.927	0.931	0.918	0.92
$_{0.90}$*AUC*	0.725	0.809	0.778	0.52
Spec /98% sens	0.686	0.649	0.690	0.52
ACC_{max}	0.85	0.856	0.863	n/a

IC ARTMAP	*s39*	*s17*	*s14*	*Radiologist*
AUC	0.776	0.821	0.860	0.92
$_{0.90}$*AUC*	0.215	0.282	0.384	0.52
Spec /98% sens	0.045	0.059	0.081	0.52
ACC_{max}	0.831	0.850	0.856	n/a

As long as AUC provides an overall estimation of the model, the partial area is more relevant to the domain of computer-aided diagnosis, and particularly where p=0.9. Another clinically relevant metric used in the application domain is sensitivity at a very high level of specificity (98%).

Table 1 summarizes results from numerous experiments where networks were trained and tested with three different sets of descriptors: s39 that contains the original 39 variables; a selection of 14 variables (s14) proposed

by Jesneck et al. [2007] as a result from stepwise feature selection technique, and our set of 17 variables (s17) we obtained by using the subset size forward selection of Guetlien et al. [2009].

Figure 4 illustrates the results. We obtained highest prediction accuracy of 91.1% by using CCNN with s17. The Default ARTMAP outperforms all other models in terms of overall performance measured by AUC. Here again, s17 is best performer. The figure also illustrates, that Default ARTMAP is the only model (among the studied here) that outperforms the average radiologist performance [Jesneck et al. 2007] in terms of AUC, but more important from a clinical viewpoint are the metrics pAUC and sensitivity at very high specificity. Figures show that again Default ARTMAP beats the others with s17 in terms of pAUC, and again is best performer in terms of sensitivity at high specificity, no matter which subset of descriptors is used.

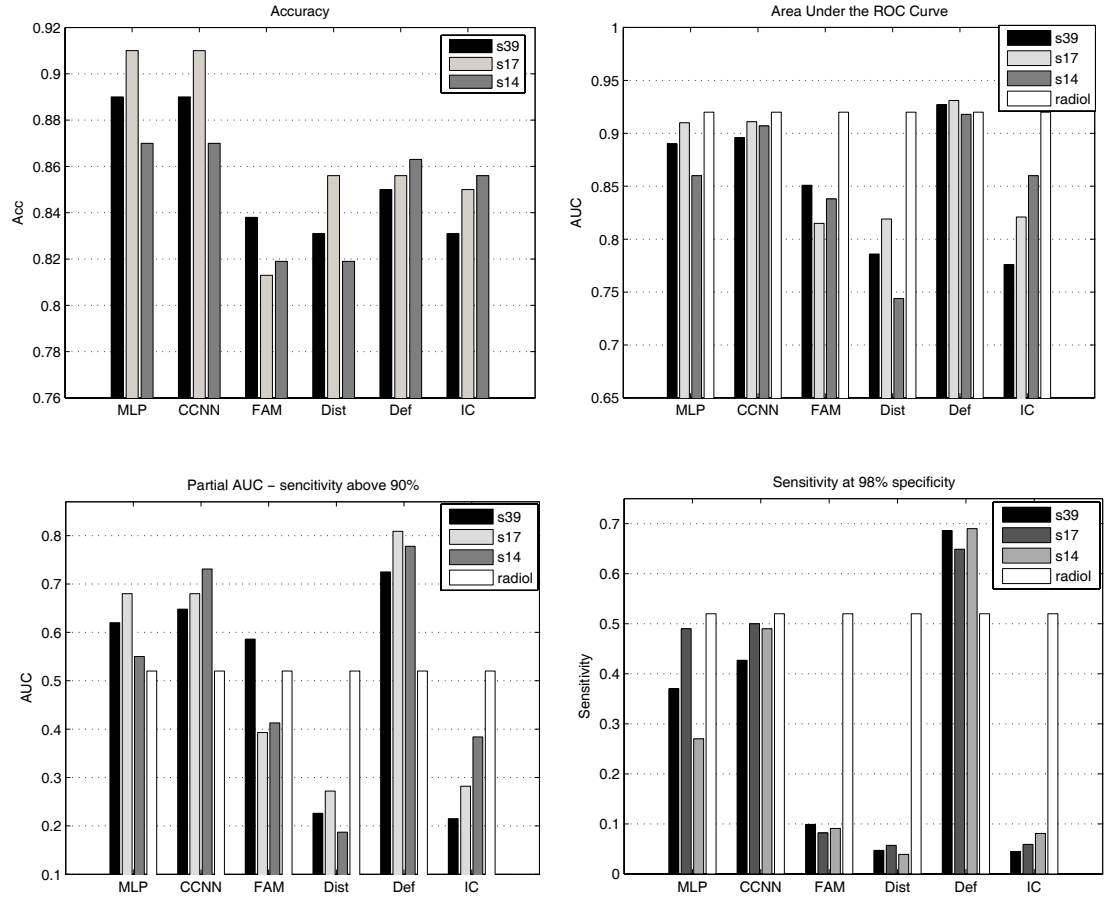


Fig. 4. Performance of three-layer perceptron (MLP), cascade-correlation NN (CCNN), Fuzzy ARTMAP (FAM), Distributer ARTMAP (Dist), Default ARTMAP (Def), and IC ATRMAP (IC), measured by max accuracy (Acc), area under the ROC curve (AUC), partial area under the ROC curve where sensitivity is above 90%, and sensitivity at 98% specificity. Models were tested with all available descriptors (s39), a selection proposed by Jesneck et al. [2007] (s14), and a new proposed set of 17 descriptors (s17). Results were also compared with a typical radiologist performance (radiol)

Figure 3 gives further details on the Default ARTMAP neural network performance obtained by ROC analysis. Bold line represents the best descriptor set. Finally, we find that Default ARTMAP is appropriate for solving the task as its clinically relevant characteristics are good, however a limitation of the model is that is that it requires a very careful tuning.

Conclusion

Many CAD systems for breast cancer screening improve lesion detection sensitivity, but improving specificity is still challenging. This study explores and compares predictive abilities of six types of neural networks: MLP, CCNN, Fuzzy, Distributed, Default, and IC ARTMAP by using a recently proposed combination of BI-RADS mammographic, sonographic, and other descriptors. We also focused our study on how various feature selection techniques influence predictive abilities of those models and found that a subset obtained by subset size forward selection provides best overall results. Our performance estimations were based on ROC analysis and metrics, such as max accuracy, area under the ROC curve and convex hull. We paid particular attention on clinically relevant metrics, such as partial area under the ROC curve with sensitivity above 90%, and specificity at 98% sensitivity, as a higher specificity reduces false positive predictions, which would allow decreasing unnecessary benign breast biopsies while minimizing the number of delayed breast cancer diagnoses. In conclusion, our results show that among all neural networks we explored for this application domain, highest prediction accuracy of 91.1% can be obtained by cascade-correlation neural network, but Default ARTMAP outperforms all other models in terms of overall performance and clinically relevant metrics. All the results confirm that the set of descriptors we propose outperforms the ones used in previous studies.

Bibliography

[BI-RADS, 2003] American College of Radiology. BI-RADS: ultrasound, 1st ed. In: Breast imaging reporting and data system: BI-RADS atlas, 4th ed. Reston, VA: American College of Radiology, 2003

[Carpenter et al, 1991] Carpenter, G., Grossberg, S., & Reynolds, J.: ARTMAP: Supervised Real-Time Learning and Classification of Non-stationary Data by a Self-Organizing Neural Network. Neural Networks, vol. 6, pp. 565-588 (1991)

[Carpenter et al., 1992] Carpenter, G.A., Grossberg, S., Markuzon, N., Reynorlds, J.H. and Rosen, D.B. "Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps", IEEE Transaction on Neural Networks, 3(5), pp. 698-713, 1992.

[Carpenter, 1997] Carpenter, G.. Distributed Learning, Recognition, and Prediction by ART and ARTMAP Neural Networks. Neural Networks, vol. 10:8, 1473-1494, (1997)

[Carpenter & Markuzon, 1998] Carpenter, G. & Markuzon, N.. ARTMAP-IC and Medical Diagnosis: Instance Counting and Inconsistent Cases. Neural Networks, vol. 11:2, pp. 323-336 (1998)

[Carpenter, 2003] Carpenter, G.: Default ARTMAP. Proceedings of the International Joint Conference on Neural Networks (IJCNN'03), Portland, Oregon, pp. 1396−1401, (2003)

[Grossberg, 1976] S. Grossberg, "Adaptive pattern classification and universal recoding. II: Feedback, expectation, olfaction, and illusions." Biological Cybernetics, 23, 1976.

[Fahlman & Lebiere, 1990] Fahlman, S. and Lebiere C. "The Cascade-Correlation Learning Architecture" in D. S. Touretzky (ed.), Advances in Neural Information Processing Systems 2, Morgan Kaufmann, 1990.

[Fawcett, 2006] Fawcett, T. "An introduction to ROC analysis"; Pattern Recognition Letters, Vol. 27 Issue 8, pp. 861-874, 2006.

[Guetlein et al., 2009] Guetlein, M., Frank, E., Hall, M., Karwath, A. "Large Scale Attribute Selection Using Wrappers"; In Proc. IEEE Symposium on CIDM, pp.332-339, 2009.

[Jacob, 1988] Jacob, R. "Increased Rates of Convergence Through Learning Rate Adaptation"; Neural Networks, Vol. 1 Issue 4, pp. 295-307, 1988.

[Jemal et al., 2005] Jemal, A., Murray, T, Ward, E., Samuels, A., Tiwari, R., Ghafoor, A., Feuer, E., Thun, M.. "Cancer statistics", Ca-Cancer J. Clin 2005;55:10–30, 2005.

[Jesneck et al., 2006] Jesneck, J., Nolte, L., Baker, J., Floyd, C., Lo, J. "Optimized Approach to Decision Fusion of Heterogeneous Data for Breast Cancer Diagnosis."; Medical Physics, Vol. 33, Issue 8, pp.2945-2954, 2006

[Jesneck et al., 2007] Jesneck, J., Lo, J., Baker, J. "Breast Mass Lesions: Computer-Aided Diagnosis Models with Mamographic and Sonographic Descriptors"; Radiology, vol.244, Issue 2, pp 390-398, 2007.

[Lacey et al., 2002] Lacey, J., Devesa, S., Brinton, L. "Recent Trends in Breast Cancer Incidence and Mortality."; Environmental and Molecular Mutagenesis, Vol. 39, pp. 82–88, 2002.

[Nachev & Stoyanov, 2010] Nachev, A. and Stoyanov, B., "An Approach to Computer Aided Diagnosis by Multi-Layer Preceptrons", In Proceedings of International Conference Artificial Intelligence (IC-AI'10), Las Vegas, 2010.

[Rumelhart & McClelland, 1986] Rumelhart, D., McClelland, J. "Parallel Distributed Processing"; Explorations in the Microstructure of Cognition, Cambridge, MA: MIT Press, 1986.

[Stavros et al., 1995] Stavros, A., Thickman, D., Rapp, C., Dennis, M., Parker, S., Sisney, G. "Solid Breast Modules: Use of Sonography to Destinguish between Benign and Malignant Lesions"; Radiology, Vol. 196, pp. 123-134, 1995.

Authors' Information

Anatoli Nachev *– Business Information Systems, Cairnes Business School, National University of Ireland, Galway, Ireland; e-mail: anatoli.nachev@nuigalway.ie*

Major Fields of Scientific Research: data mining, neural networks, support vector machines, adaptive resonance theory.

Mairead Hogan *– Business Information Systems, Cairnes Business School, National University of Ireland, Galway, Ireland; e-mail: mairead.hogan@nuigalway.ie*

Major Fields of Scientific Research: HCI, usability and accessibility in information systems, data mining.

Borislav Stoyanov *– Department of Computer Science, Shumen University, Shumen, Bulgaria; e-mail: bpstoyanov@abv.bg*

Major Fields of Scientific Research: artificial intelligence, cryptography, data mining.

Mechanical Engineering

DESCRIPTION OF SURFACES HAVING STRATIFIED FUNCTIONAL PROPERTIES

Wiesław Graboń

Abstract*: The article presents the characteristic of surfaces having functional properties. It discusses the parameters used to describe the roughness of this type of surfaces. It proposed the method of constructing software which calculates probability parameters.*

Keywords*: two-process surfaces, roughness parameters.*

ACM Classification Keywords*: Algorithm, Measurement.*

Introduction

During friction in the presence of a lubricant too smooth surfaces hold lubricating oil poorly (it can lead to erosion of a coupling), whereas too rough surfaces are worn away intensively. The opposing influence of small and large heights of surface roughness of a cylinder liner on the functional properties of the piston-rings-cylinder assembly caused that the researches began to conduct studies in order to find surfaces combining sliding properties of smooth surfaces with the oil storage capacity which is typical of porous surfaces. Thanks to those works, in the 1980s structures of cylinder liner surface achieved after two processes (plateau honing surfaces) came into being. They should be similar to the geometrical structure of cylinder liner surfaces which is created during running-in period, then the time of running-in process and wear should be smaller. The example of these kinds of surfaces is the surface of a cylinder liner after plateau honing. The basic tasks of this surface are: to ensure leak-tightness as well as to provide the piston-rings-cylinder assembly with optimal greasing of gear. The most difficult functioning conditions among all of tribological systems of an internal combustion engine are precisely in piston-rings-cylinder assembly [Niewiarowski, 1983]. In this system the particularly difficult conditions are found in the area of the top dead centre position of the first piston ring, where the thickness of an oil film between a packing ring and smooth surface of a cylinder comprises 0-3 μm. The coefficient of friction between the packing ring and a cylinder comes to 0.1-0.15. The piston-rings-cylinder assembly should assure mileage up to 500000 km with reference to personal cars and 1500000 km to trucks. Between the smooth surface of a cylinder and the surface of piston rings there are various greasing conditions from the boundary lubrication up to hydrodynamic lubrication in the middle of the piston's distance line [Shin, 1983], [Sudarshan, 1983]. To the dominant types of wearing-out of cylinders in an internal combustion engine the researches include: abrasive wear, corrosive wear, adhesive wear and from time to time fatigue wear.

According to the author [Kozaczewski,1986] of the research articles, the geometrical structure of cylinders surface influences engine properties, mainly in the initial stage of its functioning (the period of running-in). It is considered that the rough surface of cylinders causes little tendency to erosion, whereas smooth surface ensures their little wear in the period of running-in. At first, the researches stated that little wear in the period of running-in

was found in cylinders characterised by great smoothness. As a result, this kind of cylinders was recommended. However, as progression in engine construction was marked (most of all in the load growth), erosions of cylinders' smooth surface happened. The author of publication [Wiemann, 1971] claimed that the bigger height of roughness of cylindrical liner which does not have additional surface treatment is, the greater its erosion resistance. It appeared that considerable increase of roughness height is also unfavourable because it causes acceleration of chromium plated piston rings wear. The researches [Sreenath, 1976] proved that above the optimal roughness height of Ra=0.8 μm parameter, the linear wear of cylinder rises. Duck [Duck, 1974] determined the advisable value range of parameter Rt for spark-ignition engines: 2-5μm, whereas for diesel engine 4.7 μm. The authors of the review work [Day, 1986] came to the similar conclusions. The advisable greater roughness in diesel engine results probably from the greater loads in this kind of engine. The difference in functioning conditions of various types of engine is the cause of discrepancy with regard to honing cross-hatch angle α (fig.1), usually smaller in spark-ignition engines in comparison to diesel ones.

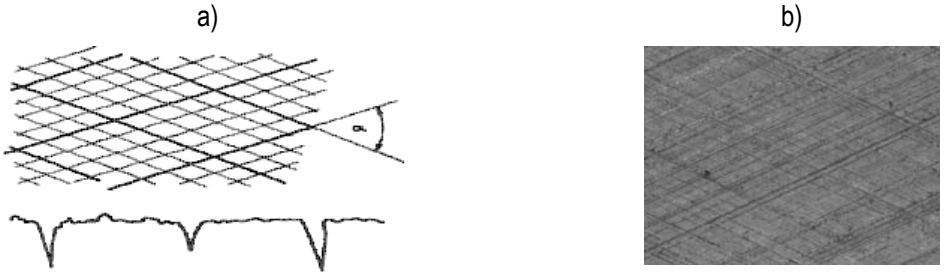


Fig.1 Schematic diagram of a cylinder after honing (a) [Zwierzycki, 1990] and a photograph of the surface of a cylinder liner (b)

The smaller honing cross-hatch angle affects the decrease in consumption of oil which is particularly aimed at in spark ignition engines, but in the case of compression-ignition engines more important issue is to eliminate the galling. Decreasing of the honing cross-hatch angle causes oil film thickness reduction which leads to greater loss of energy and increasing of wear.

The authors of the publications [Pawlus, 1994], [Willis, 1986] confirmed that plateau honing provides lesser linear wear during running-in period and the time of this process might be shortened in comparison to one process honing. In his work [Campbell, 1972], Campbell affirmed that the achievement of linear wear corresponding to 30% of bearing ratio requires two times less of volumetric consumption in the case of plateau honing cylinders in comparison to the similar surface of the same roughness height after one process honing. The usage of plateau honing caused significant reduction of running-in time [Willis, 1986]. Santochi and Vignale [Santochi, 1982] employed plateau honing with reference to air cooled motorcycle engines. They reached geometrical structure od surface characterized by Ra=1μm, Rz=12 μm, Rv/Rp=2 parameters, however traditional structure was characterised by the following parameters: Ra=2.4μm, Rz=18 μm, Rv/Rp=1.1. One obtained faster stabilization as well as improvement of functional parameters of engine after running-in by replacing the traditional honing with the plateau honing (fig. 2). Dolecki et al. in their research [Dolecki, 1983] proved smaller oil consumption (at about 20%) by engine in the case of equipping it with plateau honing cylinders. The research conducted on Polonez engine shows that the plateau honing has a positive influence on the linear wear of cylinders during running-in. It allowed to conclude that the linear wear in the period of running-in is proportional to roughness height as well as to emptiness coefficient Rp/Rt [Pawlus, 1999]. The evidence for the supremacy of plateau honing over one-process honing are also works [Essig, 1990], [Barber, 1987]. One of a few comprehensive works concerning the influence of two process~~es~~ surface on fricative tribological properties is Jeng's work [Jeng,1996]. He highlighted

that the research described in literature was not able to grasp only the influence of microgeometry of surface on the wear, therefore he was not sure about the need of using additional process of honing.

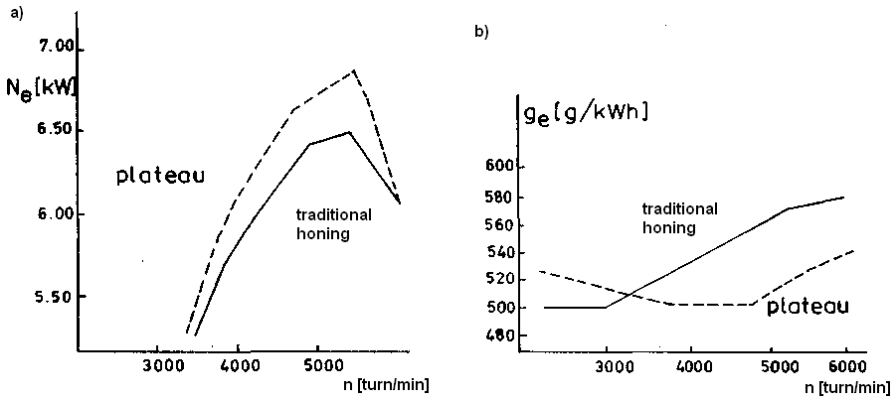


Fig.2 The influence of geometrical structure of surface on power output Ne (a) and individual fuel consumption ge (b) [Santochi, 1982]

He studied the one process honing surfaces and plateau honing surfaces, both kinds of the surfaces had the same values of the Rq parameter. On the basis of experiments carried out on the special simulation stand he found that the two processes surfaces have a shorter running-in period and less galling resistance in relation to the one process surfaces (the resistance may be increased by the use of oil additives).

The initial wear of these surfaces during the running-in period was greater, however they quickly attained a constant intensity of wear. Therefore, during the work of piston-rings-cylinder assembly they ensure less wear in comparison with surfaces after one process honing. He also performed friction coefficient test and found that in the course of the work in a homogenous hydrodynamic lubrication conditions, the coefficient of friction two-process surfaces is the same as the one-process surfaces. However, the two-process structures of surface are more advantageous from the point of view of the influence on the mixed friction coefficient.

Nosal considers [Nosal, 1998] that the increase in resistance of galling of the plateau honing structure is caused by the increase of oil surface capacity (which causes lubricating layer thickness increase, the reduction of friction resistance and temperature in the contact zone), and frequent interruptions of two surfaces being in contact caused by valleys (which reduces the probability of a galling centre).

The authors of the publications [Sudarshan, 1983] paid attention to the possibility of accumulation of abrasive particles in valleys created during honing process. It should lead to reduction of intensity of abrasive wear. The considerable deterioration in lubrication conditions since the disappearance of valleys was observed by the authors of the publications [Stout, 1990]. This kind of situation causes possibility of intensification of the abrasive wear or generation of galling danger. The specific danger appears in the high loaded diesel engines, in this case on the cylinder liner a very smooth texture comes into existence, it resembles bore polishing surface of a significant tendency to galling. This is the case when together with the increase of smoothness friction increases as well. The authors of the publications [Michalski, 1994] studied the influence of roughness of cylinders liner surfaces on the value of abrasive wear which significantly exceeds the initial roughness height. As a result of the conducted studies, they claimed that the abrasive wear of the cylinder liner is proportional to the distance between the honing valley and to the value of the emptiness coefficient Rp/Rt. Excessive increase of height roughness causes oil consumption increase in the engine [Kozaczewski, 1986]. Figure 3 shows an example of the impact of roughness height of the cylinder liner on the oil consumption.

During recent years there have been a few changes in requirements relating to cylinder liner roughness: from very smooth, mirror-surface (Ra = 0.20 mm), by changing the roughness to higher one (Ra in range 0.8-1.2 mm) caused by the increase in power of an engine, up to plateau structure which is now almost universally applicable.

The increased roughness diminishes the ability to galling, however, it causes the increased consumption of oil and increased toxicity of exhaust gases.

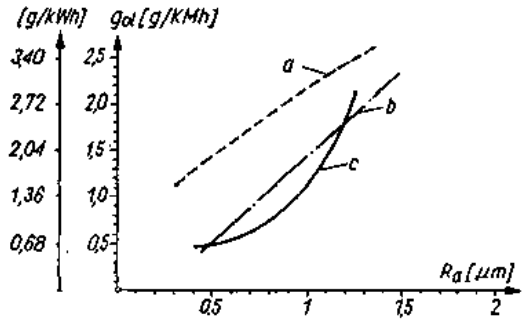


Fig. 3. Influence of cylinder liner roughness on the oil consumption in internal combustion engine [Kozaczewski, 1986]: a) compression-ignition engine, Vs (engine cubic capacity) = 6 dm3, b) spark ignition Vs = 1.3 dm3, c) compression-ignition, V s = 4 dm3

We can also find the opposite statement. For example, Gruszka [Gruszka, 1983] in his doctoral thesis maintains that too low roughness increases the consumption of oil.

During the construction process of new engines manufacturers of trucks must take into consideration very stringent requirements which are constituted by the European Union standards included in the series EU III/IV/V relating particularly to toxicity of exhaust gases. Therefore, they strive to reduce the roughness of cylinder liner, which leads to reduction of oil film thickness and consequently to smaller oil consumption and exhaust gases emissions. Currently, there are the following ways of honing development of cylinder liner surfaces:

- glide or slide honing,
- honing using laser beams to cracks cutting,
- manufacturing of lubricating pockets on the smooth honing surface.

There are two types of glide honing due to the honing cross–hatch angle:

- measuring 60^o angle,
- measuring 140^o angle, so called – spiral [Cieślak, 2008].

In currently manufactured (according to plateau standard) cylinder liners , cylinder bearing surface is described by the following roughness parameters: Rpk < 0.3 μm, Rk = 0.8-1.4 μm, Rvk = 1.7-3.2 μm. In order to decrease the thickness of oil film and reduce oil consumption as well as exhaust gases emissions it has been proposed to reduce the roughness to the following parameter levels: Rpk < 0.2 μm, Rk = 0.2-0.5 μm, Rvk = 1.4 to 3.0 μm. New requirements focus on even greater reduction of parameter Rk value.

The researchers from the Volvo and the University of Halmstad (Sweden) followed the program Piston Simulation for the analysis of impact of the geometrical structure of cylinder liner surface on the oil film thickness and friction force. It was found that the increase of the value of the Rk parameter has an impact on increasing both analysed values in a top dead centre position of the piston [Johansson, 2008].

The authors of the article [Ohlsson, 2003] explored the correlation between roughness parameters of a cylinder liner and consumption of oil by the engine. Oil consumption is proportional to the value of many parameters of geometrical structure of cylinders liner surface, but only Rq, Rvk and Rk parameters in 2D and 3D correlation coefficients are included within the limits of 0.9-1.

The Authors of the articles [Hassis, 1999], [Schmid, 2006], [Schmid, 1999] from Nagel company think that producers should aim at minimizing the value of the parameters Rk and Rpk in order to reduce oil consumption.

They presented the results of the studies in accordance with which the change of the classical plateau honing (Rpk < 0.2 μm, Rvk = 1.4 to 2.0 μm, Rk = 0.6-0.8 μm) to the glide honing (Rpk < 0.1 μm, Rvk = 0.8-1.2 μm, Rk < ¼ Rvk) leads to the reduction in consumption of oil by the engine (over 60%). The reduction of oil consumption is even more affected by the application of spiral honing.

It is essential to modify the honing cross-hatch angle by the application of spiral honing. Application of angle less than 90 degrees counteracted too easy blow-by of oil into the combustion chamber. However, the increase of the honing cross-hatch angle (fig. 1) leads to the decrease of abrasive wear of cylinder liner. The applied honing cross-hatch angle less than 180 degrees allows rotations of the piston rings. The laser honing process is a perspective technology, since it allows to reach the geometrical surface structure with the guarantee of small oil consumption and adequate lubrication surface of cylinder liner at the most loaded place - in piston's top dead centre position. The remaining surface of cylinder liner can have small height of roughness. For example, Klink [Klink, 1997] received the cracks with a width of 40-80 μm, depth 5-25 μm, and the mutual distance of 300 μm. Figure 4 shows examples of cylinder liner surfaces made by different methods.

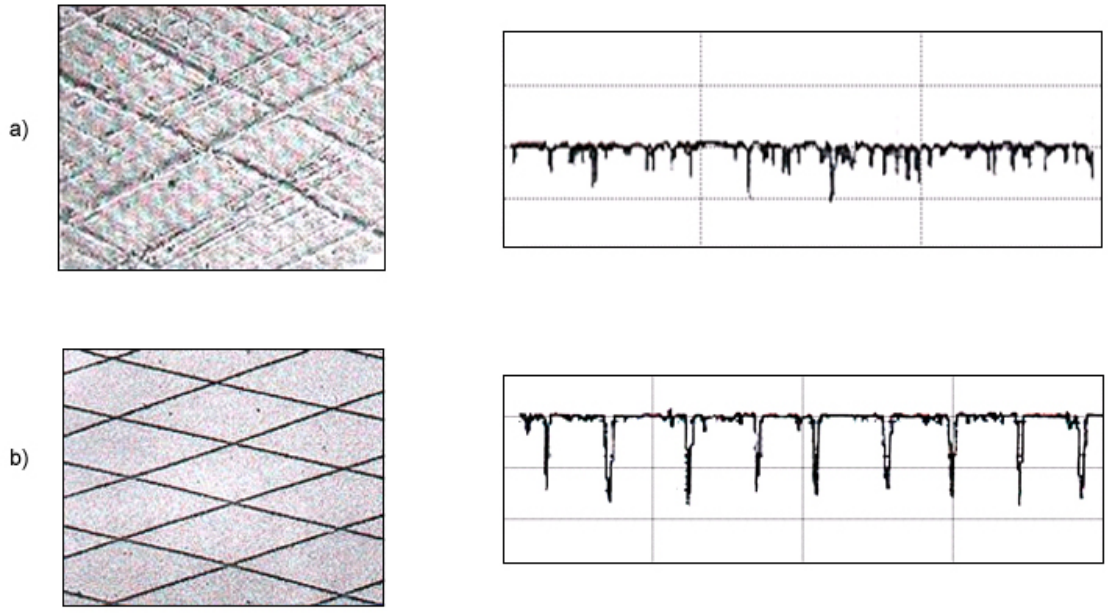


Fig.4. The surface of cylinder liner after glide honing (a) and laser honing (b) [Cieślak, 2008]

As a result of literature analysis it can be concluded that the topography of the cylinder liners' surfaces formed by honing process has a significant impact on the operating parameters of the combustion engines, particularly in the early period of their work.

Description of two-process surfaces

In accordance with the information contained in the works of [Chusu, 1975], [Nowicki, 1991] the profiles of cylinder liner surfaces after plateau honing are irregular, they usually have random character. When on this kind of profiles periodical irregularity appears then we can qualify them to the mixed profiles. Periodic irregularities depending on oscillation and kinematics of machining process may be a consequence of the earlier boring process of cylinders. Due to the explicit directionality of cylinders' patterns the authors [Wieczorowski, 1996] rated the cylinders' surfaces after honing to the mixed surfaces. According to the author of the work [Michalski, 1998] waviness of surface of honing cylinders has the random characteristics.

In connection with an important impact of the geometrical structure of surface of cylinder liner on the combustion engines exploitation properties, their producers have high requirements concerning the roughness parameters of plateau honing surfaces.

These requirements are mainly related to the proper determination of material ratio curve (so called bearing curve or Abbot Firestone curve) of roughness profile, the profile height and the distance between the deep valleys. The requirements of Berliett company from the 1970s (see fig. 5.) are typical example.

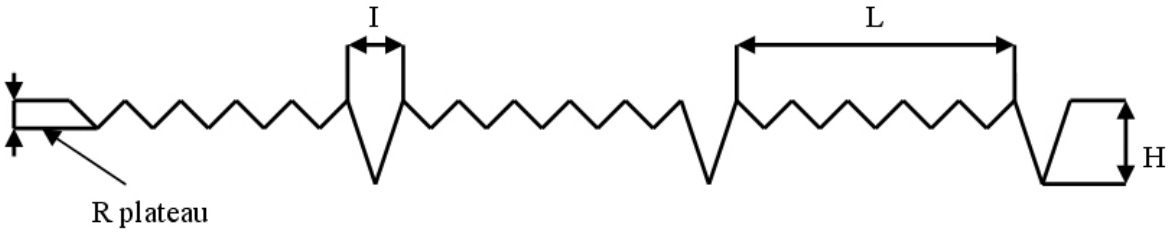


Fig 5. Roughness parameters describing the cylinder liner surface profile [Cieślak, 2008]

Parameters should take the following values:

- average roughness depth: 2.5 μm ≤ R ≤ 6.5 μm
- "plateau" roughness depth: 1 μm $_{plateau}$ ≤ R ≤ 3 μm
- depth of valleys: 6 μm ≤ H ≤ 16 μm
- width of the plateau: 125 μm ≤ L ≤ 600 μm
- width of valleys: 10 μm ≤ l ≤ 65 μm

Important parameters are the horizontal ones, it is demonstrated by the fact that the average distance between valleys and their dimensions are often included in the requirements of engine manufacturers.

The method of determining the width of the valleys previously required by the GOETZE company was subjective. More objective manner, similar to the method described in the article [Michalski, 1994], is shown in the work [Lenhof, 1997]. It is based on the number of intersections of profile with the line described in DIN 4776 standard. Other parameters, used by researchers because of their statistical significance are the statistical moments of the third and fourth order: Rsk (skewness or asymmetry) and Rku (eksces or kurtosis). Willn [Willn, 1972] said, however, that these parameters are correlated with each other, therefore, he proposed additional parameters based on analysis of distribution of the number of peaks or distribution of cross sections of the profile lines parallel to the geometric mean line.

Many researchers associated with automotive companies use bearing area curve to cylinder liner surface analysis. In this curve we can distinguish 3 basic parts: peak, central and valley area, responsible for the different properties of surface. Abbott and Firestone [Abbott, 1993] considered that part of the peak corresponds to 2-5% of the bearing ratio, the central 25-75%, valley 75-98%.

German researchers proposed the profile roughness description method described in DIN 4776 (and later ISO 13565-2). Nielsen thought that honing process can be controlled by changing of Rk parameter value [Nielsen, 1988]. Authors of the work [King, 1994] tried to determine the value of the five parameters from the group "Rk" on the basis of the value of the parameters Rsk, Rku Rq. It is possible only for certain types of ordinate distributions, which are characterized by small asymmetry.

Criticizing the method defined in DIN 4776, Zipin in his work [Zipin, 1983], discredited its usage for analysis of surface's profiles having Gaussian's ordinates distribution. Also the authors of the work [Malburg, 1993] express doubts as to the correctness of the determination of Mr2 parameter. It depends on the slope of the material

bearing ratio curve in its middle area. The parameters contained in ISO 13565-2 are used by most European manufacturers of internal combustion engines.

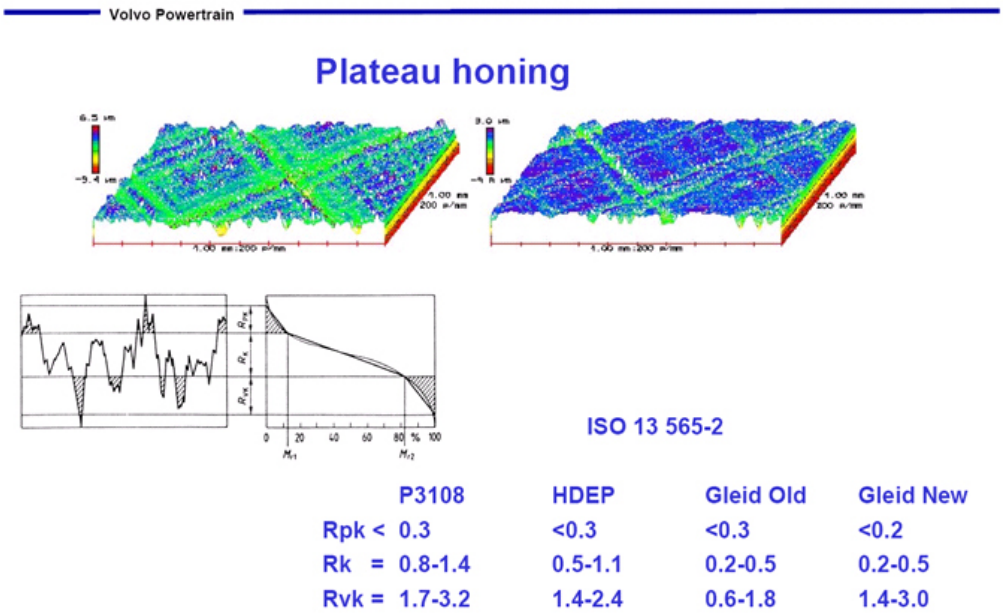


Fig 6. The values of cylinder liner surface parameters used by the Volvo company [Cieślak, 2008]

Figure 6 shows the requirements applied by the VOLVO company for compression ignition engines used in trucks. Authors in the work [Michalski, 1992] suggested a method based on the analysis of approximated material ratio curve. They applied the following equation (1):

$$R = 0.35[1 + (2/\pi)arctg(A\{tg[(\pi/2)(2tp-1)] - tg[(\pi/2)(2X0-1)]\})] + 0.3tg(Btp)/tg(B) \quad (1)$$

where A, X, B-independent parameters, R - standardized height of the roughness , tp - bearing ratio.

This approach allows you to specify the minimum and maximum curvature of the coordinates (xrk1, yrk1, xrk, yrk), coordinates of a point of inflexion bearing curve (xpp, ypp) and tangent of slop of this curve at that point (del).These parameters and the ones resulting from ISO 13565-2 are shown in Figure 7. The parameter yrk1 is analogous to yr1, yrk-yr2, ypp-pp, xrk1-Mr1, xrk- Mr2, del-Rk/Rt, Vok-Vo2.

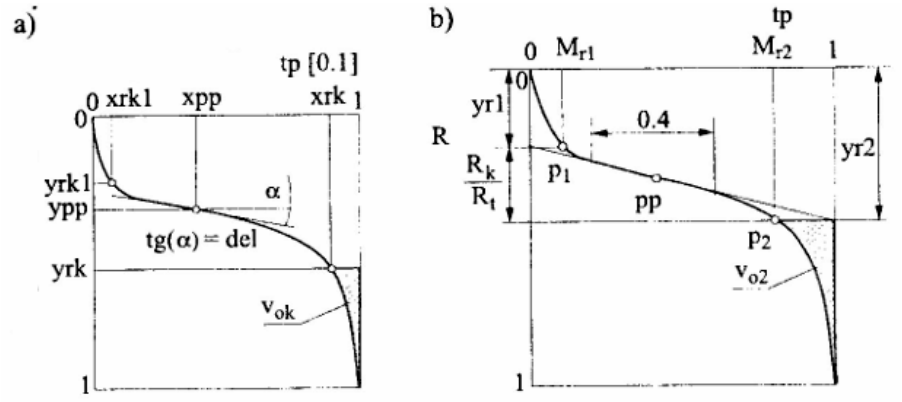


Fig 7. Some parameters of roughness resulting from the approximation of the Abbott curve, (a) and ISO 13565-2 standard (b) [Pawlus, 1999]

The authors of the publications [Malburg, 1993], [Sanna-Reddy, 1997] proposed a method of analysis the surface obtained after many processes. It was used in the American company Cummins producing internal combustion engines. This method is described in ISO 13565-3 standard (fig. 8).

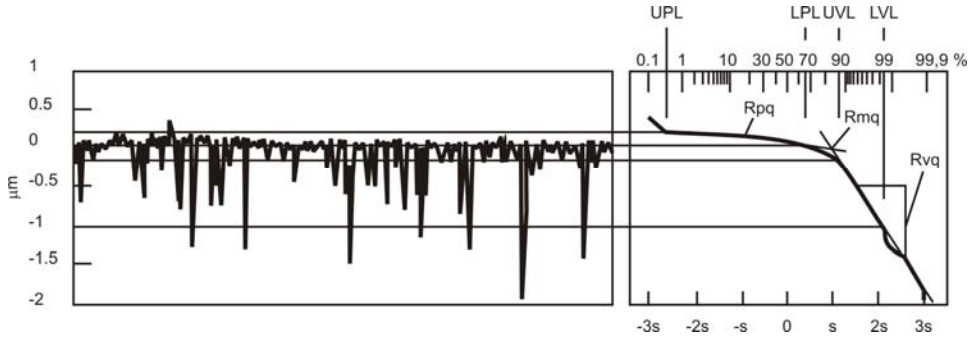


Fig. 8. Graphical interpretation of parameters contained in PN-EN ISO 13565-3: 2002 standard [Jakubiec, 2004]

This method called the probabilistic one, is based on an analysis of the data becomes from the bearing area curve plotted on normal probability paper. In this co-ordinate system, surface which has the Gaussian ordinate distribution is described by a straight line, but the surface after two-process honing, both of them having Gaussian ordinate distribution too, is described by two straight lines of different slopes. The intersection point of those lines in this graph according to the authors of the work [Malburg, 1993] separates the plateau and base texture.

In this graph the abscissa of this intersection point is defined as Rmq parameter (see fig.8) which is important feature of the model because it depends on the honing time. Rpq parameter is the slop of a regression line drawn through the plateau region, but Rvq – through the walleyes region. Rpq, Rvq and Rmq are three parameters independently characterizing each area of plateau honed surfaces, therefore the honing process controlled with their use should be precise.

This method was recommended in the work: [Whitehouse, 1985] [Zipin, 1983], it was also used for the analysis of the zero wear process and to study running-in process. It is conceptually simpler and more elegant than the method described previously (ISO 13565-2), but the practical difficulty of the parameters Rpq, Rvq and Rmq computations exists [Pawlus, 2009].

The method has been used to model the geometrical surface structure after two processes. The authors of the works: [Rosen, 2004], [Anderberg, 2009] analyzed the connections between parameters from ISO 13565-2 (Rk) and ISO 13565-3 (Rq) standards with honing process parameters. However, the conclusive answer to the question which group of parameters is more associated with the manufacturing process was not found. Parameters contained in ISO 13565 standard make a major contribution to an analysis of profiles after several processes. These parameters can be used also in 3D system. The analysis of the relative differences between parameters contained in the standards ISO 13565-2 and ISO 13565-3 and their three-dimensional counterparts is interesting. It can give the answer to the question whether these parameters describe the statistics (average) properties of the surface or they are susceptible to the presence of accidental valleys and peaks. In order to find the intersection point between plateau and valley areas, methodology described in ISO 13565-3 standard and other methods (for example [Sannareddy, 1998]) used different curves to the approximation probability plot of cumulative distribution. Those methods have some imprecisions which were noticed by the authors of the works [Jakubiec, 2004], therefore, the author proposes a different way of solving the problem. This method will be presented in the following subsection.

Implementation

To automatize the determination process of parameters Rq in plateau and valley area, computer program was created. This program was partly based on algorithm described in ISO 13565-3 standard. The main problem is to determine transition point between two random regions. To find this point material probability curve graph was rotated by ψ angle anticlockwise according to the following equation:

$$x' = x\cos\psi - y\sin\psi$$

$$y' = x\sin\psi + y\cos\psi$$

ψ angle is the slope of straight line passing by the first and the finishing point of the material ratio curve (see Figure 9).

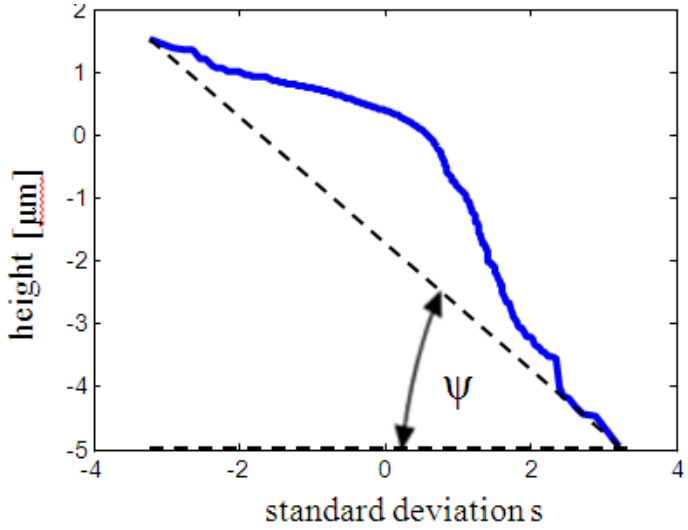


Fig.9. Material probability plot with straight line passing by the first and the finishing point and ψ angle In rotated diagram C point of the highest ordinate was determined (see Figure 10). This point is treated as transition between two random regions.

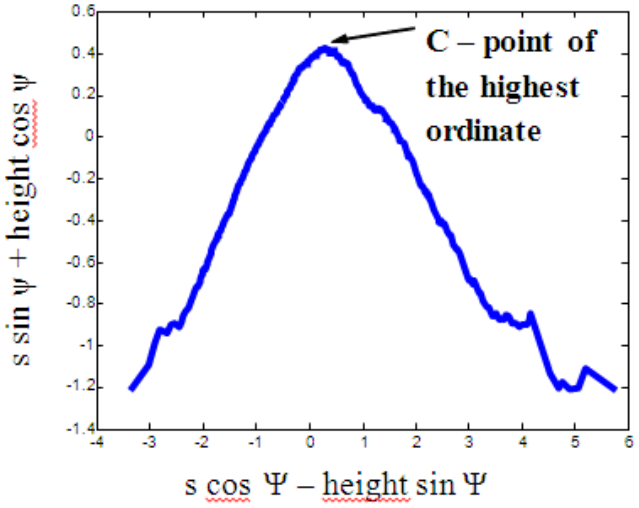


Fig.10. Material probability plot rotated by ψ angle

According to methodology which is recommended in ISO 13565-3 standard, nonlinear material probability curve graph regions were eliminated and Upper Plateau Limit (UPL), and Lower Valley Limit (LVL) points were assigned.

In material probability curve graph the lower boundary of the region plateau (LPL) and upper region valley (UVL) were determined by elimination of a few points which are situated partly right and partly left from transition point. The number of eliminated points was determined from the value of curvature material ratio curve in transition area. Afterwards, linear regression lines between points UPL and LPL and between points UVL and LVL were determined. The values of directional coefficients of these lines were assigned as values Rpq and Rvq. Parameter Rmq was assigned as value of abscissa in intersection of regressions lines drown in plateau and valley region - Fig.11. Areal Spq, Svq or Smq parameters can be also determined in this way.

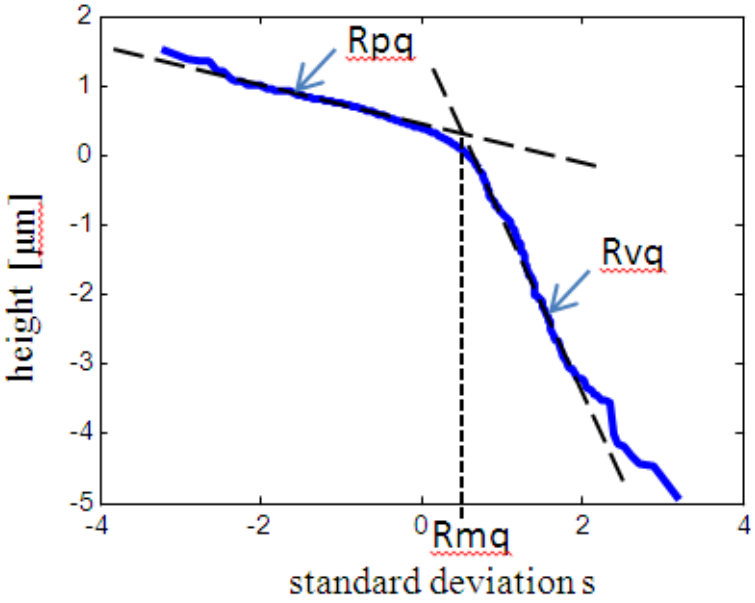


Fig.11. Material probability plot with regression lines passing through random regions and transition point between those regions

Conclusion

Parameters describing roughness profiles of surfaces having stratified functional properties were proposed in the standards ISO 13565-2 (Rk) and ISO 13565-3 (Rq). Parameters from Rq group can be used to profile simulation and zero-wear evaluation or to control of manufacturing process. The number of those parameters is lesser than parameters from Rk group and they are not determined arbitrarily as in ISO 13565-2 standard. However, in European industry the parameters from ISO 13565-2 are used. Methodology of determining the parameters included in ISO 13565-3 standard is imprecise, therefore its incorrect usage can lead to significant mistakes. That is why different methodology to determine those parameters was suggested. After the analysis of many surfaces it was found that this method is useful.

Bibliography

[Abbott, 1993] E.J.Abbott, F.A.Firestone: Specifying surface quality. Mech. Eng., 55, 1993, 556-572.

[Anderberg, 2009] C.Anderberg, P.Pawlus, B.G.Rosen, T.R.Thomas: Alternative descriptions of roughness for cylinder liner production. Journal of Materials Processing Technology 209, 2009, 1936-1942.

[Barber, 1987] G.C.Barber, J.Lee, K.C.Ludema: Materials and surface finish effects in the breaking-in process of engines. ASME Journ. of Eng. for Gas Turbines and Power, 109, 1987, 380-387.

[Campbell, 1972] J.C.Campbell: Cylinder bore surface roughness in internal combustion engines: its appreciation and control. Wear, 19, 1972, 163-168.

[Cieślak, 2008] T.Cieślak: Wpływ wybranych parametrów procesu gładzenia na strukturę geometryczną powierzchni cylindrów. Rozprawa doktorska, Politechnika Rzeszowska, Rzeszów 2008.

[Chusu, 1975] A.P.Chusu, Ju.R.Witenberg, W.A.Palmow: Szerochowatost powierchnostiej. Moskwa, Nauka 1975.

[Day, 1986]R.A.Day, T.J.Reid, D.C. Evans: Developments in liner technology. AE Technical Symp, paper 2, New York, USA, 1986.

[Duck, 1974] G. Duck: Zur Laufflachengestaltung von Zylindern und Zylinderlaufbahnen. MTZ, 35 (10), 1974, 339-340.

[Dolecki, 1983] W.A.Dolecki, W.N.Buntów, A.K.Argusow, B.A.Opoczyński, M.A.Grigorjev: Zwiększenie trwałości maszyn metodami technologicznymi. WNT, Warszawa 1983.

[Essig, 1990] G.Essig, H.Kamp, E.Wacker: Diesel engine emissions reductions - the benefit of low oil consumption design. SAE, paper 900591, 1990.

[Gruszka, 1983] J.Gruszka: Badanie wybranych cech warstwy wierzchniej żeliwnych tulei w oparciu o próby zużyciowe. Rozprawa doktorska, Politechnika Poznańska, Poznań 1983.

[Hassis, 1999] G.Hassis, U-P.Weigmann: New honing technique reduces oil consumption. Industrial Diamond Review, 3/99, 205-211.

[Jakubiec , 2004] W.Jakubiec , J.Malinowski: Metrologia wielkości geometrycznych. PWN, Warszawa 2004.

[Jakubiec , 2004] W. Jakubiec, M. Brylski: Validation of software for calculation the surface roughness parameters according to ISO 13565-3. Proceedings of 9th International Conference on Metrology and Properties of engineering Surfaces, Halmstad University, Sweden 2004, 77-84.

[Jeng,1996] Y.Jeng: Impact of plateaued surfaces on tribological performance. Tribology Transactions, 39/2, 1996, 354-361.

[Johansson, 2008] S.Johansson, P.H.Nilsson, R.Ohlsson, C.Anderberg, B.-G.Rosen: New cylinder liner surfaces for low oil consumption. Tribology International, 41, 2008, 854-860.

[King, 1994] T.G.King, N.E.Houghton: Describing distribution shape: Rk and central moment approaches compared. Proceedings of the 6-th Conference on Metrology and Properties of Engineering Surfaces, Birmingham, UK, 1994, 110-118.

[Klink, 1997] L.Klink: Laserhonen fur Zylinderlaufbagnen. MTZ, 58/9, 1997, 554-556.

[Kozaczewski,1986] W.Kozaczewski: Konstrukcja złożeń: tłok-cylinder silników spalinowych. WKŁ, Warszawa 1986.

[Lenhof, 1997] U.Lenhof, A.Robota: Assessment of honed cylinder bore surfaces for IC engines. AE GOETZE GmbH, Braunschweig 1997.

[Malburg, 1993] M.C.Malburg, J.Raja: Characterization of surface texture generated by plateau-honing process. CIRP Annals, 42/1, 1993, 637-640.

[Michalski, 1992] J.Michalski, P.Pawlus: Description of the bearing length curve of the inner surface of piston engine cylinders. Wear, 157, 1992, 207-214.

[Michalski, 1994] J.Michalski, P.Pawlus: Effects of metallurgical structure and cylinder surface topography on the wear of piston ring-cylinder assemblies under artificially increased dustiness conditions. Wear, 179, 1994, 109-115.

[Michalski, 1998] J.Michalski: Badania poprawności odwzorowania nierówności powierzchni cylindrów po gładzeniu płasko-wierzchołkowym. Materiały konferencji: „Mechanika '98", część 2, Rzeszów 1998, 149-158.

[Michalski, 1994] J.Michalski, P.Pawlus: Description of honed cylinders surface topography. Int. J. Mach. Tools Manufact., 34/2, 1994, 199-210.

[Nielsen, 1988] H.S.Nielsen: New approaches to surface roughness evaluation of special surfaces. Precision Engineering, 10, 1988, 209-213.

[Niewiarowski, 1983] K. Niewiarowski: Tłokowe silniki spalinowe. WKŁ, Warszawa 1983.

[Nosal, 1998] S.Nosal: Tribologiczne aspekty zacierania się węzłów ślizgowych. Rozprawy nr 328, Wydawnictwo Politechniki Poznańskiej, Poznań 1998.

[Nowicki, 1991] B.Nowicki: Struktura geometryczna. Chropowatość i falistość powierzchni. WNT, Warszawa 1991.

[Ohlsson, 2003] R. R.Ohlsson, B.-G.Rosen: Surface texture knowledge – ISM. In: Advanced Techniques for Assessment Surface Topography. L. Blunt, X. Jiang (Eds.), Kogan Page, London and Sterling 2003, 325-336.

[Pawlus, 1994] P.Pawlus: A study on the functional properties of honed cylinder surface during running-in. Wear, 176, 1994, 247-254.

[Pawlus, 2009] P. Pawlus , T. Cieslakb, T. Mathia The study of cylinder liner plateau honing process Journal of Materials Processing Technology 209.

[Pawlus, 1999] P.Pawlus: Struktura geometryczna powierzchni cylindrów podczas eksploatacji silnika spalinowego. Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 1999.

[Rosen, 2004] B.G.Rosen, C.Anderberg, R.Ohlsson, Characterisation of cylinder roughness for manufacturing control and quality assurance. Addendum to Proceedings of XI International Colloquium on Surfaces, Chemnitz (Germany) 2004.

[Sannareddy, 1998] H. Sannareddy, J. Raja, K.Chen: Characterization of surface texture generated by multi-process manufacture. Int. J. Mach. Tools Manufact. Vol 38. Nos 5-6, pp.529-536, 1998.

[Sanna-Reddy, 1997] H.Sanna-Reddy, J.Raja, K.Chen: Characterization of surface texture generated by multi-process manufacture. Transactions of the 7th International Conference on Metrology and Properties of Engineering Surfaces, Gothenburg, Sweden, 1997, 111-118.

[Santochi, 1982] M.Santochi, M.Vignale: A study on the functional properties of a honed surface. CIRP Annals, 31, 1982, 431-434.

[Schmid, 1999] J.Schmid: Honing technology for optimal cast-iron cylinder liners. Engine Technology International, 1(2), 1999, 108-109.

[Shin, 1983] K.Shin, Y.Tateishi, S.Furuhama: Measurement of oil film thickness. SAE, paper 830068, 1983.

[Sudarshan, 1983] T.S.Sudarshan, S.B.Bhaduri: Wear in cylinder liner. Wear, 91, 1983, 269-277.

[Sreenath, 1976] A.V. Sreenath, N.Raman: Running-in wear of a compression ignition engine: factors influencing the conformance between cylinder liner and piston ring. Wear, 38, 1976, 271-289.

[Sudarshan, 1983] T.S.Sudarshan, S.B.Bhaduri: Wear in cylinder liner. Wear, 91, 1983, 269-277.

[Stout, 1990] K.J.Stout, EJ.Davis, P. I.Sullivan: Atlas of machined surfaces. Chapman and Hall, London 1990.

[Schmid, 2006] J.Schmid: Optimized honing process for cast iron running surfaces. VDI Symposium “Piston running surfaces, pistons and conrods”. Boblingen, Germany, Schmid.

[Willis, 1986] E.Willis: Surface finish in relation to cylinder liners. Wear, 109, 1986, 351-366.

[Wieczorowski, 1996] M.Wieczorowski: Stereometryczna ocena porównawcza powierzchni przy zastosowaniu różnych układów odniesienia profilometrów. Rozprawa doktorska Politechnika Poznańska, Poznań 1996.

[Willn, 1972] J.E.Willn: Characterisation of cylinder bore surface finish: a review of profile analysis. Wear, 19, 1972, 143-162.

[Whitehouse, 1985] D.J.Whitehouse: Assessment of surface finish profiles produced by multi-process manufacture. Proceeding of the Inst. Mech. Engrs, 199/4, 1985, 263-270.

[Wiemann, 1971] L.Wiemann: Die Bildung von Brandspuren auf den Laufflachen der Paarung Kolbenring - Zylinder in Verbrennungsmotoren. MTZ, 32/2, 1971, 43-49.

[Zipin, 1983] D.J.Zipin: The analysis of profile strata for surface texture specification. Appl. Surface Science, 15, 1983, 334-360.

[Zwierzycki, 1990] W.Zwierzycki: Wybrane zagadnienia zużywania się materiałów w ślizgowych węzłach tarcia. PWN, Warszawa 1990.

Authors' Information

Wiesław Graboń *– Department of Computer Science, Rzeszow University of Technology; W. Pola 2, 35-959 Rzeszów, Poland ; e-mail: wgrabon@prz.edu.pl*

Major Fields of Scientific Research: Software engineering information systems, Modeling of manufacturing processes.