

BIO-ALGORITHMS AND MED-SYSTEMS



Vol. 2, No. 3
2006

INFORMATICS

MEDICAL
SCIENCES

STATISTICS

COMMERCE

LAW

JOURNAL EDITED BY MEDICAL COLLEGE - JAGIELLONIAN UNIVERSITY - KRAKOW - POLAND

<http://www.bams.cm-uj.krakow.pl>

OPENING PAPER/ ARTYKUŁ OTWIERAJĄCY

- 3 Automatic Image Understanding a New Paradigm for Intelligent Medical Image Analysis
Ryszard Tadeusiewicz, Marek R. Ogiela

ARTIFICIAL INTELLIGENCE / SZTUCZNA INTELIGENCJA

- 11 Artificial intelligence computer algorithms in interpretation of selected CNS visualization
Lidia Ogiela, Artur Gądek

BIOINFORMATICS / BIOINFORMATYKA

- 17 Complex-network-based methodology for analysis of biomedical data
Sławomir Goryczka, Tomasz Arodź
- 25 The approach for computer analysis of results obtained in a two-dimensional electrophoresis of serum protein complexes with Congo red
Jacek Śmietański

BIOMEDICAL ENGINEERING / INŻYNIERIA BIOMEDYCZNA

- 33 Automated detecting symptoms of selected gallbladder illnesses based on a static ultrasound images analysis
Sławomir Bodzioch
- 43 Grid systems and their applications to biomedical science
Maciej Malawski, Tomasz Szepieniec, Irena Roterman-Konieczna
- 47 Stress state analysis in the femur after the implantation of an anatomical centega type stem
Wojciech Więckowski

MODELING AND SIMULATION / MODELOWANIE I SYMULACJA

- 53 Picture grammars in 3D modeling of medical structures
Marek R. Ogiela, Ryszard Tadeusiewicz, Mirosław Trzupek

TELEMEDICINE / TELEMEDYCYN

- 61 E-learning as a new tool in medicine teaching on the example of an introductory statistics course
Andrzej A. Kononowicz, Irena Roterman-Konieczna

E-LEARNING / E-NAUCZANIE

- 67 A tool for resuscitation training
Paweł Krawczyk, Grzegorz Cebula, Wiesław Pyrczak, Janusz Andres
- 71 Breast cancer – e-learning guide
Łasocha Bartłomiej, Matuszewski Przemysław, Pyrczak Wiesław
- 75 Computer science for pharmacy students. Verification of earned abilities by using the Internet
Jacek Śmietański

SHORT COMMUNICATIONS / KOMUNIKATY

- 79 Virtual reality – e-Education
Jan Trąbka
- 81 Małopolska Information Technologies Cluster – supporting the regional development
Sławomir Kopec

Automatic Image Understanding A New Paradigm for Intelligent Medical Image Analysis

RYSZARD TADEUSIEWICZ, MAREK R. OGIELA

*AGH University of Science and Technology, Institute of Automatics, Bio-Cybernetics Laboratory,
al. Mickiewicza 30, 30-058 Krakow, Poland*

Medical images are nowadays a very important and very popular source of information, used by medical doctors both for diagnosis and for planning of appropriate treatment. Therefore medical imaging as the methodology of physical investigations have become one of the most important tools in the whole of biomedical engineering, and also the market of imaging devices can be found as a very fast developing parts of Med-systems. Based on the variety of modern medical imaging devices, a physician can obtain different kinds of images, which present interesting organs and analyzed tissues in many different forms. One can use typical X-ray images (for general purposes Fig. 1 or specialized, like mammography Fig. 2, dentist pantomography Fig. 3), CT visualizations (Fig. 4), USG presentations (Fig. 5), nuclear images (Fig. 6) or thermograph plots (Fig. 7) – and many others. Not all of these images are easily readable, but an expert radiologist with good experience can manage all of them.



Fig. 1. Typical X-ray medical image for general purpose

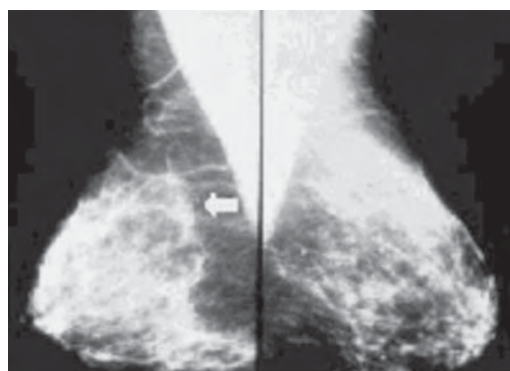


Fig. 2. Special kind of X-ray visualization: mammography



Fig. 3. Dentist pantomography

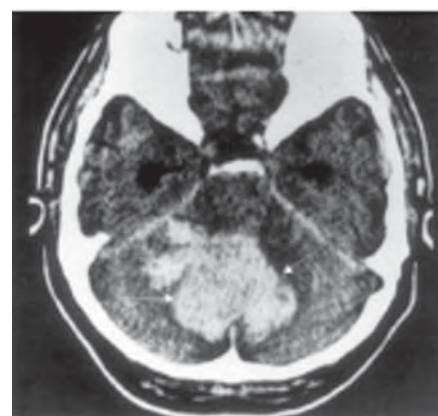


Fig. 4. CT visualization

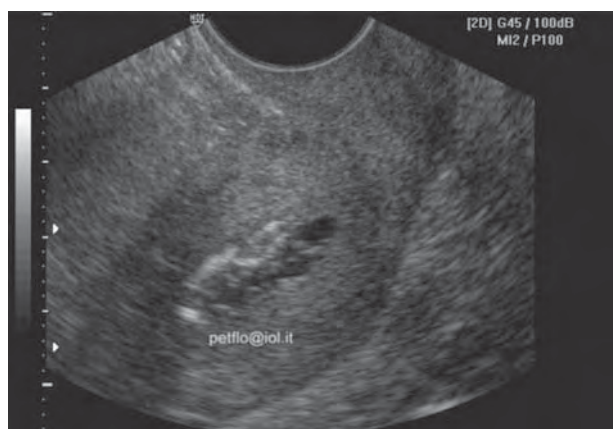


Fig. 5. USG presentation

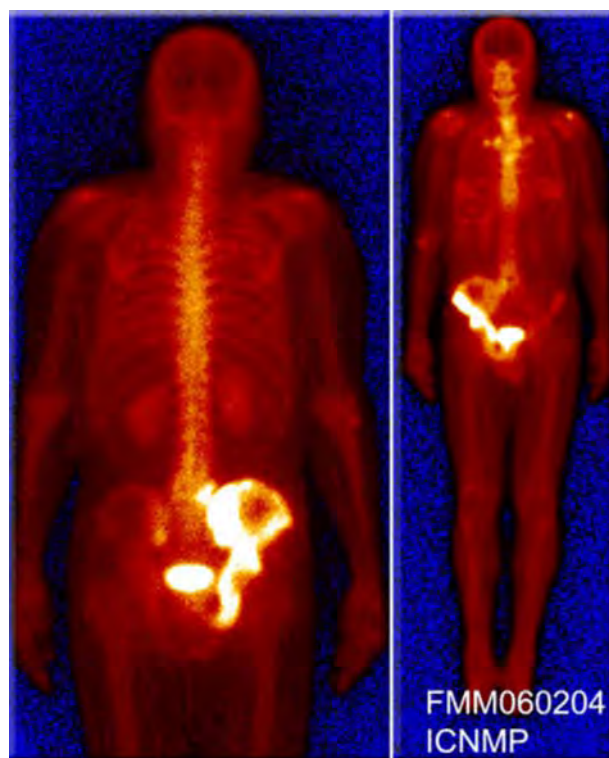


Fig. 6. Nuclear imaging

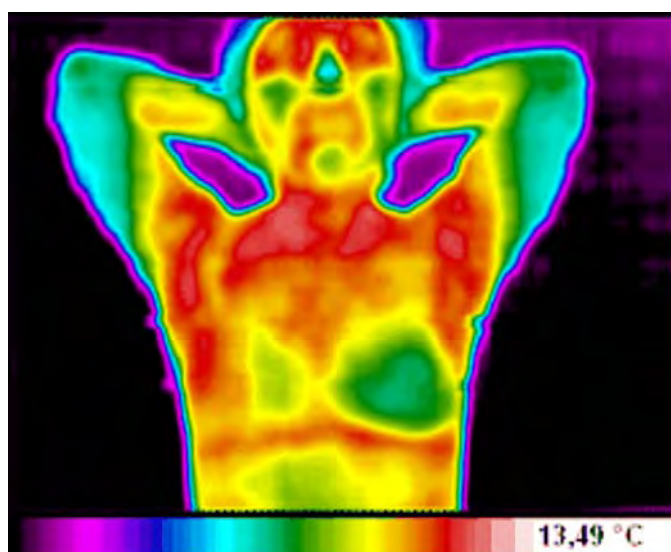


Fig. 7. Thermograph medical image

The abundance of forms and methods of medical imaging is a big opportunity of contemporary medicine, but at the same time, is a source of many problems, especially for young doctors or for specialists adapted to a interpretation of images produced by old-fashioned imaging devices, pushed now to using imagines of a new type, objectively much better than traditional ones, but in fact foreign to physicians and therefore difficult for an interpretation. The many sources and many kinds of medical images available now give a lot of visual data for every investigated patient. For the eyes of a radiologist expert such images become deep and valuable information about the morphology of organs under examination and about the health status of a patient. But not every physician is an expert radiologist, and a well educated end experienced doctor sometimes can be tired or absent-minded. It means a loss of information, which can be deducted from the medical image, but is neglected or omitted because of subjective (visual) methods used for its evaluation and interpretation.

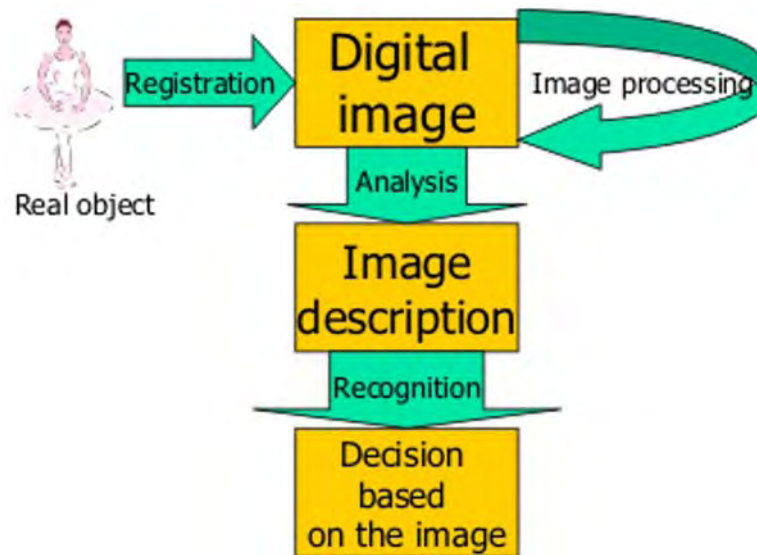


Fig. 8. Typical elements of image processing, analyzing and recognizing

One of the ways leading to an improvement of the analysis of medical images is using computers as advisors for the doctors. A computer course can not replace the doctor, but it can be very helpful in medical images' interpretation, especially for doctors who are not expert in a selected form of medical imaging. The computer vision method used traditionally to help medical images' interpretation performs three kinds of operations (Fig. 8).

The first are systems for **preliminary image processing**. Good and properly performed processing can be very helpful, because it eliminates a majority of noises present in the source image and can enhance the image quality especially in a sense of presentation interesting features of the image (e.g. sharpening of the view of the organ under consideration while all other organs are removed from the image or included in the background).

Second is **image analysis**. Under this very voluminous term, we can understand any calculation and measurement based on digital image, which gives parameters, useful for an interpretation of the image. Very important and valuable is, a reduction of information volume connected with such a step of image analysis. A typical medical image represents at last megabytes (sometimes gigabytes) of information volume and needs a lot of memory and computer power for storing and evaluation, while after properly performed analysis all necessary information about the important features of the objects on the image can be captured in the form of a vector containing a few dozen parameters.

The third step is **pattern recognition** named also **classification** of the image, and connected directly with the main goal medical image interpretation as a whole, e.g. medical diagnosis. Using very known methods of automatic classification, for example Bayesian rules, learning algorithms, k-mean or SVM methods and also the neural networks model – we can try to prepare a decision, which sometimes can automatically solve the problem under consideration.

Modern methods of computer vision, which is a field of informatics connected with artificial intelligence, allows a thorough **analysis** of images bringing out their parameters and numerically defining the selected features of the registered objects. We can also recognize defined types of objects. But all these conventional techniques of computerized handling of images (collected in the form of a compact diagram in fig. 8) do not move us nearer to an understanding of the meaning of what is revealed in the form of the image, as they refer to its form only.

Consequently, we do not know what it means that certain features and parameters of objects measured in the image have certain values calculated by the computer, because we are not able to link any meritorious **meanings** to these objects, parameters, or values. Of course, analysis of the image and parameters defined in result of such an analysis are often very useful in the case of simple assessments. We can compare, for

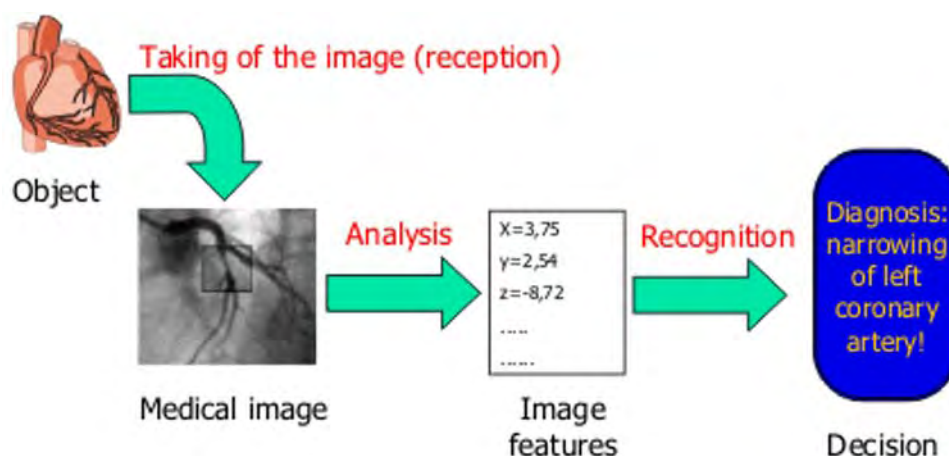


Fig. 9. Information flow diagram in normal image processing and recognition

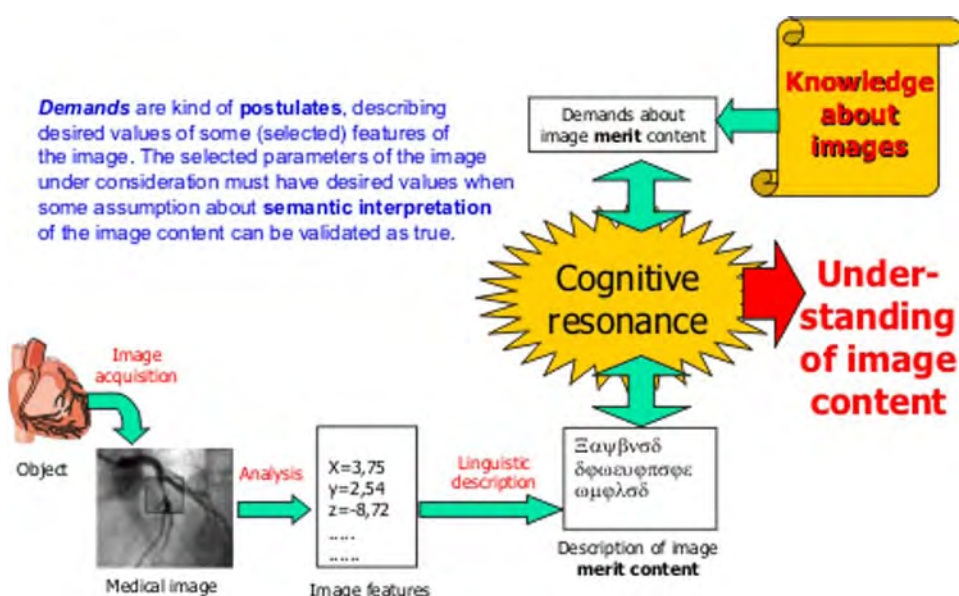


Fig. 10. Two-way information flow during the test of image understanding

instance, values referring to a given patient with the values recognized as "physiological standard". Image analysis gives us also very helpful **numerical values** of the features of selected objects (i.e. shapes and dimensions of organs or parts of organs). This is very important and useful.

Nevertheless all traditional steps of automatic processing and interpretation of medical images are not sufficient in the more difficult problems. In such problems we do need something more, over the processing, analysis and recognition, because for performing a full interpretation of a complex problem we must have the semantic interpretation of the image merit content. In fact the doctor's activity during the image interpretation is not devoted to measurement of some parameter or doing some classification. These are only the first steps on the long way leading to an **understanding** of a patient's problem and the nature of the disease.

The understanding process is always based on the physician knowledge – an this is the main difference between every method of the image processing, analysis and also recognition, which is ever **data-driven** procedure (Fig. 9), and the task performed by the doctors mind, which is all the time **knowledge-based** reasoning (Fig. 10).

Automatic image understanding allows to answer the question: What appears from what we see? What is the meaning of the fact that visible objects have certain features? What consequences arise from the fact that objects may be assigned to certain selected classes?

The following set of features of image understanding is consistent with the concept:

- A natural way of thinking of a human expert is imitated as closely as possible;
- The creation of a linguistic description of the meritorious content of the image allows its semantic representation without an aprioristic definition of a number of the recognized classes;
- The created linguistic description is of a structural character and allows us to analyze the meaning of the image, as detailed as needed for classification or indexation purposes.

Taking into account all the facts mentioned above we try to build mathematical models and also practical algorithms for an **automatic understanding of medical images**. The proposed method is based on the linguistic description of the images, which must be prepared for every kind of image under consideration (e.g. coronarography or urography images) on the base of a specially designed artificial *image content describing language*. After designing the structure of artificial language devoted to the description of the **merit sense** of the image under consideration, we must define a special kind of graph-grammar, describing the rules of the proposed language and of course prepare automatic procedures for extracting the necessary elements (so called graphical primitives and graphical relations) playing the role of “nouns” and “verbs” of defined grammar.

The structure of designed grammar must include as much as is available, elements of the doctor’s knowledge both about the illnesses under consideration, and about the structures and its morphological deformations, which can be visible on the images under consideration. It is difficult, and in fact it is the bottleneck of the presented methodology – but in numerous previous papers we prove, that this is possible and that it works properly with great efficiency.

When we have the proper (knowledge based!) grammar and we can convert each particular medical image (of selected type) to the merit based linguistic description, after which we can perform an automatic parsing (structural and semantic analysis). The parsing process used in our methodology to automatic understanding of the images is very similar to the process of automatic translation the text of the computer program form the selected algorithmic language (e.g. C++) to a binary code which can be performed directly by a microprocessor. As is commonly known, in the case of automatic translation of the programs as the input for the translation process we have the text of the program written by the programmer (not applicable to the microprocessor structures) and as the output we can obtain the same program transformed to the meaningful form of binary codes accepted by the hardware and ready for execution. In our case we have in the input of the parsing process a description of the analyzed image, expressed in the form of a sentence in special language, but this description is until now controlled only by the form of image and visualized medical structures. The parsing process ought to transform this description of the **form** of the image to the description of the **sense** of the image. If we can do this, we say, that we obtained the merit content of the image or that we **understood** the image.

Generally, such a structure of the image understanding system corresponds to one of the psychological models of a visual perception based on the concept of utilizing knowledge about the perceived objects, which will be discussed below.

During the parsing process two streams of information are combined and in some sense collided. The first stream starts from the analyzed image and brings all the details of the morphology of the observed organs. The second stream starts from the doctor’s knowledge represented in the from of a designed graph-grammar and developed parser. This second stream of information brings some demands or expectations to the forms which **can** be observed on the image when some interpretations of merit of the image become true. The confrontation between demands taken form the knowledge base and real parameters and features extracted from the input image is very similar to the interaction between two waves (for example during light interference). Some of the facts derived from this knowledge gains some features disclosed on the input image. Other features of the image can disclose a conflict with some expectations based on the knowledge, which leads to a changing of the working hypothesis about the understanding of the image merit sense. Because of this characteristics all

iteratively preformed process, connected with the parsing of the linguistic description of the image, based on the knowledge incorporated to the parser structure, which we called cognitive resonance.

Now, we must specify the above outlined concepts of the automatic image understanding to have them combined with the defined computing concepts and practically implemented. Looking for computable formulas which may correspond to the concepts of cognitive resonance between features of the image and expected on the basis of knowledge, and those features which are actually detectable in the topical image, we will refer to the **mathematical linguistics**.

At first, this decision may look a little odd, because a conventional application of linguistics is quite different – for instance, it is applied to create programming languages or text analysis. But such an approach is fully justified when we remember the reason why the convenient model of image *recognition* has been rejected and the search for a new paradigm, called automatic *understanding*, has begun.

As we have stated above, in the case of complex images analysis (especially medical ones), *a priori* indication of any **patterns** and an exact **number** of the recognized classes is not possible. Automatic image "understanding" must be identified with such a process of image analysis process that would result in creating *potentially unlimited* number of various conclusions. Therefore, it is not possible to select an answer from the preset collection, but we have to *generate* answers. This generation must be subjected to strict (grammar) rules and conducted in a manner enabling a precise analysis of the previously generated answers and reaching conclusions (mainly to find description classes defined in the semantic equivalent sense) – for instance, those connected with an identical diagnosis.

Thus, results of image understanding may be (according to the concept presented in this paper) expressed in the form of "sentences" recorded in a special language, and the process of concluding (about the "understood") images will be connected with a syntactic analysis of the generated sentences based on the grammar especially created for this purpose, and an automatic parser performing the necessary syntax analysis. In a majority of cases the dictionary, grammar and parser may be constructed in such a way that concepts connected with image understanding and concluding based on the understanding are naturally translated into the specific operations of a syntactic analyzer, as known from translations of formal languages. Unfortunately, the details of this process go beyond the scope of this paper, so we would like to encourage our reader to become familiar with other more extensive papers of the authors, listed below in the references.

References

1. Cios K. J., Mamitsuka H., Nagashima T., Tadeusiewicz R.: Computational intelligence in solving bioinformatics problems (Guest Editorial), *Artificial Intelligence in Medicine* (Elsevier), nr 35, 2005, pp. 1-8.
2. Ogiela M. R., Tadeusiewicz R., Ogiela L.: Intelligent Semantic Information Retrieval in Medical Pattern Cognitive Analysis, [in:] Gervasi O. et al. (Eds.): *Computational Science and Its Applications – ICCSA 2005, Lecture Notes in Computer Science*, vol. 3483, Springer-Verlag, Berlin – Heidelberg – New York, 2005, part IV, pp. 852-857.
3. Ogiela M. R., Tadeusiewicz R., Ogiela L.: Cognitive Vision Techniques in Medical Image Processing and Analysis. *Proceedings of HEALTHCOM 2005, 7th International Workshop on Enterprise Networking and Computing in Healthcare Industry*, IEEE Catalog Number 05EX1016, 2005, pp. 120-123.
4. Ogiela M. R., Tadeusiewicz R.: Picture Languages in Medical Pattern Knowledge Representation and Understanding, Program and Abstract of the MDAI 2005 -Modeling Decisions for Artificial Intelligence, Tsukuba, Japan, July 25-27, 2005, pp. 36-37
5. Ogiela M.R., Tadeusiewicz R.: Artificial Intelligence Structural Imaging Techniques in Visual Pattern Analysis and Medical Data Understanding, *Pattern Recognition*, Elsevier, 2003, vol 36/10 pp. 2441-2452.
6. Ogiela M. R., Tadeusiewicz R.: Syntactic reasoning and pattern recognition for analysis of coronary artery images, *International Journal of Artificial Intelligence in Medicine* (Elsevier), vol. 26, nr. 1-2, 2002, pp. 145-159.
7. Ogiela M. R., Tadeusiewicz R.: Intelligent Recognition in Medical Pattern Understanding and Cognitive Analysis, chapter in book: Muhammad Sarfraz (ed.): *Computer-Aided Intelligent Recognition Techniques and Applications*, John Wiley & Sons, Ltd., Hoboken, New Jersey 2005, pp. 257-274.
8. Tadeusiewicz R., Ogiela M. R.: *Medical Image Understanding Technology*, Springer-Verlag, Berlin – Heidelberg – New York 2004.
9. Tadeusiewicz R.: Automatic Understanding of Signals, chapter in book: Kłopotek A., Wierzchoń S., Trojanowski K. (eds.): *Intelligent Information Processing and Web Mining*, Springer-Verlag, Berlin – Heidelberg – New York 2004, pp. 577-590.

10. Tadeusiewicz R., Ogiela M. R.: New Proposition for Intelligent Systems Design: Artificial Understanding of the Images as the Next Step of Advanced Data Analysis After Automatic Classification and Pattern Recognition, [in:] Kwasnicka H., Paprzycki M. (eds.): Intelligent Systems Design and Applications, IEEE Computer Society Press, Los Alamitos, Washington, Brussels, Tokyo 2005, pp. 297-300.
11. Tadeusiewicz R., Ogiela M. R.: The New Concept in Computer Vision: Automatic Understanding of the Images. Chapter in book: Rutkowski L. (et all eds.): Artificial Intelligence and Soft Computing, Lecture Notes in Artificial Intelligence, vol. 3070, Springer-Verlag, Berlin – Heidelberg – New York 2004, pp. 133-144.
12. Tadeusiewicz R., Ogiela M. R.: The Processing, analysis, recognition, and automatic understanding of medical images. Chapter in Nowakowski A., Kosmowski B. B. (eds.): Optical Methods, Sensors, Image Processing, and Visualization in Medicine, Proceedings of SPIE, vol. 5505 (Progress in Biomedical Optics and Imaging vol. 5, no. 31), 2004 pp. 101-109.
13. Tadeusiewicz R., Ogiela M. R.: Structural approach to medical image understanding. Bulletin of the Polish Academy of Sciences – Technical Sciences, vol. 52, no. 2, 2004, pp. 131-139.
14. Tadeusiewicz R., Ogiela M.R.: Artificial Intelligence Techniques in Retrieval of Visual Data Semantic Information, [in:] Menasalvas E., Segovia J., Szczepaniak P.S. (eds.): Advances in Web Intelligence, Lecture Notes in Artificial Intelligence, nr 2663, Springer Verlag, 2003, pp. 18-27.
15. Tadeusiewicz R., Ogiela M. R.: Automatic Understanding of Medical Images – New Achievements in Syntactic Analysis of Selected Medical Images, Biocybernetics and Biomedical Engineering, vol. 22, nr 4, 2002, pp. 17-29.

THE ARTIFICIAL INTELLIGENCE COMPUTER ALGORITHMS IN INTERPRETATION OF SELECTED CNS VISUALIZATION

METODY SZTUCZNEJ INTELIGENCJI W AUTOMATYCZNEJ INTERPRETACJI OBRAZÓW CUN

LIDIA OGIELA*, ARTUR GADEK**

* AGH University of Science and Technology, al. Mickiewicza 30, PL-30-059 Kraków, Poland,
Faculty of Management, e-mail: logiela@agh.edu.pl

** MSWiA Hospital in Krakow, Trauma-Orthopedic Departament, ul. Kronikarza Galla 25, Kraków, Poland,
e-mail: drgadek@poczta.onet.pl

Abstract. In the paper were presented Artificial Intelligence methods, and in particular cognitive linguistic mechanisms used in semantic meaning inference, which allow one to create intelligent information system of the DSS type.

Additionally was presented computer algorithms of semantic description applied for object analysis in the case of recognition of spinal cord visualizations. The proposed semantic derivation procedures were based on cognitive resonance model, and were applicable to the task of semantic interpretation of selected central nervous system patterns. Towards effective recognition of pathological changes research were based on visual data set obtained during magnetic-resonance imaging of spinal cord.

Key words: IT systems, cognitive analysis, diagnostic support systems, central nervous system (CNS)

Streszczenie: W niniejszej pracy przedstawiono algorytmy sztucznej inteligencji, a w szczególności mechanizmy lingwistyczne służące do semantycznego wnioskowania znaczeniowego. Wykorzystanie prezentowanych mechanizmów lingwistycznych jest znaczące dla tworzenia inteligentnych systemów informacyjnych typu DSS.

Przedstawiono także komputerowe algorytmy opisu znaczeniowego analizowanych obiektów w zastosowaniu do przykładowej analizy obrazów rdzenia kręgowego. Zaproponowane procedury semantycznego wnioskowania oparte zostały o model rezonansu kognitywnego i jednocześnie zostały zaimplementowane do zadania semantycznej interpretacji wybranych zobrażeń centralnego układu nerwowego. W celu prawidłowego wykrywania poszukiwanych zmian patologicznych rdzenia kręgowego przeprowadzono analizę zbioru danych pozyskanych z badań magnetyczno-rezonansowych badanych struktur CUN.

Słowa kluczowe: systemy informacyjne, analiza kognitywna, diagnostyczne systemy informacyjne, centralny układ nerwowy (CUN)

1. Introduction

The development of civilisation has, for a number of years, been accompanied by a conviction that scientific research, when appropriately conducted, can ensure cognitive development related to all aspects connected with the mechanisms of learning and understanding. This belief was also to be found in times when the research of these phenomena was conducted with the use of instruments already available in the science. The pioneers of this approach were Hermann von Helmholtz, Wilhelm Wundt and Gustav Theodor Fechner. They were the first of those to lay down the grounds for psychology to be treated as a separate science [4]. From among many areas of psychology, the authors of this paper will quote mainly **cognitivism**, an experimental field of knowledge examining various manifes-

tations of the human mind's cognitive activities. Cognitivism became the foundation and the grounds for the design, thus for the development of one kind of IT system proposed not only for the needs of contemporary IT but also for the needs of medicine. This is due to the fact that cognitivism attempts to explain the perception and interaction taking place between a human being and a computer, optical and acoustic illusions, the rights governing multimedia-aided learning as well problems with human memory, its development, learning and knowledge structures [4].

The use and approach based on cognitivism in the field of IT systems is very widespread since we are dealing with the development of 'computational intelligence' methods. These, among others are the methods of machine recognition and pattern classification, multi-dimensional statistics that take into account issues such as clustering, discrimi-

nation and methods based on similarity, etc. Cognitive science can be considered to include also research on the optimisation and modelling of uncertainty, research in the field close to technical cybernetics such as sets and approximated logic plus the control and steering theory [9].

Cognitivism is therefore more and more frequently identified with many science and technology disciplines, including widely understood fields of information systems. For this reason the development of IT systems inherently connected with the above-described discipline of understanding and learning has become a certain challenge for the modern world of science. It is, however, impossible to propose and develop entire range of IT systems with no application and/or use in practice. Hence, an innovative attempt at creating an IT system for the needs of medicine. The system proposed by the authors is an intelligent system, based on computer diagnostic operational methods, currently very popular due to their extensive diagnostic capacities. This paper will demonstrate an example of the possibilities to develop a system that not only diagnoses, but one that is also oriented towards the issues of cognitive analysis and machine understanding of the pathological lesions taking place in the central nervous system, with particular reference to spinal cord disease lesions.

2. Description concept of the proposed diagnostic system

There are some description concepts (models) inherently associated with advanced and intelligent IT systems. The said models characterise the here-described systems. The most frequently used concepts are currently IT tools used to define such systems (it is worth mentioning here the UML language as one of the more important tools of this kind; in

connection with object methodology it is the most commonly used tool). Nevertheless to formulate modern solutions one also uses mathematical and logic IT description methods. There are even some philosophical concepts proposed [1]. Every one of these concepts uses different theoretical formalisms to execute the above-mentioned description (figure 1).

In order to be able to make a correct holistic analysis of the proposed IT system, one should take into consideration the above-presented description concepts. The authors of this paper propose, for intelligent IT systems, in particular those serving to understand image medical data, to use the most useful formalism based on a mathematical concept of IT systems. This is due to the fact that a mathematical formalism has the advantage of being precise and unique [6, 8]. It should be emphasised, however, that the two remaining concepts could also be used. Their additional advantage is that they are more legible than the mathematical concept.

In the system developed here, every medical image subject to analysis, constituting a kind of primitive component in the diagnostic IT system will be subject to analysis. As a result of the analysis we expect a response to the question as to whether a given patient has any disease lesions in the analysed organ. If so, whether the lesions are important, or, whether there are no lesions at all. If disease lesions can be found, then an expert will analyse them with a view to their kind. The latter equals finding an answer to the question about exactly what disease a given patient suffers from. The proposed diagnostic system functions on the basis of three main imaging analysis stages:

- **Image processing** performed in order to obtain the best quality possible and to expose the content the image carries
- **Image analysis** performed in order to get the properties of a given image in the form of features vector

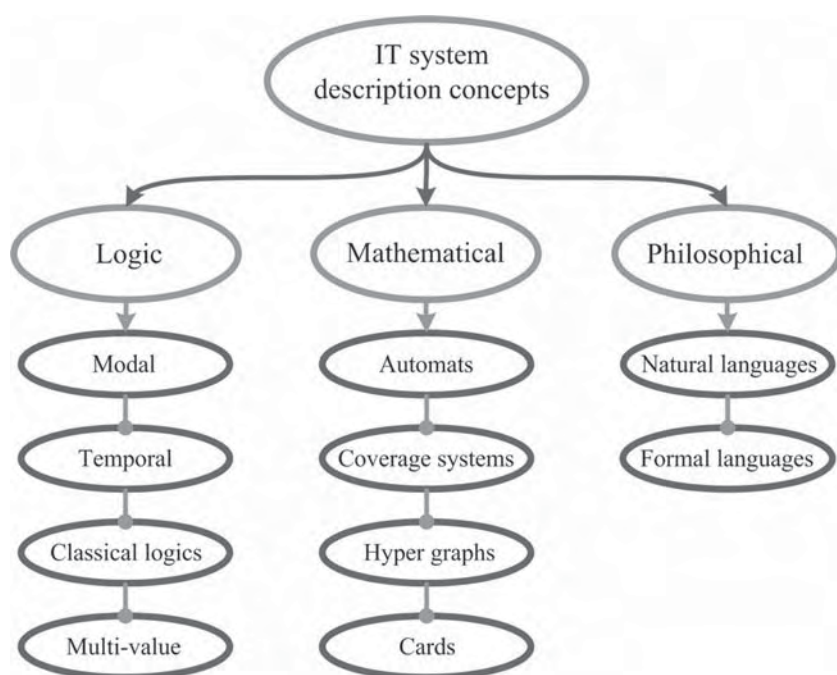


Fig. 1. IT system description concepts

- **Image recognition** leading to a classification of all features of the analysed image [7].

Intelligent diagnostic systems proposed in earlier research were used, among others, to diagnose pancreas, kidney and heart diseases. Their functioning is based on medical image recognition methods [8].

Due to the fact that diagnostic information systems are developed extremely rapidly, an attempt was made to construct a new class of such systems that would use cognitive analysis mechanisms for their operation, directed at attempts of automatic understanding of the analysed image semantics, thus at the interpretation of their contents.

3. Intelligent, cognitive IT systems of the diagnostic type

Cognitive IT systems of the diagnostic type have been developed based on systems whose purpose is not only to perform a simple analysis based on recording, processing and interpreting data, but first of all the analysis is based on understanding and reasoning performed in connection with the semantic content of the data processed.

Every IT system that analyses a selected image or information based on some features characteristic for it, has in its base some **knowledge**, necessary to conduct a correct analysis. This knowledge is the basis to generate a system of expectations relating to the conducted analysis.

This analysis is very frequently based on the syntactic approach [4, 6-8], which for image meaning interpretation uses an operation of image pre-processing, usually composed of:

- **Filtration and processing** of the input image,
- Image **coding** by means of terminal elements of the introduced language,
- Analysed object shape **approximation** [4, 7, 8].

After such stages have been completed it is possible to obtain a new image representation in the form of hierarchical structures of the semantic tree and successive steps of this representation derivation from the grammar start symbol [8]. The proposed system, already at the stage of image data pre-processing, must in most cases perform image segmentation, identify primitive components and determine what are the spatial and semantic relations between them. A correct classification is the recognition whether a given input image representation belongs to a class of images generated by a language defined by one of grammars possible to introduce. Such grammars can be included into sequential, tree and graph grammars while recognition with their use is performed in the course of the syntactic analysis conducted by the system [7].

In the most recent research on intelligent information systems, one can notice that the mere recognition of the analysed image is no longer insufficient and therefore there are more and more frequent attempts made at automatic understanding of the image semantics. In order to make this process feasible, there are attempts to use artificial intelligence techniques which enable researchers to extract important semantic information leading to meaning inter-

pretation of the analysed image, i.e. machine understanding [8].

This is an extremely complex process due to the fact that information flow in such cases is clearly two-dimensional. In the proposed model leading to image type data analysis, the stream of empirical data interferes with the stream of generated expectations [4, 8].

The application and use of systems described in this paper could be multiple due to wide possibilities offered to them by contemporary science. Nevertheless the greatest possibilities to use cognitive information systems are currently offered by medicine. This is due to the fact that there are an increasing number of disease lesions in developing pathological processes of individual organs as well as a growing ability to recognise these diseases. Medical images belong to some of the most diversified data. They also have an extremely deep and important meaning interpretation (for example for the patient's fate). Cognitive information systems could certainly serve in also many other fields of science and everyday life – if an attempt at the development of intelligent IT systems were made so as to add to other areas the process of understanding the analysed information or data.

The currently developed IT systems can try to understand data with some a-priori approach at what this understanding is supposed to serve. This must be differentiated from a situation in which a person, coming across a new situation analyses it in many respects. The results of this analysis can be completely unexpected conclusions such as (this is the case of medical images) diagnosing (the diagnosis being the same the computer came up with or a different one), or understanding that the image presented for understanding is bad quality because the machine taking it was out of focus and that the examination has to be repeated.

The terminological and conceptual base of the approach presented in this paper is the cognitive analysis (described above). At the current stage, however, it is better known in the context of psychology specialist considerations of human cognitive processes and in the context of human reason and wisdom hypotheses, as dealt with by philosophers specialising in the foundations of epistemology, gnoseology and semiotics as well as criteriology by D. J. Mercier and other advanced thought schools. This issue has so far been examined with a view to technical sciences to a smaller degree [4, 8].

4. A diagnostic IT system model proposed for spinal cord image analysis

An example of an intelligent IT system oriented at a complex analysis of image-type data is a certain model supporting the diagnostics of the central nervous system proposed in this paper. It deals in a special way with selected spinal cord diseases [4]. Due to the fact that the issue of occurrence of disease in the spinal cord is extremely extensive, we shall present below some selected pathological phenomena, typical for some central nervous system disease units.

The main element of a correctly functioning IT system supporting medical image diagnostics is to prepare a method aimed at the analysis of diseases and pathological lesions occurring in the spinal cord. The system would propose correct automatic interpretation methods of extremely complicated medical images, which are created as a result of imaging nervous system fragments. It is therefore extremely difficult to analyse such images due to the fact that there are various morphologies of the imaged organs in various patients. This is true both in the correct state and in the case in which some disease lesions appear. It is worth remembering that currently the anatomy and patomorphology differentiate a whole range of central nervous system developmental anomalies. Nevertheless it happens also that the system operates correctly for the first years and only after some time serious disorders in its functioning begin to appear.

In central nervous system image examinations the main role is now attributed to tomography (*Nuclear Magnetic Resonance, NMR*), also referred to as the MRI (*Magnetic Resonance Imaging*). This is imaging based on the phenomenon of magnetic nuclear resonance. Magnetic resonance makes it possible to obtain thickness maps (the so-called topography) primarily of hydrogen atoms (protons) and their relaxation times. Owing to the use of projection corresponding to the tomographic technique (computational reconstruction of the examined parameter location based on multiple multi-dimensional probing), the NMR image can be obtained on any body cross-section. Hydrogen is such an element of water found in abundant quantities in living organisms (60-70%) yet it is also a component of all organic compounds. It is worth remembering that extremely large quantities of hydrogen are found in lipids. The information obtained about hydrogen distribution inside the organism is the basis for image construction. Tissues are differentiated with respect to the water quantity or fat content.

Proton density and their relaxation times are mirrored by brightness (i.e. greyscale) of points found on a given map. The magnetic resonance method ensures a lot better-contrasted images of soft tissues than X-ray images. In the case of many diseases it also shows better the difference between the healthy and disease-struck tissue [2].

In order to perform analysis of disease lesions in the spinal cord we have proposed the following attributed grammar with the following form:

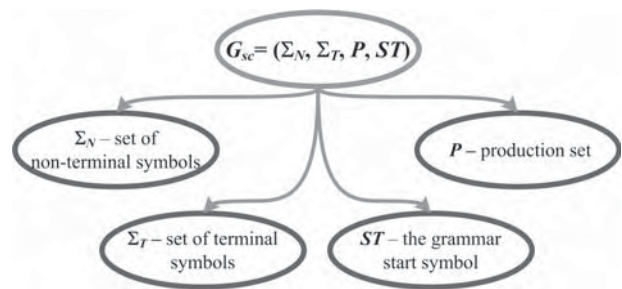


Fig. 2. G_{sc} sequential grammar

where:

$\Sigma_N = \{SPINE_LESION, SPINAL_STENOSIS, SPINAL_DILATATION, SPINAL_TUMOR, N, D, S\}$

$\Sigma_T = \{n, d, s\}$

Apart from these, the following meaning was given to terminal elements present in the description:

$n \in [-11^\circ, 11^\circ]$, $d \in (11^\circ, 180^\circ)$, $s \in (-180^\circ, -11^\circ)$,
 $ST = SPINE_LESION$

P production set has been defined as in table 1 below.

The proposed grammar makes it possible to detect various medullary or meningeal stenoses, medulla dilations: cysts and neoplasm tumours characteristic for various kinds of neoplasm lesions and spinal cord inflammatory processes.

Table 1 Production set defining changes in the G_{sc} grammar

Pathological lesion	Grammar rules	Semantic actions
Dilation/cyst	1. $SPINE_LESION \rightarrow SPINAL_DILATATION$ 2. $SPINAL_DILATATION \rightarrow D N S$ $D N$ $D S$	$Lesion = \text{spinal dilatation}$
Neoplastic tumours	3. $SPINE_LESION \rightarrow SPINAL_TUMOR$ 4. $SPINAL_TUMOR \rightarrow D S D S$ $S D S N$ $S D S D$ $D S D N$	$Lesion = \text{spinal tumor}$
Stenosis, compression	5. $SPINE_LESION \rightarrow SPINAL_STENOSIS$ 6. $SPINAL_STENOSIS \rightarrow S N D$ $S D$ $S N$	$Lesion = \text{spinal stenosis}$
Elements of the detected lesions	7. $N \rightarrow n \mid n N$ 8. $D \rightarrow d \mid d D$ 9. $S \rightarrow s \mid s S$	$Lesion \text{ features} =$ location; length; diameter, quantity, severity

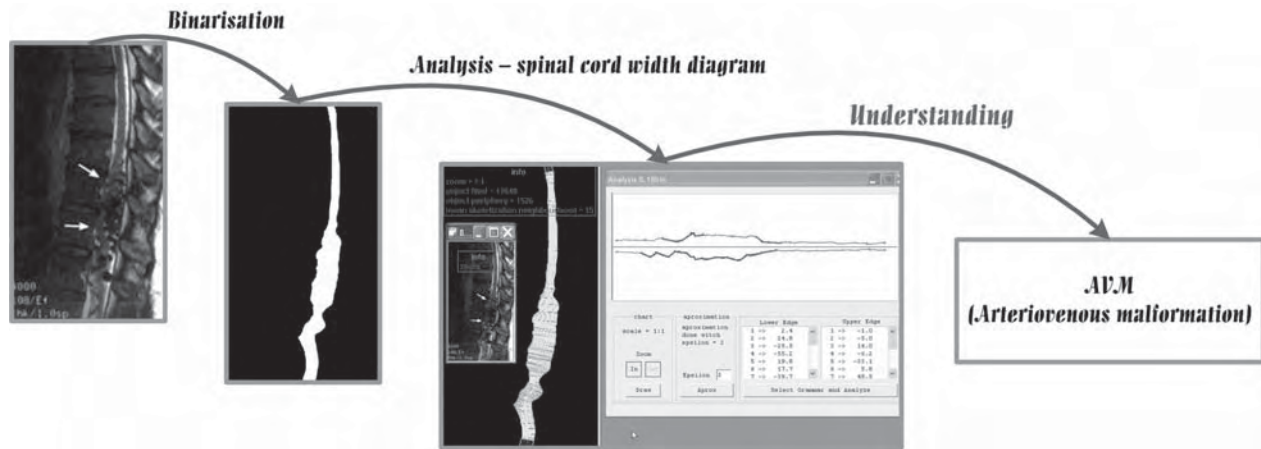


Fig. 3. Diagnostic descriptions of central nervous system lesions, containing a detected AVM syndrome

Figure 3 shows an image of the spinal cord with a visible deformation, the spinal cord image after binarisation and the spinal cord diagram. The bold line on the diagram shows the place in which there are anomalies in the structure of the spinal cord. A set of chords visible on figure 3, crossing the spinal cord in many successive points perpendicular to its axis shows how the width diagram was generated.

The spinal cord width diagram (figure 3) shows, in its most concise form, the results of that spinal cord morphology analysis. From the point of view of the performed medical image analysis, this is the most precious source of information since it can be used for looking for pathological lesions. It contains all-important data about the examined fragment of the central nervous system and at the same time it 'ignores' all such details of the spinal cord image that from the diagnostic point of view are meaningless.

For the example of spinal cord MR presented above, on figure 3, we have also supplied a diagnostic description of pathological lesions detected in the spinal cord. We have presented an example of a result obtained in the course of examinations performed for this disease. The results presented have been obtained by means of use of the attributed grammar and they are an example of approach based on the understanding of the analysed images containing medical data. The type of lesion detected has been specified based on its location and morphometric parameters determined by grammar semantic procedures.

The above-quoted example (and many others obtained in the course of the conducted research [4]) shows the results of semantic, meaning interpretation of the analysed and detected pathological lesions occurring in the spinal cord area.

5. Summary

In order to perform meaning analysis on spinal cord images with the use of the linguistic mechanism described in this paper, the MISA (Medical Image Syntax Analyser) computer system has been developed. It enables the analysis and classification of spinal cord images analysed in this paper.

The application efficiency of procedures proposed in this paper, for the analysis of central nervous system medical

images with the use of this system is very high, since in the case of the 84 analysed images containing spinal cord pathological lesions it amounted to 90.5%. This means that 76 images were correctly recognised as a result of the method proposed in this paper.

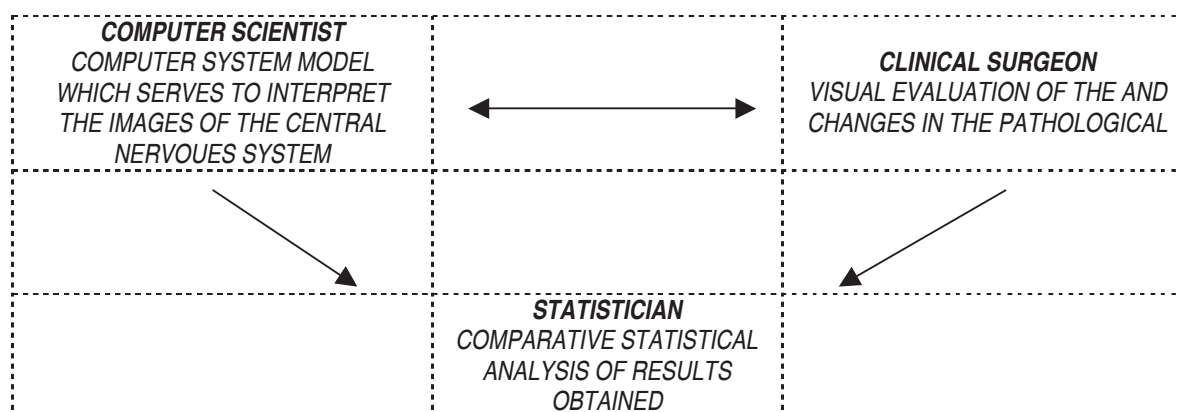
These results are obtained as a result of the application of semantic analysis algorithms conducted in reasoning modules of the proposed system, based on semantic actions assigned to structural rules.

The research conducted by the authors, based on analysis of images with pathological lesions in a part of the central nervous system, the spinal cord, have demonstrated that cognitive data analysis can be a factor significantly enriching possibilities of contemporary information systems. In particular, the described research has demonstrated that an appropriately built image grammar makes it possible to conduct a precise analysis and to describe medical images, to gain from them important semantic information on the nature of processes and pathological lesions as found in the patient's spinal cord. It is worth emphasising that the results described in this paper have been obtained owing to the cognitive process, simulating an expert's method of thinking: if he/she sees a deformation of the organ shown by the medical image used, he/she tries to understand the pathological process that was the reason for the appearance of deformations found. He/she does not perform a mechanic classification only to point to the most similar sample of pathological image. Moreover, the research conducted has demonstrated that for cognitive analysis attempts (on the central nervous system) it is possible to apply sequential grammar-based linguistics.

References

1. J. S. Albus, A. M. Meystel: Engineering of the Mind – An Introduction to the Science of Intelligent Systems. A Wiley-Interscience Publication John Wiley & Sons Inc, 2001.
2. F. A. Burgener, S. P. Meyers, R. K. Tan, W. Zaunbauer: Differential Diagnosis in Magnetic Resonance Imaging. Georg Thieme Verlag, 2002.
3. A. Gądek: Wytrzymałość trzonów kręgowych a ich mikrostruktura i gęstość mineralna. Rozprawa doktorska, Collegium Medicum UJ, Kraków 2001.

4. L. Ogiela: Ocena użyteczności metod analizy kognitywnej w wybranych systemach informacyjnych. Rozprawa doktorska, AGH, Kraków 2005.
5. L. Ogiela: Metody analizy kognitywnej w medycznych systemach informacyjnych typu DSS. Computer Science, 7 (2005).
6. M. R. Ogiela, R. Tadeusiewicz, L. Ogiela: Intelligent Semantic Information Retrieval In Medical Pattern Cognitive Analysis. Computational Science and Its Applications – ICCSA 2005, 3483, pp. 852-857.
7. R. Tadeusiewicz, M. Flasiński: Rozpoznawanie obrazów, Wyd. Naukowe PWN, Warszawa 1991.
8. R. Tadeusiewicz, M. R. Ogiela: Medical Image Understanding Technology, Artificial Intelligence and Soft-Computing for Image Understanding. Springer-Verlag Berlin Heidelberg 2004.
9. K. Wenta: Kognitywistyka w multimedialnych i sieciowych systemach informacyjnych, w: Daniłowicz Cz. (red.), Multimedialne i sieciowe systemy informatyczne, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2002, s. 65-70.



COMPLEX-NETWORK-BASED METHODOLOGY FOR ANALYSIS OF BIOMEDICAL DATA

METODOLOGIA ANALIZY DANYCH BIOMEDYCZNYCH BAZUJĄCA NA SIECIACH ZŁOŻONYCH

SŁAWOMIR GORYCZKA, TOMASZ ARODŹ

*Institute of Computer Science, AGH University of Science and Technology,
al. Mickiewicza 30, 30-059 Kraków, Poland, tel: +48 12 617 3497, fax: +48 12 633 9406,
slaweg@student.agh.edu.pl, arodz@agh.edu.pl*

Abstract: Following the rapid growth of high-detailed information on genomes and proteomes, methods for deducing properties of living organisms at the system level are in need. One approach for capturing the global characteristics of biological entities is to model them as complex networks. Such networks can then be analysed with descriptors tailored to reveal their static and dynamic properties. With metabolic and protein-protein interaction networks of two bacteria as examples, we show here the methodology for carrying such an analysis using a set of informative descriptors.

Key words: complex networks, graph descriptors, metabolic network, protein-protein interaction network

Streszczenie: Wraz ze wzrostem ilości zgromadzonych informacji o genomie i proteomie, wzrasta zapotrzebowanie na metody pozwalające na wnioskowanie o cechach badanego organizmu na poziomie systemowym. Jedno z podejść do tego problemu ujmuje globalne cechy organizmu jako złożoną sieć. Może ona być następnie analizowana przy pomocy deskryptorów zaprojektowanych pod kątem uwypuklania statycznych oraz dynamicznych cech systemu. Bazując na przykładach dotyczących sieci metabolicznych oraz zbudowanych w oparciu o interakcje białek dla dwóch mikroorganizmów, w niniejszej pracy prezentujemy metodologię analizy opartej na deskryptorach sieci złożonych.

Słowa kluczowe: analiza sieci złożonych, deskryptory grafowe, sieci metaboliczne, sieci interakcji białek

1. Introduction

In recent decades, life sciences have gone to great lengths to discover and understand the molecular basis of biology and medicine. This effort has resulted in a spectacular growth in our knowledge of detailed mechanisms of action and the structural organization of cells, tissues and organs. The vast quantities of available information on genes, proteins and metabolites, while far from being complete, nonetheless allow for making some generalising observations on the nature of processes and structures in biology.

Unlike during previous attempts at understanding the organisms in a holistic way, the scientists today have at hand databases rich with genomic, proteomic or metabolomic data. Thus, the understanding of a cell, tissue or organ as a well-defined biological system can now be deeply founded in high-detail molecular data. On the other hand, the methods for effectively inferring the properties of any highly complex biological system as a whole, from the partial knowledge about its components and their inter-relations, are still under development.

One of the models used to represent the individual constituents of a biological system and their relations is that of a network. In such a model, the information on e.g. the cell is encoded as a set of vertices with connecting edges. Furthermore, for a given network, the graph that defines it may be evolving in time, reflecting the changes in the system. An intuitive example of such a network-based model is the description of protein-protein interactions, with individual proteins as vertices, and the fact of forming a complex modelled by an edge linking the two proteins.

Modelling the biological system as a network allows for the use of various elaborate mathematical techniques for statistical analysis of graphs [1,17]. The analysis of networks of biological origin, among others, has led to a change of perspective in the mathematical analysis of networks in itself. Until recently, the network analysis field focused on graphs resulting from e.g. industrial applications. Their distinctive feature was that the vertices had similar number of neighbours. For example, the Erdos and Renyi [10] model assumed a random process of graph creation, which results in normal distribution of the number of neighbours of

vertices. Other investigated networks were of relatively regular structure, e.g. the railroad network.

In [2], Barabasi and Albert have shown that in many real-world systems, the probability $P(k)$ that a vertex in the network interacts with k other vertices decays as a power law, i.e., scales like $k^{-\gamma}$. This result indicates that large networks self-organise into a scale-free state, a feature unpredicted by previous random network models. Scale-free networks have few vertices known as hubs, each with a large number of neighbours. The hubs form the skeleton of the entire network, which makes them very fragile to malicious attacks. If we remove few random vertices from the network, usually nothing will change. However, if we start to remove hubs, after few steps we will change the topology of the network significantly. This feature of scale-free network has been investigated e.g. in [8] by Crucitti et al. Another phenomenon observed in biological networks is the small-world property [23]. The distances between vertices, even in large graphs, are usually relatively small, as a result of many shortcuts present in the structure.

The above-mentioned quantities, the number of neighbours of a vertex and the number of vertices in paths between a pair of vertices, are just two of a broader family of graph descriptors. Here, we analyse a number of graph descriptors in the context of biological systems. We show their usefulness in capturing the general properties that emerge from the simple, local interactions modelled by a graph.

The rest of the paper is arranged as follows. In Section 2, we define the network more formally, and then focus on presenting the descriptors of graph structure that can be used to capture the properties of the network. In Section 3, we apply these tools to metabolic and protein-protein interaction networks. We show how can they be employed in analysing large networks. Finally, Section 4 summarizes the conclusions of the paper.

2. Analysis of Complex Networks

A complex network is defined here as a graph, i.e., a set of vertices with edges connecting them and rules introducing dynamic changes of the graph structure. More formally, a network N is a triple (V, E, S) , where V is a set of vertices in specified point of time, E – set of edges in the point of time and S is a set of operations (also randomised operations), which alter the topology of the network. Algorithms from the set S are allowed to use atomic operations to change the structure of the network, e.g. addition of a new vertex or edge to the network. The algorithms define the evolution of the network in time, e.g. under some form of perturbation of its constituting vertices or edges.

The networks in which complex behaviour emerges are usually of considerable size. Thus, the inspection of their properties is not straightforward. To address this issue, various methodologies have been used [17]. In particular, the analysis using descriptors that capture statistical properties of a graph, and their evolution in time has been recently employed [2]. These descriptors often reuse or extend concepts successfully used e.g. in defining topologi-

cal quantitative structure-activity relationships (QSAR) in drug discovery [9]. However, in order to be useful for analysing complex networks, the descriptors have to meet several criteria. First, they have to capture the dynamic changes in the structure of a network, which is not present e.g. in static chemical structures used in QSAR. Second, the size of analysed networks may be very large, constraining the computational complexity of the descriptor calculation.

The analysis of complex networks requires descriptors that allow for analysis of both its static and dynamic properties. To this aim, we provide two types of information. For capturing of static structural properties, we employ distribution plots of relative frequencies of quantities describing the nature of the graph in perspective of a neighbourhood of its constituting vertices and edges. For capturing dynamic behaviour, we resort to single-number descriptors and investigate their changes in relation to the changes in the network.

2.1. Descriptors of Static Structure of Complex Networks

For analysing the static nature of the complex network, we can use the distribution of quantities related to both the local neighbourhood of the vertices, as well as of the edges. Below, we present three of such descriptors.

Vertex degree distribution

The vertex degree $\deg(v)$ of a given vertex v in a graph is the number of edges linking it to other vertices. This information allows for creating a highly informative method for inspecting the graph's topology, by creating a plot with empirical distribution of degrees of vertices in a graph. Such a distribution is useful in capturing the scale-free nature of the graph.

Edge connectivity distribution

The edge connectivity descriptor, also referred to as the product of the degrees, shows the local weakness of each edge separately.

Let $G=(V, E)$ be a graph and e an edge in E . The edge connectivity for edge $e=(v, w)$, $EConn(e)$, is a product of degrees of incident vertices v and w in V

$$EConn(e)=\deg(v) \deg(w). \quad (1)$$

If connectivity of an edge is high, the edge is then strongly connected because it connects two vertices with high degrees. For these vertices loss of the edge is not so important for their connectivity with other vertices compared to a weakly connected edge. The weakly connected edge is incidence to vertices with low degrees. For these vertices every edge is important to keep links with other vertices in the graph. The complexity of computation values of this descriptor for one edge is $O(1)$ and for all edges $O(m+2n)$, where $n=|V|$ and $m=|E|$.

Edge coherence distribution

In [23], Duncan J. Watts has proposed the range of an edge, or edge coherence, as a new descriptor. This descriptor is supplementary to the edge connectivity. Its

values show how important for two neighbouring vertices is the edge incident to both of them.

Let $G=(V, E)$ be a graph and e in E and $G'=(V, E\setminus\{e\})$. The range of edge $e=(v, w)$, $g(e)$, is a distance, i.e., the number of vertices to be traversed on the path, between vertices incident to e in graph G'

$$g(e) = d_{G'}(v, w), \quad (2)$$

If value of descriptor is low, the edge is locally weak. A weak edge can be removed with minimal increase in the distance between the two vertices it connects. An edge with low range could be substituted by the shortest path, which connects the endpoints of this edge.

If the value of the range for an edge is high then this edge is very important for vertices connected by it (locally strong edge). When this edge will be removed then the shortest path connecting these vertices becomes quite long. If $g(e)=8$, then edge e is required to keep the graph connected.

Computational complexity of the algorithm used for calculating the value of this descriptor for one edge is $O(n+m)$ and consequently for all edges in the graph is $O(m(n+m))$, where $n=|V|$ and $m=|E|$.

2.2. Descriptors Used in Analysis of Network Dynamics

Analysing the distributions of values for a larger set of graphs that capture the evolution of the network is not feasible. Instead, we will use a more compact representation, in which each graph is quantified by a set of single-number descriptors of its overall structure.

Connectivity

Connectivity of a graph G , $Conn(G)$, is a ratio of number of edges to maximal possible number of edges. Let $G=(V, E)$ be an undirected graph without self-loops and multiple edges and let $n=|V|$, and $m=|E|$.

Connectivity of graph G is given by the following formula

$$Conn(G) = 2m/(n(n-1)). \quad (3)$$

This descriptor is very easy to compute and is interpreted as the density of edges in the graph. The minimal estimated computational complexity for this index is $O(1)$ if we have edges and vertices in a graph counted. Otherwise we will have to count all vertices and edges and complexity increase to $O(n+m)$.

Total Walk Count

In [21], Gerta and Christoph Rucker have proposed a new measure of the graph's complexity, named Total Walk Count (TWC). The TWC is obtained by counting all paths with l edges originating from any vertex v , $l w_v$, of all lengths l . Let $G=(V, E)$ be a graph and $P(v, x)$ be a path linking vertices v and x in graph G

$$l w_v = |\{P(v, x) : d(P(v, x)) = l\}|. \quad (4)$$

The maximum length of any walk w is limited by the number of edges in the graph.

Randić Index

In [20], Milan Randić has introduced a new connectivity index. It was used for the first time in describing chemical molecules, and has been later generalised by Lemont Kier and Lowell Hall [15]. The Randić connectivity index for graph $G=(V, E)$ follows the formula

$$\chi(G) = \sum_{(v, w)} (\deg(v) \deg(w))^{-1/2}, \quad (5)$$

where the sum is taken over all possible edges (v, w) between vertices v and w in E . The computational complexity for calculating this index is $O(m)$.

Zagreb Indices

In [12], Ivan Gutman and Nenad Trinajstić have introduced a pair of indices M_1 and M_2 referred to as Zagreb indices. For graph $G=(V, E)$, the Zagreb indices are defined as follows

$$M_1(G) = \sum_v \deg(v)^2, \quad (6)$$

$$M_2(G) = \sum_e w_e, \quad (7)$$

where w_e is the weight of edge e . For unweighted graphs we can assume that for each edge e , $w_e=1$.

A problem with Zagreb indices is that their contributing parts give greater weights to inner (interior) vertices and edges and smaller weights to outer (terminal) vertices and edges of a graph. In [18], Sonja Nikolaiæ et al. have introduced a way to amend Zagreb indices by inserting inverse values of the vertex-degrees into equations above. They call these indices: modified Zagreb indices and denote them as ${}^m M_1$ and ${}^m M_2$. The following formulas are necessary for calculating ${}^m M_1$ and ${}^m M_2$

$${}^m M_1(G) = \sum_v \deg(v)^{-2}, \quad (8)$$

$${}^m M_2(G) = \sum_e w_e^{-1}. \quad (9)$$

This set of indices is very easy to compute with computational complexity $O(n)$.

Subgraph Count

The Subgraph Count for a graph G , ${}^k SC(G)$, is the number of connected subgraphs with k edges

$$\begin{aligned} {}^k SC(G) &= |\{H = (V_H, E_H) \subset G : |E_H| = k \wedge \text{diam}(H) < \infty\}| \end{aligned} \quad (10)$$

These descriptors show how many subgraphs with k edges occur in the graph. We can also normalize these descriptors and denote it as, ${}^k SC_n(G)$

$${}^k SC_n(G) = {}^k SC(G) / {}^k SC(K), \quad (11)$$

where K is a complete graph (clique) with n vertices.

High-order descriptors that belong to this set are not computable in most cases. The most important descriptors from this group are presented below. In [19], John R. Platt has shown that number of paths composed of two edges is a good descriptor of molecules in chemistry. The Platt index for graph G is the subgraphs count descriptor, ${}^2 SC(G)$. The Platt index

for the complete graph K is maximal. This value is used to normalize the value of the Platt index for any graph with the same number of vertices, which is equal to n . The minimal estimated computational complexity for this index is $O(n)$.

In [11], M. Gordon and G.R. Scantlebury noticed that also a subgraph count descriptor ${}^3SC(G)$ is a good factor for analysing graph topology. The Gordon-Scantlebury index for a graph G , ${}^3SC(G)$, is the number of connected subgraphs with three edges. For the complete graph K with n vertices, the value of the Gordon-Scantlebury index reaches a maximum. Contrary to the Platt index, subgraphs with 3 edges can be of three types, i.e., looks like a triangle, line or a star. We count the number of every kind of three-edge subgraphs separately. The computational complexity of this descriptor is: $O(nm(2m-1))$. In [3], the previously used subgraph counts have been extended to the Total Subgraph Count descriptor, which is the sum of subgraph counts of all possible degrees.

It is worth mentioning that connectivity is the normalized first-order term in the series of subgraph counts

$$Conn(G) = {}^1SC_n(G). \quad (12)$$

Information-theoretic descriptors

In [6], Danail Bonchev and Nenad Trinajstić have introduced an approach to characterizing graph structures using the Shannon's theory of information [22]. They studied molecular branching as a basic topological feature of molecules. This approach was generalized by constructing a finite probability scheme for a graph [4].

Let the graph be represented by some kind of N primitives (vertices, edges, distances, cliques, etc.); let us also assign a certain weight (degree, value, magnitude) w to each of the N elements. The probability that a randomly chosen element i has weight w_i is defined as follows:

$$p_i = w_i / w, \quad (13)$$

$$w = \sum_{i=1}^N w_i. \quad (14)$$

The most important information-theoretic descriptor stems from the equation that describes information content of the vertex degree distribution in the graph $G=(V, E)$, $I_{vd}(G)$

$$I_{vd} = \sum_v \deg(v) \log_2 \deg(v). \quad (15)$$

In [5], Danail Bonchev et al. have shown, that value of I_{vd} increases with the connectivity and with other complexity factors, such as the number of branches, cycles, cliques, etc. The minimal computational complexity estimated for calculating this index is $O(n)$.

Cluster Coefficient

The vertex degree, $\deg(v)$, is not the only local connectivity descriptor. More detailed information on the vertex neighbourhood is contained in the cluster coefficient, c_v . Vertex w is a k^{th} -order neighbour of vertex v in graph $G=(V, E)$ if, and only if, $d(v, w)=k$.

Let $G=(V, E)$ be a graph and ${}^1E_G(v)$ be the number of edges between i^{th} -order neighbour of vertex v in V . Cluster coefficient of vertex v is defined as the ratio of ${}^1E_G(v)$ and the respective number of edges in the complete graph K , ${}^1E_K(v)$,

$$c_v = 2 {}^1E_G(v) / (\deg(v) (\deg(v)-1)) \quad (16)$$

Complexity of this calculation for one vertex v is $O(\frac{1}{2}\deg(v)(\deg(v)-1))$ and for all vertices in graph G is $O(\sum_v \frac{1}{2}\deg(v) (\deg(v)-1))$.

If we consider the handshaking lemma [13] and maximal value of degree for one vertex we can limit value of this complexity by a more easily interpretable term $\frac{1}{2} n (n-1)^2$.

When we take into account the second and the following neighbourhoods, we obtain a large group of cluster coefficients. The local connectivity is very important because it shows density of connectivity in the entire graph or network. The mean or distribution of this density is very important information in investigating networks.

3. Analysis of Complex Biological Networks

We have inspected three biological networks to show the potential applications of complex-network-based analysis. Two types of networks are investigated – the metabolic network and the protein-protein interaction network. We focus on two Gram-negative bacteria, *E. coli* and *H. pylori*. The data on metabolic network of *E. coli* were obtained from [16]. The vertices represent metabolites, which are connected by an edge if they are present in the same reaction. The *E. coli* protein-protein interaction maps were obtained from [7] and are a result of mass spectrometric analysis of complexes of proteins. For the *H. pylori* protein-protein interaction network, the data based on the DIP database, gathered by the European COSIN¹ project were used.

3.1. Scale-free Nature of the *E. coli* Metabolic Network

Recently, it has been shown [16] that the metabolic network of *E. coli* is scale-free with respect to the vertex degree. That is, the abundance of vertices is negatively correlated with their degree, with the lowest number of highly connected metabolites. While the above is not surprising,

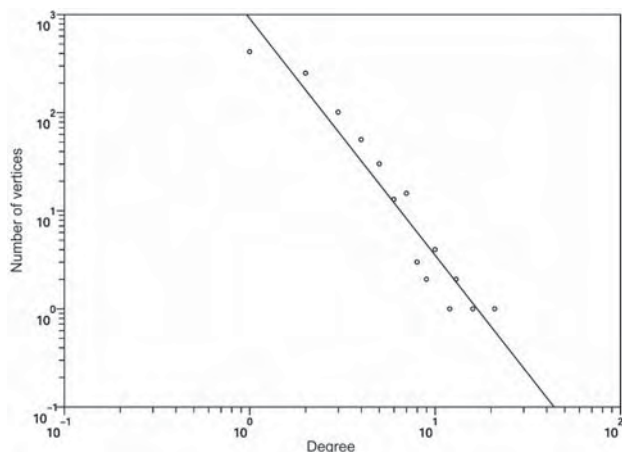


Fig. 1. Relative frequencies of vertex degrees in *E. coli* metabolic network (log-log scale)

¹ <http://www.cosin.org/extra/data/proteins/>

the studies have shown great regularity in this distribution, as shown in Fig. 1. The relation closely follows the power law, i.e., is linear on a log-log plot.

We argue here that the scale-free nature of complex biological networks is captured not only by vertex degree distribution. We show that it is also apparent in information on edges and paths in the network. In particular, the edge information is especially informative for metabolic network, as the edges model the reactions that give rise to the metabolites represented by the vertices.

For *E. coli* metabolic network, we have calculated the distribution of abundances of various values of edge coherencies (2). As noted above, for two adjacent vertices, this descriptor measures the averaged smallest paths between them not leading through the edge connecting those vertices directly. The resulting plot, shown in Fig. 2, shows that in most cases, the removal of a certain metabolic link can be circumvented with only few links. Cases where larger number of additional links is required are scarce. Again, these results allow for establishing a power-law governing the distribution of lengths of circumventing paths in *E. coli* metabolic network.

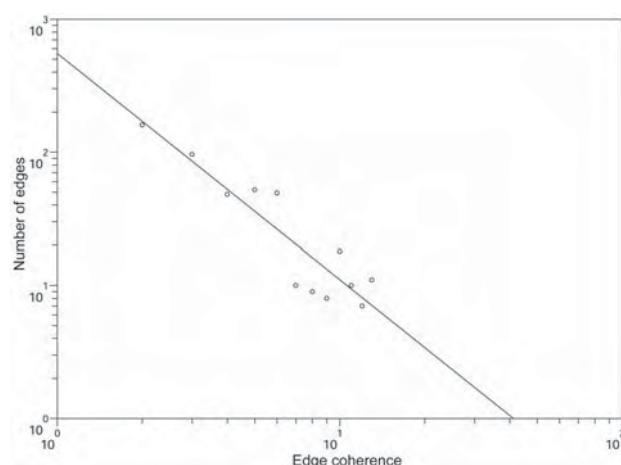


Fig. 2. Relative frequencies of edge coherencies in *E. coli* metabolic network (log-log scale)

This finding is supported by another edge-related distribution, of edge connectivity (1) values (Fig. 3). The linear log-log form of the distribution supports the information gained from distribution of vertex degree, that poorly connected vertices are the most frequent. However, the vertex degree distribution reveals only some limited information on the structure of the network. The edge connectivity information provides additional insight, showing that the most frequent edges are between vertices that are both poorly connected. This suggests, that the network differs from typical scale-free networks such as airport connections. In addition to hub-like structure evident from vertex degree distribution, it contains many links between non-hub, low-connectivity vertices that bypass the hubs. This information correlates with the information from the edge coherence graph, that a lot of short alternative paths are present in the *E. coli* metabolic network.

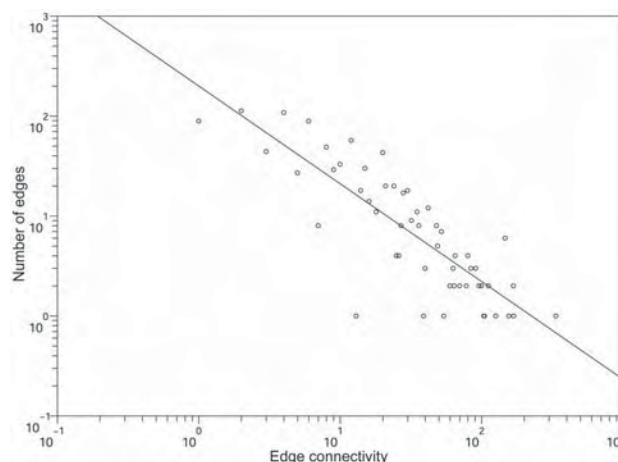


Fig. 3. Relative frequencies of edge connectivities in *E. coli* metabolic network (log-log scale)

3.2. Investigating the Resilience of *E. coli* Protein-Protein Network

The graph descriptors described in Section 2 allow for easy inspection of the behaviour of the network in case of certain events that disturb its normal structure. To show this, we investigate the resilience of the protein-protein interaction networks towards eliminating a small number of proteins, e.g. by the knockout or mutation of the corresponding genes. We show that the set of descriptors is geared towards detecting events that drastically change the structure of the network, altering significantly the connectivity patterns in the interaction network.

We analyse two types of vertex elimination – random and aimed at the most connected proteins. In the first scheme, we recalculate the descriptors for networks with 2, 4, 6, 8, and 10 randomly chosen vertices eliminated from the graph. The results are gathered in Table 1. One can see, that such random events result in a response that is proportional to the magnitude of introduced perturbation measured by the number of removed vertices.

On the other hand, an attack at the key vertices, i.e., the ones being the most highly connected, results in a response that is non-linear with the magnitude of the stimulus (see Table 2). Only two descriptors, the average cluster coefficient (16) and the Randić index (5), remain proportional to the magnitude of the introduced perturbation. However, these descriptors capture the most local information on the structure of the graph.

The analysis of other descriptors reveals two other groups of graph indicators. The first one, which includes connectivity (3) and 2nd Zagreb index (7), follows the decrease in the number of edges in the graph. Indeed, the definitions of those two entities are proportional to the number of edges in the graph. The rest of the descriptors respond with a decrease in value stronger than both the drop in the number of vertices and edges, indicating a structural change in the network.

The methodology of analysis shown above can be used to estimate the importance of the given vertex, e.g. a protein, for the inspected network. The more non-linear is the

Table 1. *E. coli* protein-protein interaction network descriptors for original graph (G) and after removal of 2, 4, 6, 8 and 10 random vertices. Percent change in the value of descriptors for 10 vertices removed

Descriptor	G	G-2	G-4	G-6	G-8	G-10	G-10 Change [%]
V	270	268	266	264	262	260	-3.7
E	716	713	707	695	687	684	-4.4
c_v	0.579	0.579	0.581	0.579	0.581	0.581	0.4
l_{vd}	4661	4651	4624	4527	4469	4452	-4.4
X	110.7	109.7	107.8	106.9	106.2	105.7	-4.4
Conn	0.019	0.019	0.020	0.020	0.020	0.020	3.0
M_1	19910	19886	19800	19292	19006	18922	-4.9
M_2	1432	1426	1414	1390	1374	1368	-4.4
2SC	181610	181533	180880	174318	171406	170693	-6.0
3SC	9239	9230	9193	8951	8816	8777	-5.0
$^2SC_n (x 10^{-6})$	9.493	9.698	9.879	9.841	9.917	8.885	6.4
$^3SC_n (x 10^{-8})$	5.236	5.393	5.538	5.501	5.578	5.728	9.3
TWC	24103	24061	23889	23278	22981	22867	-5.1

Table 2. *E. coli* protein-protein interaction network descriptors for the original graph (G) and network after removal of 2, 4, 6, 8 and 10 hubs. Percent change in the value of descriptors for 10 hubs removed

Descriptor	G	G-2	G-4	G-6	G-8	G-10	G-10 Change [%]
V	270	268	266	264	262	260	-3.7
E	716	647	585	529	478	435	-39.2
c_v	0.579	0.58	0.58	0.566	0.578	0.589	1.8
l_{vd}	4661	3996	3458	2957	2520	2161	-53.6
X	110.7	109.3	108.4	107.1	105.7	104.0	-6.0
Conn	0.019	0.018	0.016	0.015	0.013	0.012	-34.4
M_1	19910	15924	12928	10318	8236	6682	-66.4
M_2	1432	1294	1170	1058	956	870	-39.2
2SC	9239	7315	5879	4630	3640	2906	-68.5
3SC	181610	125353	91634	62795	42838	29305	-83.8
$^2SC_n (x 10^{-6})$	9.493	7.686	6.318	5.09	4.094	3.345	-64.7
$^3SC_n (x 10^{-8})$	5.236	3.724	2.805	1.981	1.394	0.835	-81.2
TWC	24103	19539	15838	12929	10473	8999	-62.6

behaviour of the measured descriptors when the vertex is removed, the higher is its importance for preserving the global properties of the biological system.

3.3. Applicability of the Descriptors to Different Organisms and Network Types

The applicability of the network-based methodology is not limited to the examples given above. To prove this, we

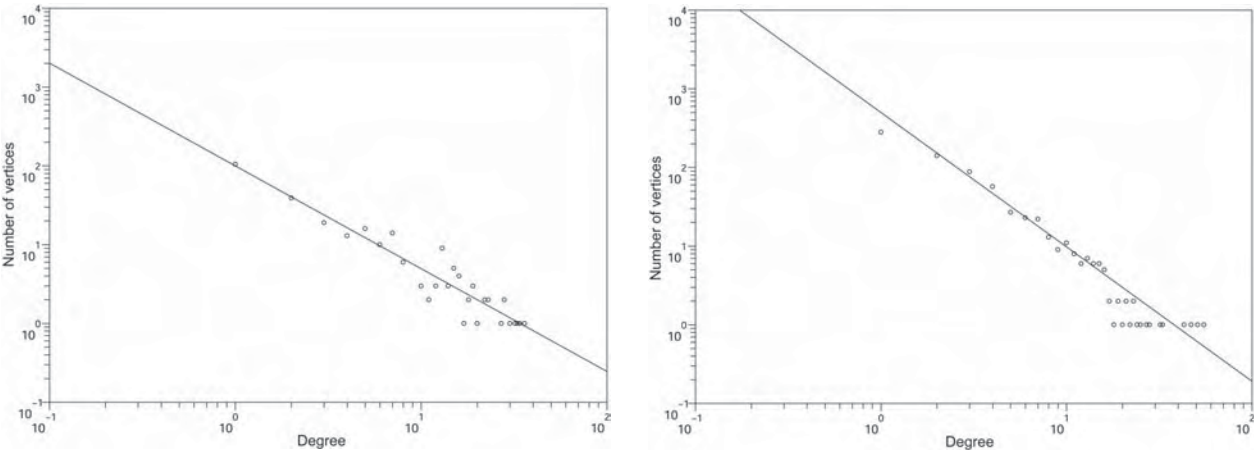


Fig. 4. Relative frequencies of vertex degrees in *E. coli* (left) and *H. pylori* (right) protein-protein interaction network (log-log scale)

Table 3. *E.coli* metabolic network descriptors before and after removing 10 random vertices or 10 hubs

Descriptor	Random			Hubs	
	G	G-10	G-10 Change [%]	G-10	G-10 Change [%]
V	896	886	-1.1	886	-1.1
E	964	936	-2.9	844	-12.4
c_v	0.52	0.519	-0.2	0.552	6.0
l_{vd}	2812	2699	-4.0	2037	-27.5
χ	415.2	407.0	-1.9	411.3	-0.9
Conn	0.0024	0.0023	-0.6	0.0021	-10.4
M_1	7072	6782	-4.1	4688	-33.7
M_2	1928	1872	-2.9	1688	-12.4
2SC	2572	2455	-4.5	1500	-41.6
3SC	11825	11300	-4.4	3739	-68.3
$^2SC_n(x 10^{-6})$	7.175	7.083	-1.2	4.328	-39.6
$^3SC_n(x 10^{-8})$	2.769	2.768	0.0	0.916	-66.9
TWC	11374	10947	-3.7	8295	-27.0

Table 4. *H. pylori* protein-protein interaction network descriptors before and after removing 10 random vertices or 10 hubs

Descriptor	Random			Hubs	
	G	G-10	G-10 Change [%]	G-10	G-10 Change [%]
V	732	722	-1.4	722	-1.4
E	1465	1425	-2.7	1081	-26.2
c_v	0.431	0.436	1.1	0.511	18.7
l_{vd}	7953	7693	-3.2	4850	-39.0
χ	270.6	266.6	-1.4	263.8	-2.5
Conn	0.005	0.005	0.0	0.004	-23.9
M_1	33572	32518	-3.1	15040	-55.2
M_2	2930	2850	-2.7	2162	-26.2
2SC	15352	14865	-3.1	6469	-57.8
3SC	240400	232942	-3.1	56512	-76.4
$^2SC_n(x 10^{-6})$	7.86	7.93	0.9	3.47	-55.9
$^3SC_n(x 10^{-8})$	1.265	1.296	2.3	0.316	-75.0
TWC	61191	59282	-3.1	28486	-53.4

present below vertex degree distribution plots for this type of network in both *E. coli* and other pathogen, *H. pylori* (Fig.4).

As can be observed, both distributions of vertex degree are linear on a log-log scale, similar to the one discussed above for metabolic network.

Also, we analyse the dynamics of *E. coli* metabolic network (Tab. 3) and *H. pylori* protein-protein interaction network (Tab. 4). One can observe that the graph descriptors used above for capturing the dynamics of the network exhibit similar behaviour as in the case of *E. coli* protein-protein interactions.

The results for the analysis using the same set of descriptors applied to different organisms and different types of network show, that the network-based approach is not limited to certain selected biological systems.

4. Conclusions

This article presents an approach to the modelling of biological systems in the form of a complex network. This

methodology has recently gained much interest from researchers from such diverse disciplines as sociology, physics, engineering and life sciences. These applications have also reinvigorated mathematical efforts in the theory of statistical analysis of graphs.

Here, we have observed that static and dynamic properties of biological networks can be captured by a set of descriptors. Moreover, the information obtained from the values and distributions of descriptors can show to what extent local change in parts of the biological network can result in global change of the network properties. We have examined metabolic and protein-protein interaction networks of two microorganisms as examples. In summary, we have shown that, the network-based methodology using a set of informative descriptors is applicable for analysis of different types of biological data and different organisms.

These conclusions can be summarised in brief as follows:

1. Biological systems capturing diverse aspects and organisms can be modelled in a form of complex networks.

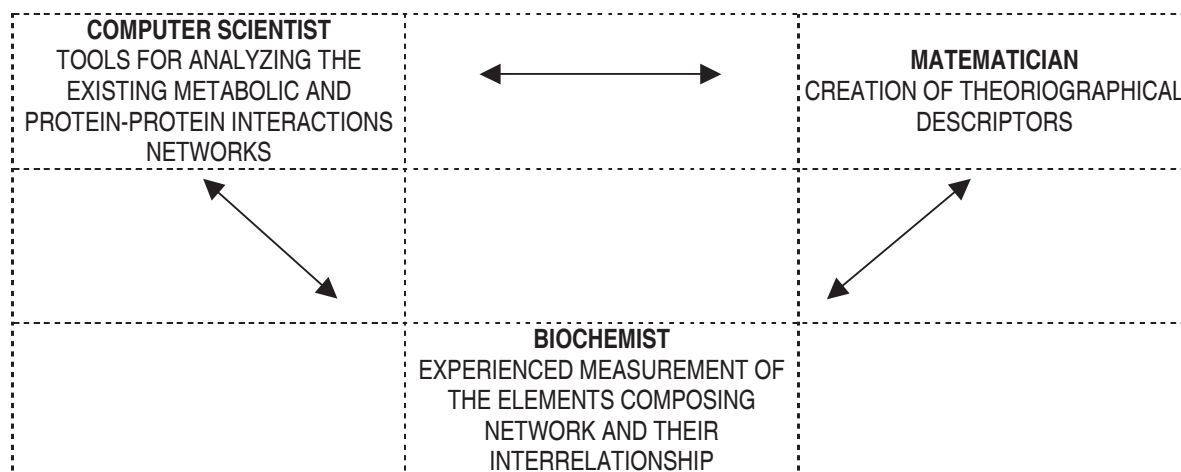
- The static and dynamic properties of the network can be captured by a set of descriptors.
- Descriptors can show the influence of local changes in the biological network on its global properties.

Acknowledgements

The authors would like to thank Prof. Dr Witold Dzwiniel for inspiring discussions and scientific guidance. The research is partially funded by Polish Ministry of Science and Higher Education grants 3 T11F 019 29 and 3 T11F 010 30. Also, T.A. acknowledges financial support from the Foundation for Polish Science.

References

- Albert R., Barabasi A.-L.: Statistical mechanics of complex networks, *Reviews of Modern Physics* 74:47-97, 2002.
- Barabasi A.-L., Albert R.: Emergence of scaling in random networks, *Science*, 286:509-512, 1999.
- Bertz S.H., Herndon W.C.: The similarity of graphs and molecules, in: *Artificial Intelligence Applications in Chemistry*, 169-175, 1986.
- Bonchev D.: *Information Theoretic Indices for Characterization of Chemical Structures*, Chemometrics Series, John Wiley & Sons, 1983.
- Bonchev D., Buck G.A.: *Complexity in Chemistry, Biology, and Ecology*, [in:] *Quantitative Measures of Network Complexity*, Springer, 2004.
- Bonchev D., Trinajstić N.: Chemical information theory, structural aspects, *International Journal of Quantum Chemistry, Symposium* 6:463-480, 1982
- Butland G., Peregrin-Alvarez J.M., Li J., Yang W., Yang X., et al.: Interaction Network Containing Conserved and Essential Protein Complexes in *Escherichia coli*, *Nature*, 433: 531-537, 2005.
- Crucitti P., Latora V., Marchiori M., Rapisarda A.: Efficiency of scale-free networks: Error and attack tolerance, *Physica A*, 320:622-642, 2003.
- Dudek A.Z., Arodz T., Galvez J.: Computational Methods in Developing Quantitative Structure-Activity Relationships (QSAR): A Review, *Combinatorial Chemistry & High Throughput Screening*, 9:213-228, 2006.
- Erdős P., Rényi A.: On Random Graphs I, *Publicationes Mathematicae* 6:290-297, 1959.
- Gordon M., Scantlebury G. R.: Non-random polycondensation: Statistical theory of the substitution effect, *Transactions of the Faraday Society* 60:604-621, 1964
- Gutman I., Trinajstić N.: Graph theory and molecular orbitals. Total π -electron energy of alternant hydrocarbons, *Chemical Physics Letters* 17:535-538, 1972.
- Harary F.: *Graph Theory*. Addison-Wesley, 1969.
- Jeong H., Mason S.P., Barabasi A.-L., Oltvai Z.N.: Lethality and centrality in protein networks, *Nature* 411:41-42, 2001.
- Kier L.B., Hall L.H.: *Molecular Connectivity in Structure-Activity Analysis*, John Wiley & Sons, 1986.
- H. Ma and A.-P. Zeng Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics*. 19: 270-277, 2003.
- Newman M.E.J.: The structure and function of complex networks, *SIAM Review* 45:167-256, 2003.
- Nikolic S., Kovacevic G., Milicevic A., Trinajstić N.: The Zagreb indices 30 years after, *Croatica Chemica Acta* 76:113-124, 2003.
- Platt J.R.: Symbol, *Journal of Chemical Physics* 15:419-420, 1947.
- Randic M.: On characterization of molecular branching, *Journal of the American Chemical Society* 97:6609-6615, 1975.
- Rucker G., Rucker C.: Substructure, subgraph and walk counts as measures of the complexity of graphs and molecules, *Journal of Chemical Information and Computer Sciences* 41:1457-1462, 2001.
- Shannon C.E.: A mathematical theory of communication, *Bell System Technical Journal* 27:379-423, 1948.
- Duncan J. Watts (1999). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton Studies in Complexity, Princeton University Press, 1999.



THE APPROACH FOR COMPUTER ANALYSIS OF RESULTS OBTAINED IN TWO-DIMENSIONAL ELECTROPHORESIS OF SERUM PROTEIN COMPLEXES WITH CONGO RED

KOMPUTEROWA ANALIZA OBRAZU ROZDZIAŁOWEGO KOMPLEKSÓW BIAŁEK SUROWICY KRWI Z CZERWIENIĄ KONGO POWSTAŁEGO W ELEKTROFOREZIE DWUKIERUNKOWEJ

JACEK ŚMIETAŃSKI

*Department of Pharmacokinetics and Physical Pharmacy, Medical College, Jagiellonian University,
Medyczna Str. 9, 30-688 Kraków, Poland; mysmieta@farmacja.cm-uj.krakow.pl*

Abstract. The research of blood serum by the method of double electrophoretical separation, where the second, perpendicular run is modified by adding Congo red, inspires a lot of hope. This paper describes the attempt to create a program which is indispensable to interpret the results.

The received on input scan of the glass plate is brought under preliminary processing which enables to get us rid of noise and useless information. The segmentation which is done in order to separate the objects from the background, was implemented by the method of linear image search and merge. The points are added from the darkest to the brightest ones to the single objects.

The recognition of previously separated objects proceeds stage by stage, thanks to that the areas which are difficult to identify can have more distinctive marks. The value of apurtenance function is calculated by the nearest neighbour method, and for the final classification the majorizing rule is used.

The calculation of percentage of each fraction proceeds by summing the brightness of the points which make the object, after the value of the background was subtracted.

The obtained results of the automatic recognition are promising. However, a greater amount of experimental images is needed for the full evaluating of the usefulness of the program.

Key words: image processing, segmentation, pattern recognition, serum protein, Congo red, electrophoresis

Streszczenie. Badanie składu surowicy krwi metodą podwójnego rozdziału elektroforetycznego, gdzie drugi, prostopadły rozdział modyfikowany jest dodaniem czerwieni Kongo, budzi spore nadzieje. W niniejszej pracy przedstawiam próbę stworzenia oprogramowania niezbędnego do interpretacji wyników. Otrzymany na wejściu skan szklanej płytki poddawany jest wstępnej obróbce, umożliwiającą pozbycie się szumu oraz zbędnych informacji. Segmentacja, mająca na celu wyodrębnienie poszczególnych obiektów od tła, zaimplementowana została metodą liniowego przeglądania obrazu i sklejeń. Punkty do poszczególnych obiektów dodawane są w kolejności od najciemniejszych do najjaśniejszych.

Rozpoznanie uprzednio wyodrębnionych obiektów odbywa się etapowo, dzięki czemu trudniej identyfikowalne plamy mają do dyspozycji większą liczbę charakteryzujących je cech. Wartość funkcji przynależności obliczana jest metodą najbliższego sąsiada, a do ostatecznej klasyfikacji stosujemy regułę majoryzacyjną.

Obliczenie procentowego udziału każdej frakcji następuje poprzez zsumowanie jasności punktów składających się na obiekt, po uprzednim odjęciu wartości tła.

Uzyskane rezultaty automatycznego rozpoznania są obiecujące, jednak do pełnej oceny przydatności programu konieczna jest większa liczba obrazów eksperymentalnych.

Słowa kluczowe: przetwarzanie obrazów, segmentacja, rozpoznawanie obrazów, surowica krwi, czerwień Kongo, elektroforeza

Introduction

The recently elaborated method of research of blood serum by double electrophoretical separation, where the second, perpendicular run is modified by adding Congo red, offers a lot of hope. However, the requirement of clinical research and admission to common use is elaboration of an algorithm which would enable the interpretation of the plate obtained as a result of the research, as well as the quantitative analysis of the results [3].

A computer program which solves the above mentioned problem will be described in this paper. The program is not limited to the mentioned quantitative analysis. There was taken an approach of automatic identification of the proteins visible on the plate [4]. This automatization can definitely shorten time and reduce costs of the research results analysis.

Image preprocessing

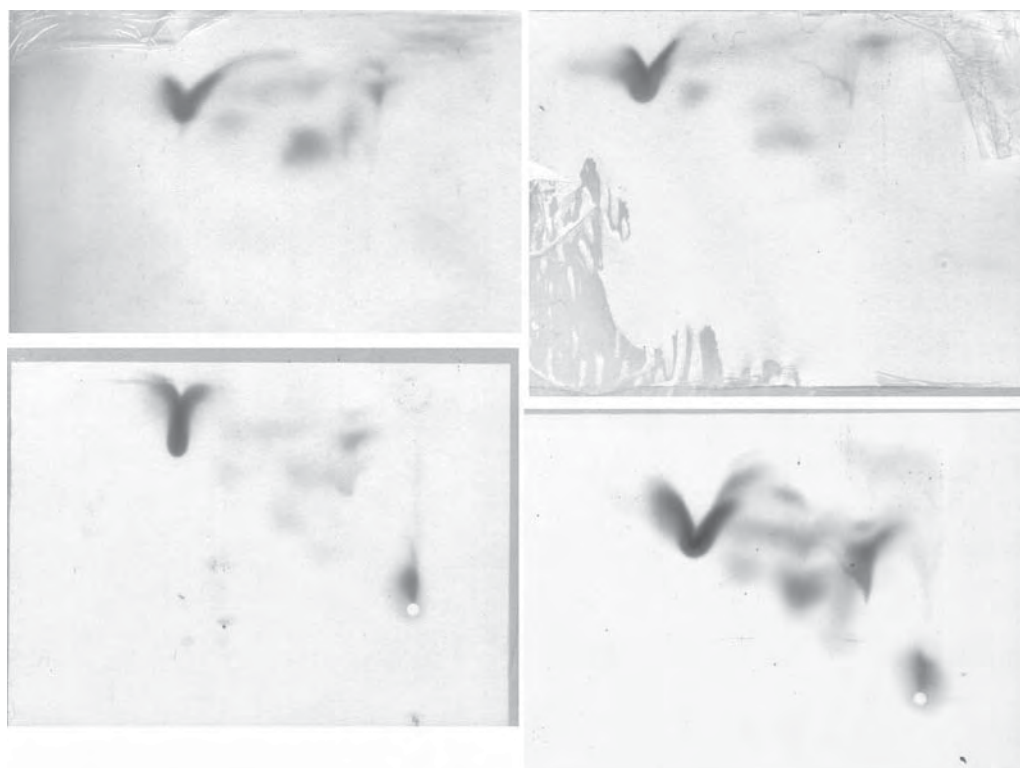


Fig. 1. Sample images.

The direct outcome of the research is a glass plate covered with spots of different size, intensity and form (Fig.1.). The program receives a scan of such a plate on input. In order not to make the question more difficult, mirror reflections or scans of reversed plates are not permitted. The image is brought to the grayscale. It does not cause any important loss of information, for all that it simplifies calculation and shortens the needed time definitely.

To enable the correct image interpretation the disturbances (noise) such as heterogeneity of background or bright points in the area of protein spots have to be removed. For this purpose there was applied an average filter [5] with the following mask:

1	1	1
1	1	1
1	1	1

This filter belongs to the family of convolution filters [2] where the brightness of the point in the filtered image g depends on the nearest neighbourhood of the analogous point in the source image f :

$$g(x,y) = w_1 f(x-1,y-1) + w_2 f(x,y-1) + w_3 f(x+1,y-1) + w_4 f(x-1,y) + w_5 f(x,y) + w_6 f(x+1,y) + w_7 f(x-1,y+1) + w_8 f(x,y+1) + w_9 f(x+1,y+1).$$

In this case the coefficients w_1, \dots, w_9 are formed as follows:

$$w_1 = \dots = w_9 = 1/9.$$

Also using of other average and median filters, as well as segmentation without previous filtration has been considered, but the obtained results were not satisfying.

Segmentation

Segmentation is understood as the separating of objects visible on an image. Indexating is also a part of this process. [1]

It is easiest to carry segmentation of a previously binarized image. On such an image one colour (white) represents the background, and the second one (black) – objects. The binarization is in general carried by the thresholding method, and the key problem is the assignation of the threshold value θ , it means such brightness level which allows to separate the background (brighter points) from the objects (darker points) unequivocally:

$$b(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq \theta \\ 0, & \text{if } f(x,y) < \theta \end{cases}$$

where: $f(x,y)$ – source image

$b(x,y)$ – binarized image.

If the single objects are easy to separate from the background, the task is not complicated; the histogram analysis of such an image allows as a rule to determine the proper level of binarization. Unfortunately, the spots which represent single proteins here, have no distinct edges – the brightness level changes gradually from the center of a spot until it fluently becomes a background. The background, in spite of partial smoothing, does not make a homogeneous texture, either. Often the spots representing the single proteins are so close to one another that the area which separates them, is distinctly darker than objects in the other parts of the image.

An additional problem which is very important and difficult to solve, is the overlapping of some spots (Fig. 2.).

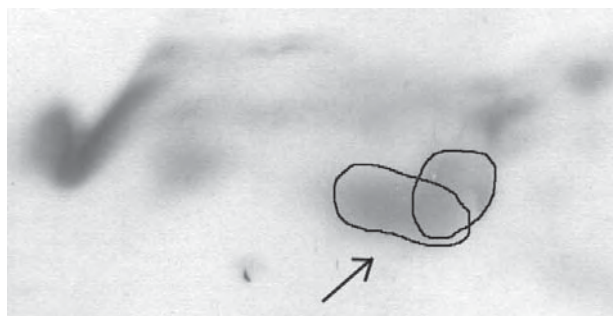


Fig. 2. Example of two proteins overlapped in one spot.

The above mentioned facts cause the unequivocal determination of the binarization level on base of histogram analysis becomes impossible. Also an attempt to find different binarization levels for different fragments of images brings no results.

For that reason the applied algorithm is more complicated – the image is searched by analysing the points from the darkest ones to the average brightness level of the whole image (this level is taken as the background brightness – it guarantees that no object will be omitted; the assigned areas can be a bit too big indeed, but it will not influence the recognition). In each run there is used the algorithm of linear image search and merge [5], and each analysed point is added to the closest contacting object which was previously identified or – if there is no such an object in the nearest neighbourhood – a new object is created. In order to avoid creating a big number of very small objects, the ones which are too small are deleted after each run (the limit largeness has been assigned experimentally), and the points which make them, will be analysed together with the brighter ones in the next run.

After the finished segmentation, before the proper recognition, the objects which surely do not represent any protein (artefacts), are deleted. These objects are too small to represent any protein (the limit size has been assigned experimentally, too) or they border on the edge of the image.

Rules

The preliminary classification can be achieved by analyzing the mutual location of each particular couple of objects. It is known that i.e. β -fraction proteins cannot take a stand to the left from albumin. This and other facts are the knowledge base – a set of certain information. On its base a set of rules, which allow us to eliminate impossible situations, is created. It can happen that some proteins are recognized already on basis of the rules. For the remaining spots the below described algorithm is run.

In order to write the rules precisely a neighbourhood relation is defined.

Two objects X and Y are in the neighbourhood relation ($S(X,Y)>0$) if the sector binding the centres of these objects

does not pass across any other objects. As the centre of an object the most intensive point is assumed.

The value of function S depends on the direction of the relation, or mutual location of researched objects (Fig. 3.).

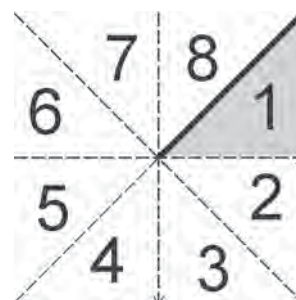


Fig. 3. Direction of neighbourhood relation. If the centre of object X is in the centre of the drawing, the direction of relation $S(X,Y)$ depends on which triangle includes the centre of object Y. Point (0,0) is in the left-top corner of the image. Gray area illustrate belongings of the edges.

To determine the direction of the mutual location of objects, the knowledge of their centres' location and the direction coefficient of line A is needed. The number A is calculated using the following formula:

$$A = \frac{y_i - y_j}{x_i - x_j},$$

where (x_i, y_i) – location of the centre of object X, (x_j, y_j) – location of the centre of object Y.

$$S(X, Y) = \begin{cases} 1, & \text{if } x_i < x_j, \quad y_i > y_j, \quad |A| \leq 1 \\ 2, & \text{if } x_i < x_j, \quad y_i \leq y_j, \quad |A| < 1 \\ 3, & \text{if } x_i < x_j, \quad y_i < y_j, \quad |A| \geq 1 \\ 4, & \text{if } x_i \geq x_j, \quad y_i < y_j, \quad |A| > 1 \\ 5, & \text{if } x_i > x_j, \quad y_i < y_j, \quad |A| \leq 1 \\ 6, & \text{if } x_i > x_j, \quad y_i \geq y_j, \quad |A| < 1 \\ 7, & \text{if } x_i > x_j, \quad y_i > y_j, \quad |A| \geq 1 \\ 8, & \text{if } x_i \leq x_j, \quad y_i > y_j, \quad |A| > 1 \end{cases}$$

Using the relation mentioned above, the rules for all possible threes (protein₁, protein₂, direction) are defined. These rules decide if a given situation $S(\text{protein}_1, \text{protein}_2) = \text{direction}$ is possible or not. By reviewing the rules for each couple of identified objects, impossible situations are excluded.

Recognition

This task consists in classifying the identified objects to adequate classes. The recognition has to be led in situation of lack of any a priori knowledge about appurtenance of the objects to the particular classes. The only information enabled to use is included in the training sequence for which the correct classification is known. [6]

In this case objects are spots on the image, and classes are the names of the proteins. The purpose is the correct assignment of the name of the protein each object represents.

Let:

D – set of objects to recognize

K – classification (equivalence relation) $K \subset D \times D$

I – set of class indexes

It is presumed that K exists, then:

there exist an $A: D \rightarrow I$ such that $\forall d \in D [A(d) = i \Leftrightarrow d \in D_i]$

Function A represents the correct recognition – each object is assigned to the adequate class. In reality, it can happen that the algorithm which finds the function A does not exist at all or it can be characterized by a too high complexity. Therefore the task of recognition is defined as follows:

To construct an algorithm which realizes:

$\hat{A}: D \rightarrow I \cup \{i_0\}$ so that the measure $Q(A, \hat{A})$ was minimum,

Q – measure evaluating the quality of recognition,

$\{i_0\}$ – answer "I do not know".

Function \hat{A} is the composition of three other functions – the solid stage of the recognition process:

$\hat{A} = F \cdot C \cdot B$

where:

$B: D \rightarrow X$ – reception,

X – features space,

$C: X \rightarrow R^L$ – calculating of appurtenance function,

L – number of classes (recognized proteins),

$F: R^L \rightarrow I \cup \{i_0\}$ – making decision.

At the beginning the reception is performed, it means the coordinates of objects to recognize in the previously selected features space are determined. In general it is combined with the selecting of important information and rejecting the redundant. The essence of the recognition process is to calculate the appurtenance function. This function defines how far the recognized element belongs to the given class. The last stage is decision making. Usually the majorizing rule is applied, what means that objects have to be assigned to the class for which the function has the highest value [6].

Considering possibilities to use the characteristics of the particular objects in the features space, distance between the couples of objects, shape, size and intensity of the spots was taken into account. Finally the analysis of size and form was resigned – these parameters would not bring any important knowledge. The training sequence includes information about the brightness and the mutual distance between couples of proteins on correct interpreted images.

The recognition is led by the following scheme:

The objects are analyzed in order of their identification in the segmentation process (it means from the darkest to the brightest one). Thanks to this sequence the biggest, most characteristic proteins are qualified first. Smaller spots, which are identified later, will have more distinctive marks; as a result it will be easier to qualify them correctly.

At the beginning it is supposed that every spot can rep-

resent whichever protein. By reviewing the rules, the impossible situations are excluded. If by this opportunity it becomes sure that the object belongs to the given class, it is identified.

If the class of the analyzed object was not assigned by the method of rules, the object is classified by the minimal distance NN method [6]. The appurtenance function C defines which proteins in the training sequence were nearest to the analyzed object. The decision is made according to the majorizing method, it means the object being recognized, gets the name of the protein for which the function C reaches the highest value:

$$F(x_1, \dots, x_n) = i \Leftrightarrow x_i = \max(x_1, \dots, x_n)$$

Calculating of participation of each fraction

The calculation of proportional participation of each recognized protein fraction is brought to summing up the brightness level of the particular pixels which make the image of the given object. However, a problem here is the necessity to deduct the value of the background of which the brightness level is not clearly defined. As mentioned, the average brightness level was assumed as the background image. So far a petty error has had no great importance for the recognition process, when calculating of concrete numerical values, it is necessary to define the level more precisely. Therefore this proposal needs that the results obtained by the program should be compared with control samples of which the particular proteins location is known from other sources. As yet, such samples do not exist.

Implementation and testing

The program was written in the C++ language as application for the Windows system. The following diagrams illustrate the scheme of data flow in the application, with a particular acknowledgement of the recognition process.

After starting the program a dialog window with the demand to choose an image to analyze appears. For the selected image the automatic analysis option can be chosen or the location and largeness of objects can be fixed manually.

The first option enables a full automatic recognition of identified spots. The segmentation, as well as the recognition are properly visualized (Fig. 6.). During the segmentation there is a possibility to regulate the speed and to stop the process at any moment. This option can help to find an optimal background value.

After completing recognition it is possible to verify the obtained results. In a case of wrong definition of the protein name you can indicate the spot which has been identified incorrectly and give it an adequate name. It is also possible to edit the area of correctly identified objects. Than the table of calculated proportional participation of each fraction appears.

The manual edition option allows to assign the particular proteins to the spots representing them by oneself. The borders of the assigned object are set by moving horizontal

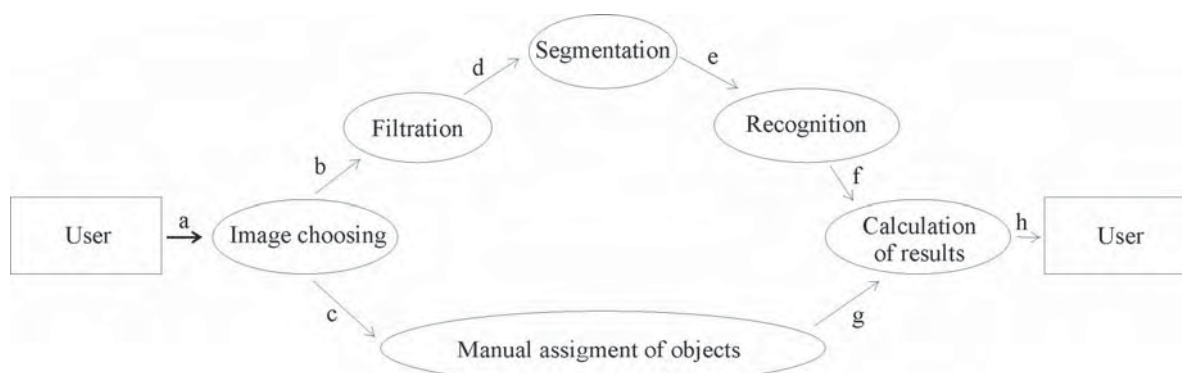


Fig. 4. DFD diagram of the created application: a – list of accesible images; b – source image, choosing the automatic recognition option; c – source image, choosing the manual edition option; d – filtered image; e – map of objects; f, g – recognized objects; h – proportions of single proteins.

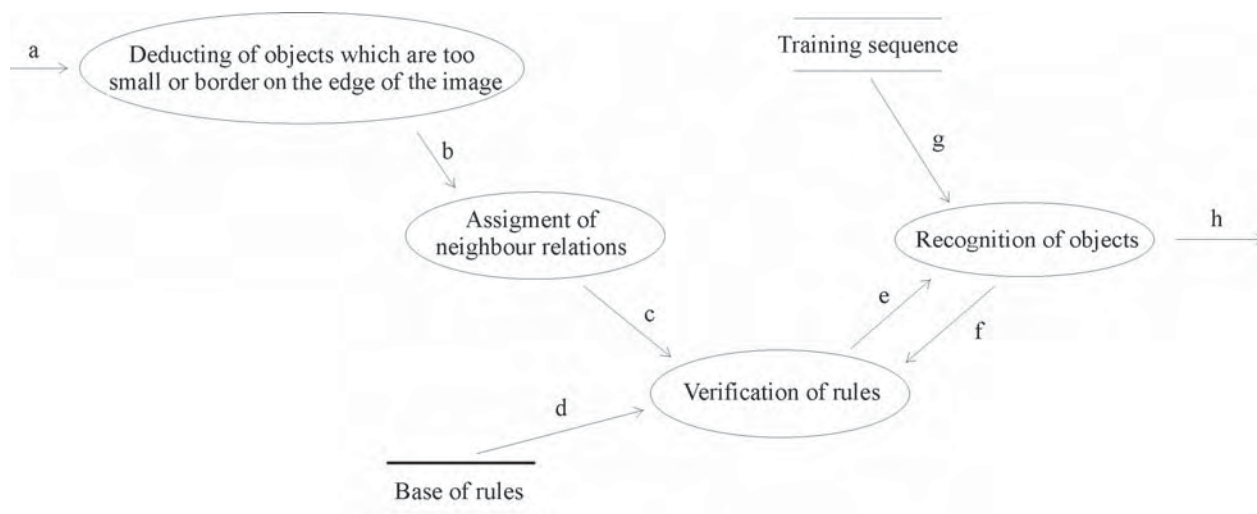


Fig. 5. DFD diagram for the recognition process: a – map of objects; b – map of objects remaining to recognize; c – list of neighbours; d – rules; e – possible belongings of the researched object; f – data of recognized objects, when there is still something to recognize; g – training sequence; h – data of all recognized objects (after the end of the recognition process)

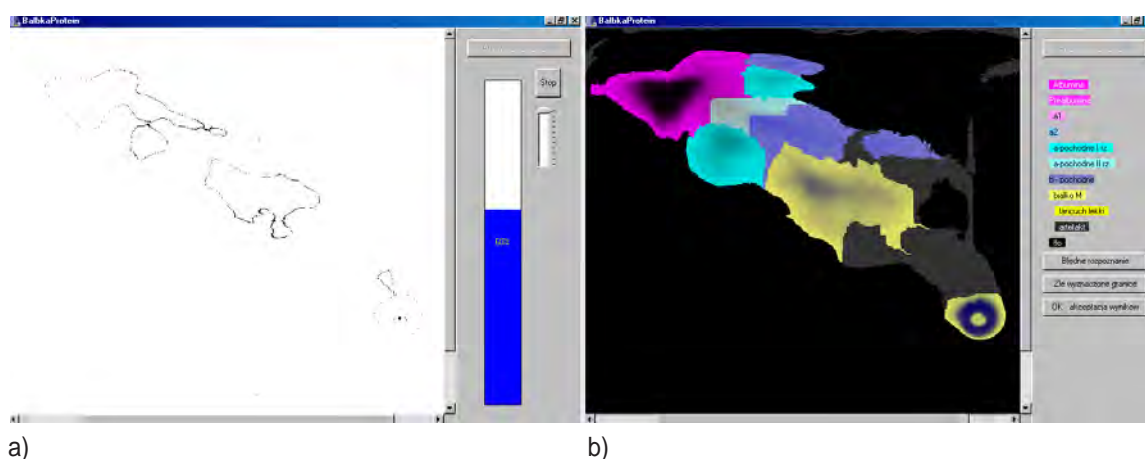


Fig. 6. a) Visualisation of the segmentation stage; b) Recognition result.

and vertical lines. The program helps to fit the object automatically to the marked area. The object can be reduced or expanded by moving the binarization level slide (Fig. 7.).

The results obtained by calculating can be printed or saved to a file.

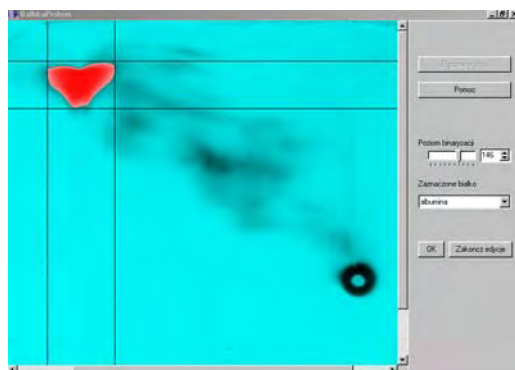


Fig. 7. Manual edition window.

Testing of correct recognition was carried on several images. It is surely insufficient to draw any statistic conclusions about the quality of the recognition process, however the amount of research, which has already been carried out, is insufficient, too. I was able to use only the images which I had got from the person who leads the researches.

The exemplary test result can be seen on the following picture (Fig. 8.). The recognition is mostly correct. A certain inaccuracy is caused by an overlapping of single spots. The erroneous classification of some objects is a result of joining of the object representing albumin with a part of fraction

α_1 and α -derivates. Unfortunately, as already mentioned, the problem of merging spots is still unsolved.

Resume

The created computer program will enable an analysis of the blood serum researches which is carried out by the double electrophoresis method with the addition of liquid crystal ligand (Congo red) before the second, perpendicular run. The automatization of the image interpretation process can entirely eliminate the necessity of participation of an experienced expert in the analysis.

The results of the automatic recognition are promising, however, for a full statistic evaluation there further research, which would enable more complex tests, is still needed.

An important problem to solve is still the overlapping of spots representing different proteins. Also the elaboration of a method which would enable to identify such areas, as well as an algorithm which would divide the overlapping spots efficaciously, seem to be a necessity.

A lengthening of the training sequence can influence the recognition results improvement, too, but a bigger number of images is desirable. It is also necessary to confront the results obtained by the program with those which would be known from another source.

Acknowledgements

I would like to thank Paweł Spólnik from the Medical Biochemistry Institute CM UJ for making the images accessible and helping to interpret them.

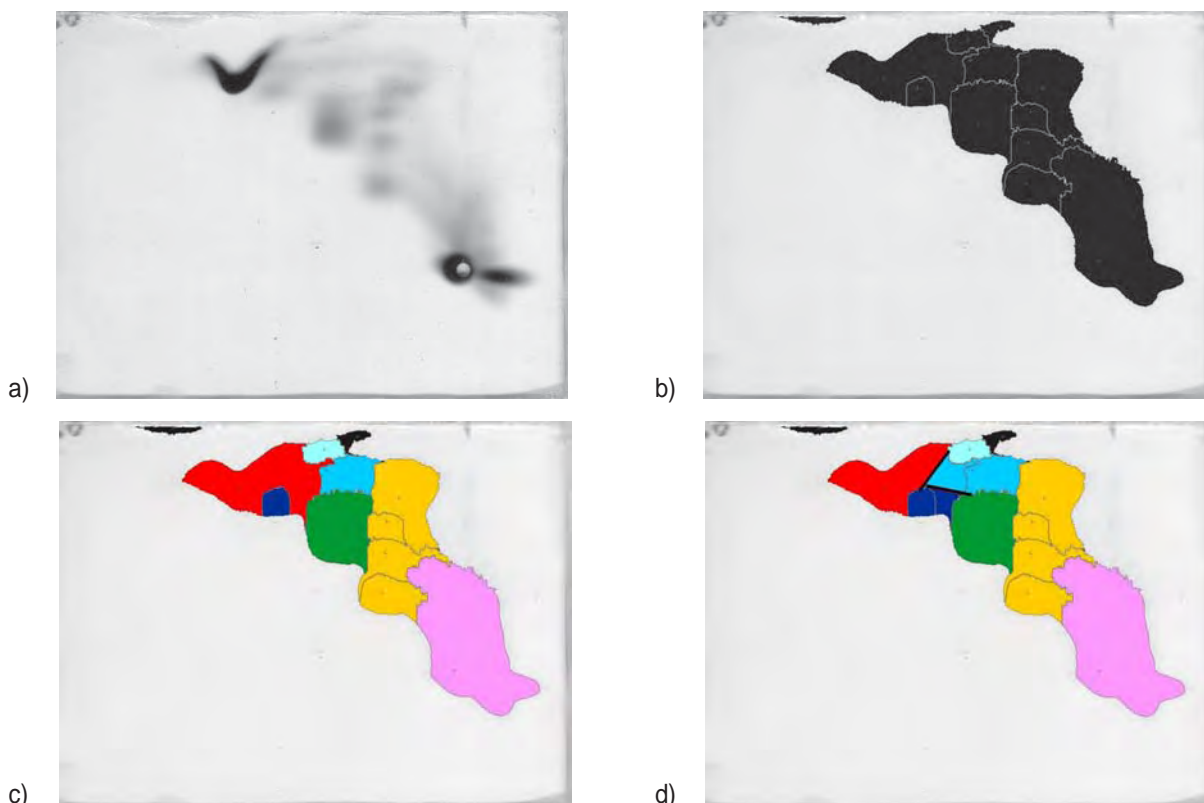
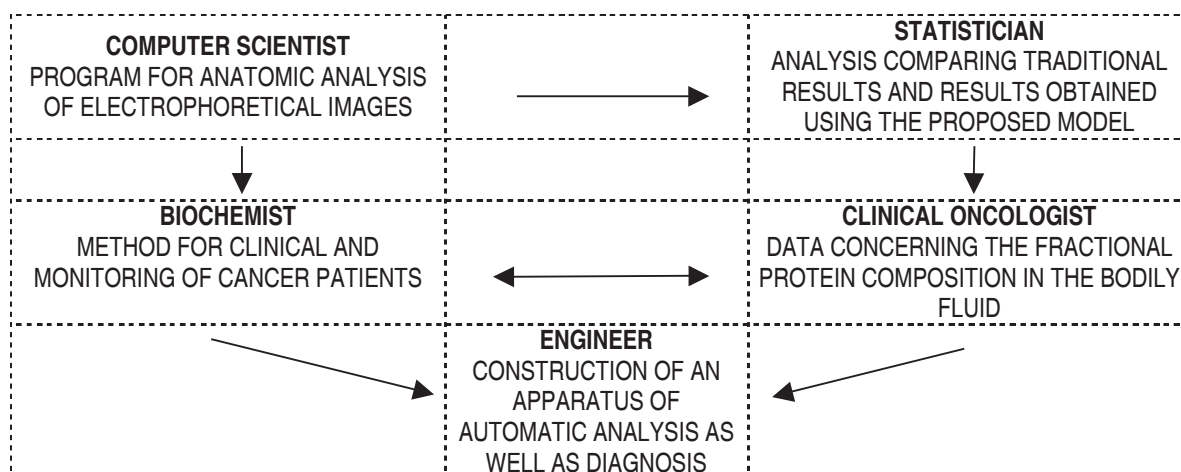


Fig. 8. Example of tested image: a) original image; b) after segmentation; c) recognition made by the program; d) correct recognition

References

1. Gonzalez R., Wintz P.: Digital Image Processing, Addison-Wesley Publishing Company, 1987.
2. Pavlidis T.: Grafika i przetwarzanie obrazów, WNT, Warszawa 1987.
3. Rybarska J., Spólnik P., Drozd A., Konieczny L., Piekarska B., Stopa B., Zemanek G., Król M., Roterman-Konieczna I.: Tworzenie i rozdział kompleksów białek surowicy z barwnikami supramolekularnymi w elektroforezie dwukierunkowej – próba opracowania komputerowej techniki analizy obrazu rozdzielowego, Bio-Algorithms and Med-Systems, Vol. 1, No1/2, pp.133-138, 2005.
4. Śmietański J., Projekt i implementacja systemu rozpoznającego białka zawarte w surowicy krwi, Kraków 2003 (thesis)
5. Tadeusiewicz R.: Systemy wizyjne robotów przemysłowych, WNT, Warszawa 1992.
6. Tadeusiewicz R., Flasiński M.: Rozpoznawanie obrazów, PWN, Warszawa 1991.



AUTOMATED DETECTING SYMPTOMS OF SELECTED GALLBLADDER ILLNESS BASED ON A STATIC ULTRASOUND IMAGES ANALYSIS

AUTOMATYCZNE WYKRYWANIE SYMPTOMÓW SCHORZEŃ PĘCHERZYKA ŻÓŁCIOWEGO OPARTEJ NA STATYCZNEJ ANALIZIE OBRAZÓW ULTRASONOGRAFICZNYCH

SŁAWOMIR BODZIOCH

Institute of Computer Science, AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków

Abstract: Automating a process of diagnosis always comes down to developing algorithms used to analyze the object of such diagnosis and verify the occurrence of symptoms related to a given affliction. Usually the final stage is to make a diagnosis based on the detected symptoms. This last stage can be carried out through either expert systems or a more classic approach, that is with the application of minimal distance methods.

In the case of gallbladder diagnostics, the most important examination still remains image ultrasound diagnostics. This paper discusses the algorithms of a computer analysis of static ultrasound gallbladder images and detecting symptoms of the most frequent illness of this organ.

Key words: ultrasound, ultrasound image, gallbladder, biomedical image analysis, analysis of object shape, histogram analysis, contrast improvement, auto leveling

Streszczenie: Zautomatyzowanie procesu diagnostycznego zawsze sprowadza się do opracowania algorytmów analizujących przedmiot diagnozy oraz weryfikujących obecność symptomów związanych z danym schorzeniem. Ostatnim etapem jest postawienie diagnozy na podstawie wykrytych symptomów. Etap ten zwykle można zrealizować za pomocą systemów ekspertowych bądź podejściem bardziej klasycznym stosując metody minimalno-odległościowe.

W przypadku diagnostyki pęcherzyka żółciowego, najważniejszym badaniem wciąż jest Obrazowa diagnostyka ultrasonograficzna. W niniejszym artykule zostaną omówione algorytmy komputerowej analizy statycznych obrazów USG pęcherzyka żółciowego oraz wykrywania symptomów dla najczęściej występujących schorzeń tego organu.

Słowa kluczowe: USG, obraz ultrasonograficzny, pęcherzyk żółciowy, analiza obrazu biomedycznego, analiza kształtu obiektu, analiza histogramowa, normalizacja kontrastu, auto-poziomowanie

1. Purpose

The purpose of this paper is to present a group of algorithms for detecting symptoms of selected gallbladder illness like cholelithiasis, cancer changes, inflammatory of gallbladder, folds and turns of the gallbladder. Illness symptoms are detected based on statistical analysis of the organ interior (e.g. gallstone detecting, so-called gall sludge, tumors and polyps), analysis of the gallbladder edge and its thickness (e.g. detecting of cancer changes and inflammatory), surroundings analysis (detecting of an acoustic shadow) and geometric analysis of an organ shape (e.g. folded gallbladder). All algorithms presented in this paper concern an analysis of static ultrasound gallbladder images.

2. Introduction

The most frequent pathologies of gallbladder are calculi, folds and turns, inflammations, and cancer changes (polyps and cancer). On ultrasound images, each of the affections is characterized by different elements [1]. Table 1 presents the characteristics of ultrasound images for particular gallbladder illness taken from paper [1].

Ultrasound examination of a gallbladder is a basic test which enables diagnosis of most of the abnormalities listed in Table 1, and therefore it seems vital to ensure support of automated diagnostics for ultrasound scanners.

Table 1. Characteristic elements on ultrasound images of gallbladder for selected diseases and pathological changes

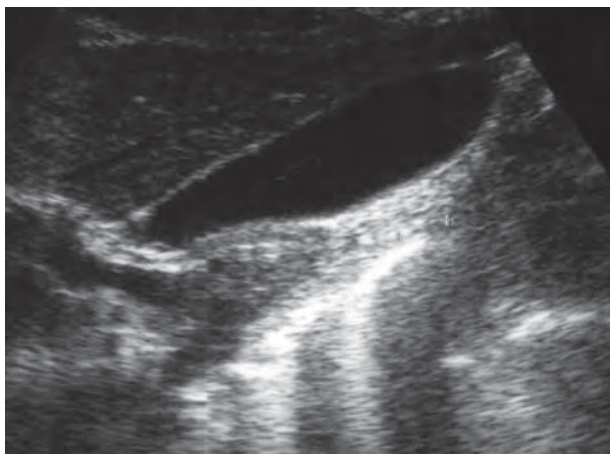
No.	Name of affection	Characteristic features in an image
1	Normal gallbladder	Pear-shaped, 3-4 [cm] in width, 6-11 [cm] in length. Wall thickness 1-2 [mm]. The inside of a normal gallbladder on ultrasound images is uniform dark. Gallbladder volume ranges from 40 to 80 [cm ³].
2	Calcareous calculi	The occurrence of light objects inside a gallbladder which cast a distinct acoustic shadow.
3	Cholesterol calculi	The presence of light objects inside a gallbladder which do not cast acoustic shadow.
4	Polyps	A round growth on the wall of a gallbladder, attached clearly to the wall.
5	Cancer	Intumescence on the wall of a gallbladder which looks like an irregular thickening on the wall of a gallbladder.
6	Hydrocele	Enlargement of a gallbladder beyond regular size, caused by excessive accumulation of bile.
7	Fold/turn	A change to the gallbladder geometry, visible on ultrasound images as folds and their accompanying pseudo-septa seen inside the organ.
8	Inflammation	Inflammation usually accompanies calculi and results from the obstruction of bile ducts. Inflammations are accompanied on ultrasound images by visible local delaminating of the wall gallbladder and the occurrence of strongly echogenic (light areas on ultrasound images) thickened bile, referred to also as sludge.

3. Gallbladder ultrasound image analysis algorithms

The algorithms provided below are based on an assumption that the location of the gallbladder on an ultrasound image is known. The problem of extracting the organ from the background has been left out in this paper; however, there are numerous papers devoted to extracting object outline in digital images, such as [3, 4, 5], including gallbladder outline extraction on ultrasound images [6, 7].

A. Algorithm for verifying a healthy gallbladder

An ultrasound image of a healthy gallbladder is characteristic of the uniform dark inside of the organ, surrounded by light diversified outline – Fig. 1.

**Fig. 1.** An example of a normal gallbladder.

The analysis of ultrasound gallbladder image, the purpose of which is to determine whether a given organ is healthy, can be brought down to statistical analysis of the grey level distribution inside the image. On ultrasound images a gallbladder always appears as the darkest object on the image.

Description of the algorithm for testing whether a gallbladder is a healthy organ:

Certain points are selected randomly inside the organ, from which vertical and horizontal segments are made. A mean value and variation is calculated for each segment. By comparing grey level variations for each section to variance thresholds, it is determined whether a given gallbladder is healthy, or abnormal. Such parameters as segment length, variance thresholds and mean value thresholds must be fitted for particular image resolution. The locations where mean value and standard deviation have high values can point to the occurrence of such objects as pseudo-septa, polyps, or calculi. Such a measurement is illustrated on Fig. 2.

If an average value and a variance value of each histogram segment are lower than assumed thresholds it means that an analyzed gallbladder is normal. Established threshold values for images without improved contrast (e.g. auto leveling) were in the following ranges: average value was in [0, 75], variance – [0, 80]. For images with normalized contrast with using auto-leveling filtration the threshold values were as follows: average value in [0, 50], variance in [0, 120]. All the threshold values were determined empirically based on ultrasound images collected earlier with an average resolution 400 × 500 pixels. Histogram segments were 20-pixels length. However these threshold values should be established for a single ultrasound scanner separately.

B. Algorithm for detecting gallbladder calculi

The detection of calculi requires extension of the analyzed area by the surroundings of the gallbladder. The analysis aimed at either confirmation or elimination of calculi has three stages:

1. Detection of light clusters of points inside an organ, forming a round shape (or light deposits). The presence of a

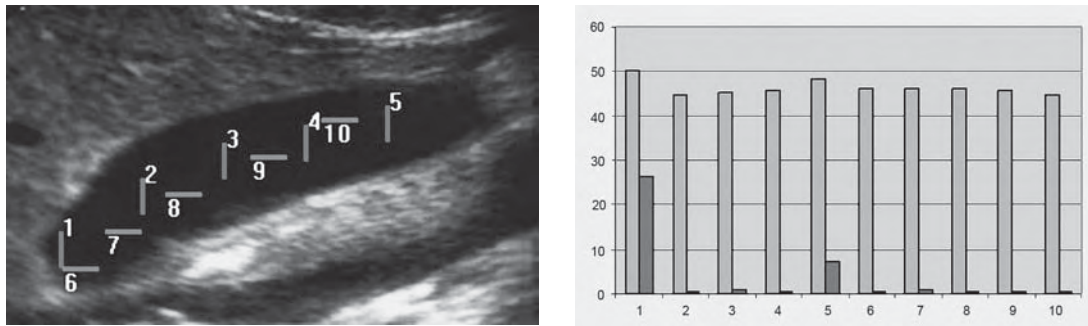


Fig. 2. Sample analysis of the inside of a gallbladder and its results.
(■ – mean value, ■ – variance)

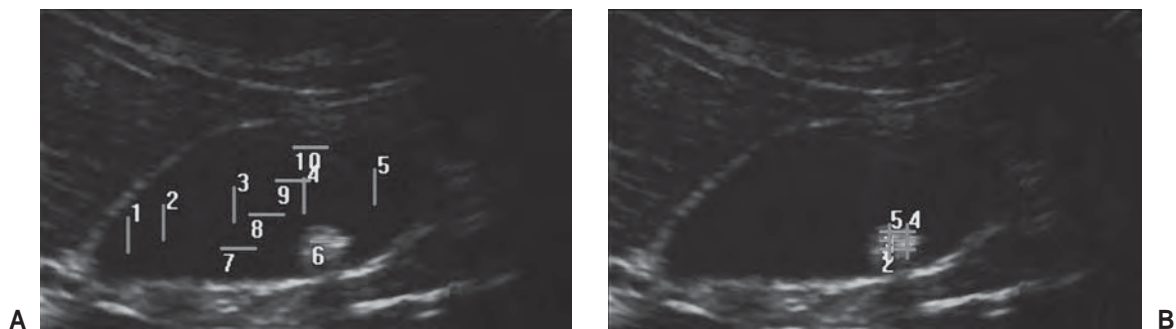


Fig. 3. An example of additional analysis of the inside of a gallbladder in order to confirm occurrence of an object inside the organ: A – initial analysis of the inside of the organ: 6th section is on a cholesterol stone, B – analysis of the surroundings of the section which pointed to occurrence of an object.

number of such objects is possible. The histogram segment method provided for in section A above is sufficient for initial tracking of objects inside a gallbladder; however, for each segment, characteristic of mean values and variance which are higher than threshold values, the verification of surroundings of such a section must be carried out, as shown on Fig. 3.

2. Checking whether a gallbladder is accompanied by either distinctly dark or distinctly light vertical streak, expanding towards the bottom of an image – a test for the occurrence of an acoustic shadow.
3. Checking the correlation between the objects found inside the gallbladder, and acoustic shadow and a rough source of sound. The condition to distinguish between cholesterol calculi and calcareous calculi – only calcified stones cast an acoustic shadow.

Detecting objects inside a gallbladder means gallstones (cholesterol gallstones). (Of course it can also mean polyps.) If detected objects throw a (acoustic) shadow it means calcareous cholelithiasis. Lack of an acoustic shadow means cholesterol cholelithiasis.

C. Algorithm for detecting hydrocele in a gallbladder

Detecting hydrocele in a gallbladder is related to measuring the volume of the organ and testing its geometry in the image plane which presents the largest cross section of the organ. As a normal organ is pear-shaped, hydrocele is characteristic of a rather protruding (ellipsoidal) shape, caused by the occurrence of considerable amount of thin bile. Detecting a disease entity of hydrocele requires:

- measuring the width of an organ,
- measuring the length of an organ,
- geometrical analysis.

Measuring organ width and length on an image must be referred to image resolution, i.e. to the information on the number of pixels in the image per 1[cm]. An approximate volume can be estimated through formula (1) taken from [8].

$$VOLUME = \frac{\pi D_1 \cdot D_2^2}{6} \quad (1)$$

When one of the dimensions (i.e. width, length, or volume) characteristic for a normal gallbladder (see Table 1) is exceeded, this may point to hydrocele.

Geometrical analyses of the gallbladder shape:

1. Geometrical analysis is based on radius-vectors which project from geometrical centre and end on gallbladder walls. The first radius is inclined to horizontal axis at an angle of 0 degrees. Successive r_i radiuses are determined by increasing the inclination angle by certain fixed value. Based on the analysis of variations in the length of r_i , it is possible to verify geometry of the gallbladder shape and determine the longest segment included in the gallbladder and crossing the geometrical centre (the segment which defines the length of the gallbladder) – see Fig. 4.
2. Another geometrical analysis which verifies shape is based on generating measurement segments at right angles to the segment which determine the gallbladder length, with their ends seated at the upper and lower

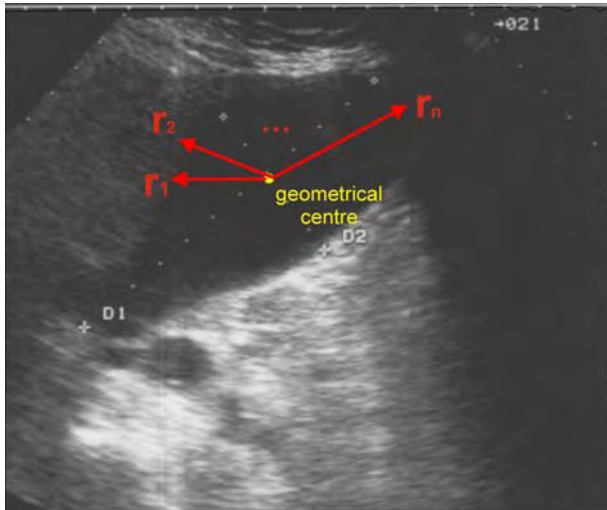


Fig. 4. Geometrical analysis no. 1 of a gallbladder: geometrical centre marked yellow, r_1, r_2, \dots, r_n – radius-vectors projecting from the point which is the geometrical centre and ended at the edge of the organ. Two successive radiuses are inclined to one another at a fixed angle.

edges of the organ. The comparison will cover the segments equidistant from the geometrical centre, but located at the opposite sides of the geometrical centre. If the difference in lengths for each pair of segments is near 0, the shape of the organ is close to elliptical. Otherwise, if the difference is significant (with a fixed mark) for each compared pair of segments, it can be assumed that the shape of the gallbladder is pear-like – Fig. 5.

Testing gallbladder shape with regard to verifying a disease entity of hydrocele.

In the case of hydrocele, a gallbladder loses its pear-like shape, starts to resemble an ellipsoid in shape, which can be verified through geometrical analysis no. 2:

1. For segments $O_i, i=1, 2, \dots, n$ (as in Fig. 5) find segment O_{gc} that crosses the geometrical centre gc ($gc \in O_{gc}$).

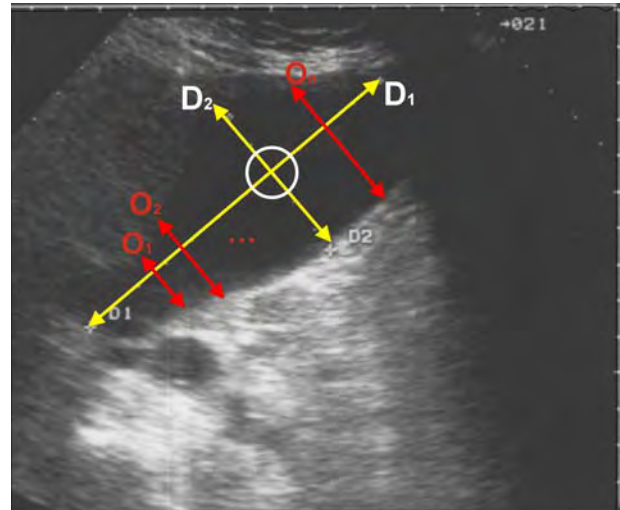
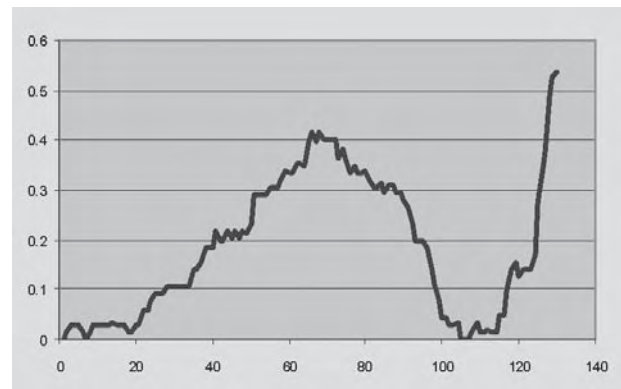


Fig. 5. Geometrical analysis no. 2 of a gallbladder: D_1 – a segment which defines the length of the gallbladder, D_2 – a segment which defines the width of the gallbladder, O_1, O_2, \dots, O_n – segments used to analyze geometry of the organ.

2. Calculate the difference $\delta_k = (O_{ii}/O_{gc}) - (O_{jj}/O_{gc})$ for each pair of segments O_{ii} and O_{jj} equidistant from segment O_{gc} : $\forall ii > 0, \forall jj < n$ and for $k=1, 2, \dots, m$ $\delta_k = (O_{ii}/O_{gc}) - (O_{jj}/O_{gc})$ – dividing the length of segment O_{ii} by O_{gc} ensures independence from the scale of imaging the gallbladder.
3. An analysis of the graph of dependence $\delta_k(k)$ – when a gallbladder loses its original shape and starts to resemble an ellipsoid, a graph of the function $\delta_k(k)$ flattens. For a normal gallbladder, depicted according to the assumptions for cross-section, the graph will be parabolic, with a distinct local extreme for an argument with its value near $m/2$. If the shape of the graph is maintained, but a value of the extreme is below 0.3, the gallbladder can be considered abnormal.



A



B

Fig. 6. An example of applying geometrical analysis no. 2 to a normal gallbladder:

A – ultrasound image of a gallbladder,

B – graph of dependence $\delta_k(k)$ for the gallbladder on image A – the graph features characteristic local extreme, usually located somewhere near the middle of the graph.

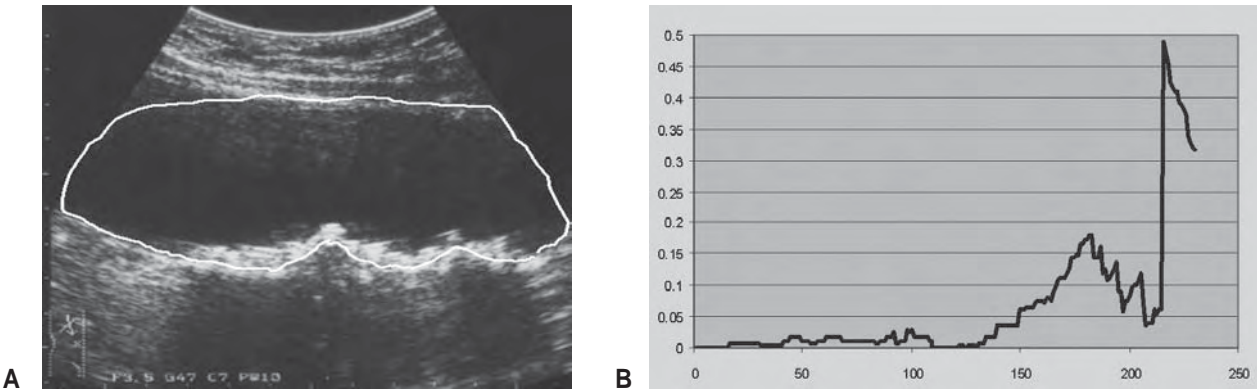


Fig. 7. An example of applying geometrical analysis no. 2 to a gallbladder hydrocele:

A – ultrasound image of a gallbladder,
B – graph of dependence $\delta_k(k)$ for the gallbladder on image A.

Examples of geometrical analysis no. 2 for the verification of a disease unit of hydrocele are shown on Figures 6 and 7.

D. Algorithm for detecting folds and turns of a gallbladder

Detecting a gallbladder fold is related to analyzing the geometry of the organ. When a turn occurs, beside (instead of) geometrical anomalies, an additional line is seen inside the gallbladder, similar to a section wall of the organ. Geo-

metrical analysis is based on the assumption that on an ultrasound image a normal gallbladder is a convex figure, and a gallbladder fold is related to verifying convex in an analyzed image. Additionally, a turn which accompanies a fold is related to gallbladder wall ripple, seen on ultrasound images as a septum inside the organ.

Detecting pseudo-septa inside the gallbladder involves carrying out a number of sections and analyzing a section histogram, as presented in Table 2.

Table 2. Illustrated verification of a turn of a gallbladder on ultrasound image.

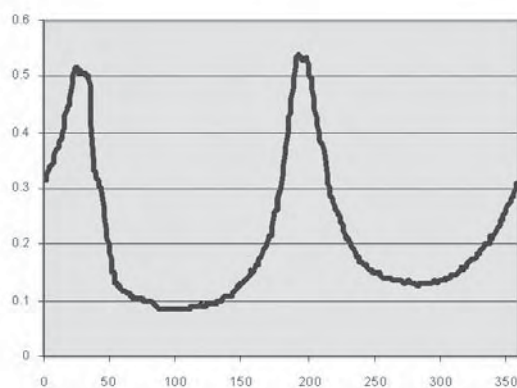
Analyzed image	Section no.	Section histogram
	1.	
	2.	
	3.	

Analysis of shape involves its prior encoding based on the geometrical analyses specified in section C. **Encoding shape:**

1. Based on geometrical analysis no. 1 (Fig. 4), a diameter $d_i = (r_i + r_{i+180})$ is found, for which a sum of $r_i + r_{i+180}$ is maximum: $d_{L_{max}} = \text{MAX}(r_i + r_{i+180})$ $i=1, 2, \dots, 360$.
2. Plotting a dependence $R(i) = r_i/d_i$ $i=1, 2, \dots, 360$, where i stands for an inclination angle of radius r_i to horizontal level, expressed in degrees. (The quotient r_i/d_i is to ensure independence of the scale applied in imaging the organ).
3. Based on geometrical analysis no. 2 (Fig. 5): O_j segments are plotted for $j=1, 2, \dots, d_{L_{max}}$ perpendicular to $d_{L_{max}}$.
4. Plotting a dependence $Q(j) = O_j/d_{L_{max}}$ for $j=1, 2, \dots, d_{L_{max}}$. (The quotient $O_j/d_{L_{max}}$ ensure independence of the scale applied in imaging the organ).
5. A shape of graph $R(i)$ for a normal gallbladder should be close to two parabolas, with two distinct local maximums of similar values. Asymmetry of the graph, a lack of two local maximums, or abrupt changes in function value stand for anatomical abnormalities of the gallbladder.
6. A graph of dependence $Q(j) = O_j/d_{L_{max}}$ for $j=1, 2, \dots, d_{L_{max}}$ for a normal gallbladder should be characteristic of a steady change of values for the growing arguments. If a graph $Q(j)$ features abrupt changes of value $Q(j)$ for successive arguments, this confirms that given gallbladder is folded.



A



B

Geometry deviation of the gallbladder causes a 2-dimensional image of this organ to lose its convex property. Lack of gallbladder convex property causes the $R(i)$ function graph to be not symmetrical (related to Y axis). Moreover on the $Q(j)$ function graph there are some abrupt value changes. Detecting of gallbladder fold/turn is associated with analysis of a value distribution of functions: $R(i)$ and $Q(j)$. Another symptom connected with gallbladder turn is detecting pseudo-septa as it is illustrated with Table 2.

Figures 8 and 9 present two cases of analyzing shape for a normal gallbladder (Fig. 8) and when a gallbladder is folded (Fig. 9).

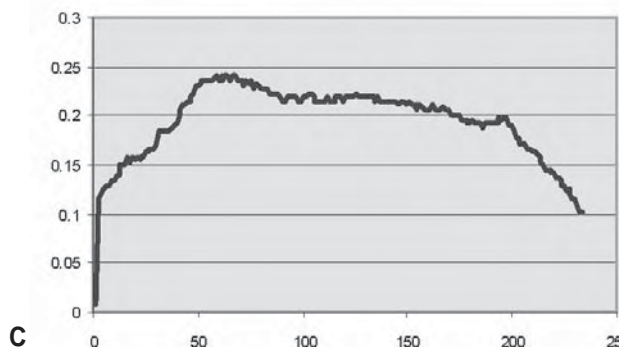
E. Algorithm for detecting cancer changes

When detecting cancer changes, it is necessary to precisely analyze the wall of the organ. To facilitate such analysis, the author of this paper puts forward a transformation of a gallbladder image based on geometrical analysis no. 2 (see C). **Gallbladder wall image transformation:**

1. Generating an image of lower wall **ed1** as follows:
 $\forall i=1, 2, \dots, n$ and $\forall j=1, 2, \dots, |O_i|$ **ed1**[i][j] = $O_i[j]$,
 where n is the number of segments O_i (taken from geometrical analysis no. 2)
2. Generating an image of lower wall **ed2** as follows:
 $\forall i=1, 2, \dots, n$ and $\forall j=|O_{max}|, (|O_{max}|-1), \dots, (|O_{max}|-|O_i|)$
ed2[i][j] = $O_i[j]$, where n is the number of segments O_i , and O_{max} is the length of maximum segment O_i .
3. Binarization of images **ed1** and **ed2** with a dynamically adjusted binarization threshold. Threshold adjustment is to present the wall of the organ and its possible pathological changes.

Detecting of cancer changes is connected with measurement of gallbladder edge thickness applied on the transformed edge image. Measurement allows us to detect tumor changes. Normal gallbladder edge has thickness in a range of 1-2 millimeters. The gallbladder cancer is visible as a wide uniform bulge 2-3 centimeters length. Polyp (Figure 13) is a local tumor on the gallbladder's edge.

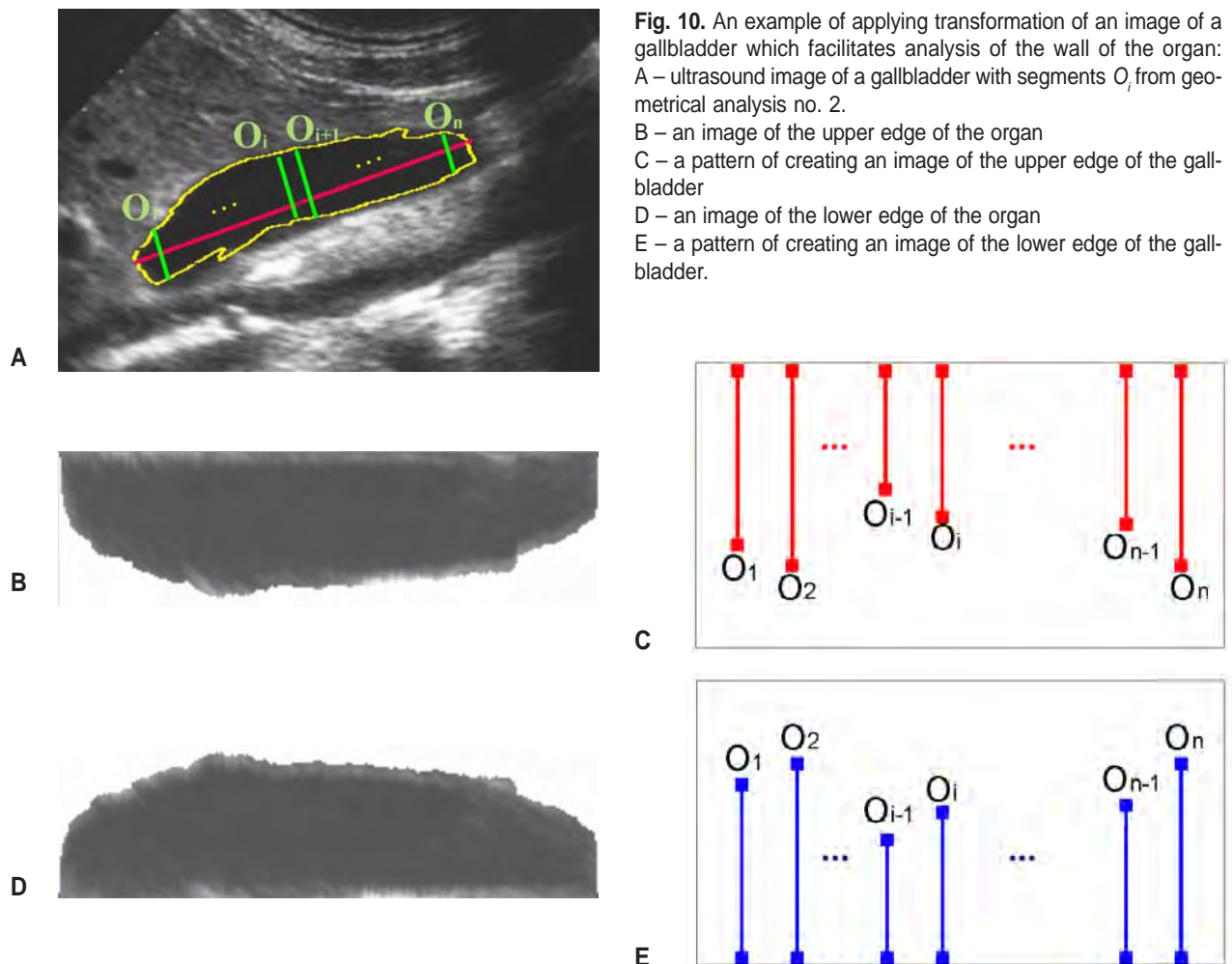
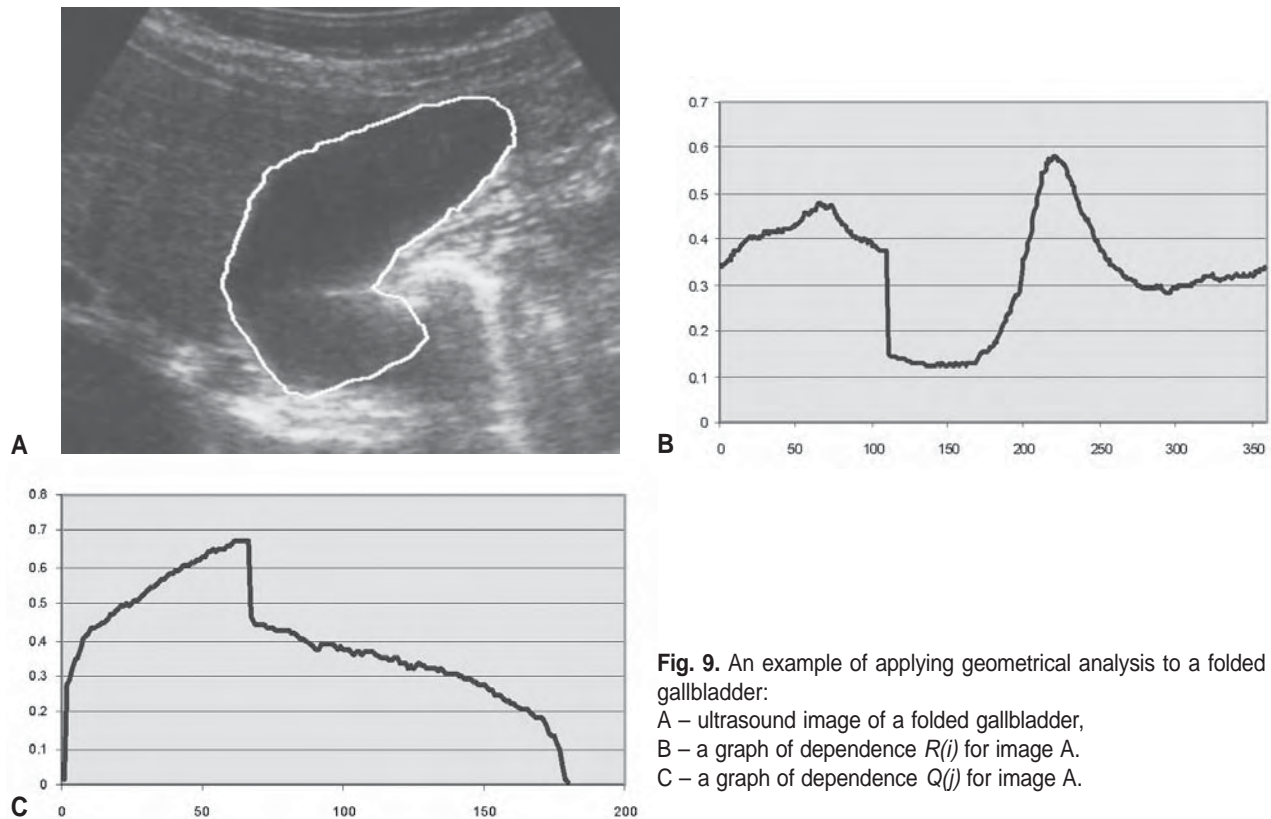
Figure 10 shows both transformations of the wall of a normal gallbladder.

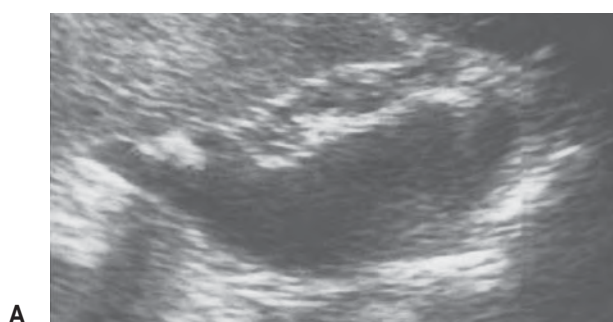


C

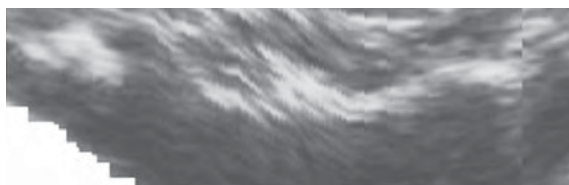
Fig. 8. An example of applying geometrical analysis to a normal gallbladder:

- A – ultrasound image of a normal organ,
 B – a graph of dependence $R(i)$ for image A.
 C – a graph of dependence $Q(j)$ for image A.





A



B

Fig. 11. An example of representing tumor intumescence on the wall of a gallbladder:

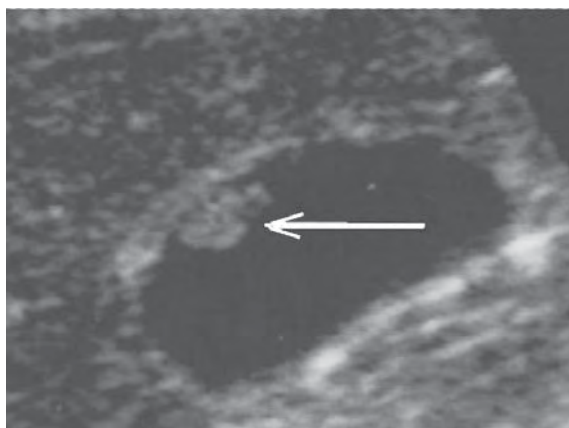
A – an original ultrasound image of the gallbladder

B – a transformed image of the upper edge of the organ

C – a transformed image of the upper edge of the gallbladder after the operation of binarization, which depicts irregularities of the wall.



C



A



B



C



D



E

Fig. 12. An example of detecting a polyp on the wall of a gallbladder.

A – analyzed ultrasound image of a gallbladder with an arrowed polyp,

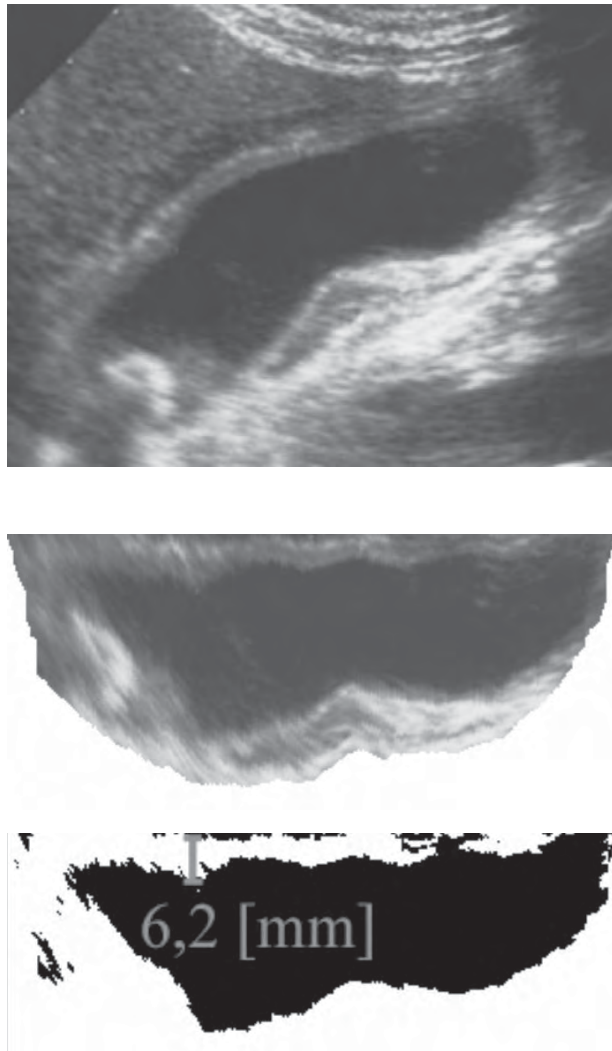
B – graph of thickness of the upper wall a gallbladder,

C – image of the upper edge with a polyp,

D – binarized image of the upper edge with a polyp,

E – image of the lower edge,

F – binarized image of the lower edge.



Figures 11 and 12 in turn depict representation of tumor changes to the wall of a gallbladder for two cases: a polyp and a cancer.

F. Algorithm for detecting inflammation of a gallbladder

Detecting the symptoms of gallbladder inflammation on ultrasound images can actually be brought down to verifying two features: delaminating of the gallbladder wall (and measuring thickness of the wall), and checking for occurrence of thick bile.

To detect delaminating of the wall of a gallbladder, it is suitable to apply an image transformation provided for in *E*-section. An example of its application, which depicts wall thickening with local delaminating, is shown on Fig. 13.

Detecting deposits of thick bile, characteristic of its high echogenicity (light images inside the organ) is actually analogous to the cases of verifying a normal gallbladder. The difference is that a section is not applied, but board samples are taken from squares the size of which depends on the resolution of an analyzed image. Mean value and variance will be calculated for each square. In cases of thick bile, the mean values will somewhat exceed the parallel values for a normal gallbladder, as the sludge is characteristic of its in-

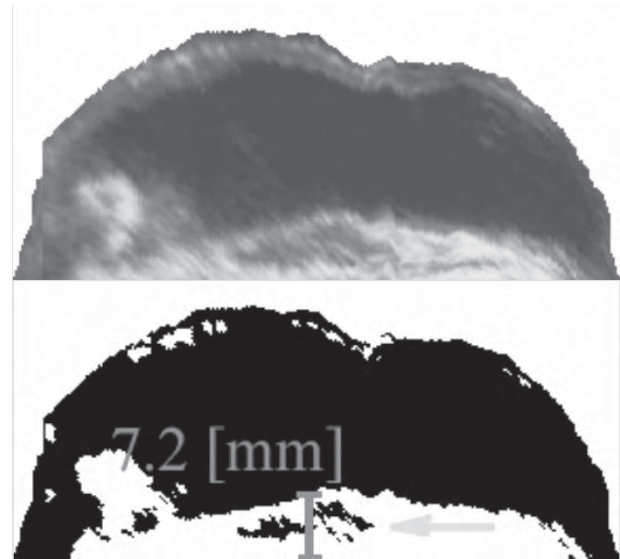


Fig. 13. An example of detecting delaminating and thickening of the wall of a gallbladder:

A – a gallbladder in an inflammatory condition.
B – transformation of an image of a gallbladder depicting the lower edge of the gallbladder,
C – image of the lower edge after binarization with marked thickening and wall delaminating,
D – transformation of an image of a gallbladder depicting the upper edge of the gallbladder,
E – image of the lower edge after binarization with marked thickening of the wall.

creased echogenicity. The same will apply to variance accordingly. (The same method can be applied in detecting stones and cholesterol deposits, characteristic of their high echogenicity depicted as light objects inside an organ).

Detecting of a gallbladder inflammatory state is connected with finding a thickened gallbladder edge (thickness higher than 2 millimeters) with stratification (diversified gallbladder edge). The presence of gall sludge also can point to an inflammation of the gallbladder however a lack of sludge does not exclude this state.

4. Discussion

A verification of the algorithms presented above is given in Table 3. The tests were executed for each illness unit. Since automation of the ultrasound gallbladder diagnosis is not effectively developed yet, so archived results presented in this paper in the Table 3 cannot be compared with other published results. Assessment of their effectiveness is left to the Reader.

The algorithms are based on an assumption that it is known in advance where exactly a gallbladder is located on an ultrasound image. Therefore, separating the organ from the background becomes a key factor for implement-

Table 3. Traceability results of selected gallbladder illness detection based on using the algorithms of static gallbladder image analysis.

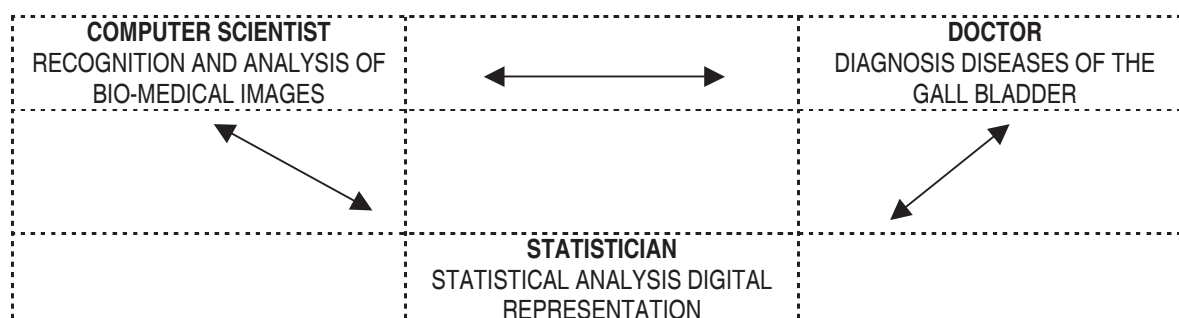
Disease entity	% detectability	number of images used in verification
A normal gallbladder	78% / 88% (an enhanced contrast)	30
Calculi	71% – detecting an object inside the gallbladder 64% – detecting acoustic shadow	40 25
Folds/turns	76% – geometrical analysis 80% – detecting a pseudo-septum	21
Tumor changes	89% – polyps 100% – tumor	9 2
Inflammation	100% – thick bile detected 100% – delaminated wall of the gallbladder	13 2
Gallbladder hydrocele	100% – measuring the gallbladder 66% – geometrical analysis	6

ing the algorithms provided for in this paper into the FirmWare of ultrasound equipment. Moreover, the algorithms are parameterized (e.g. the thresholds of mean values and variances for the method of finding objects inside an organ) and depend on image contrast, the latter being an outcome of the type of ultrasound scanner, the scanner head used, and the systems which transfer ultrasound waves into an image (referred to as presentation types [8]). The remaining parameters must be adjusted with regard to image resolution.

Therefore the algorithms which separate an organ (from both static images and a video sequence) and the problem of normalizing contrast map out the directions of further research on applying algorithms for supporting image diagnostics.

5. Bibliography

1. Dobrinski W., Kramer H. (eds.), Banholzner P., Decking R., Dobrinski W., Fießl H.S., Görndt L., Habscheid W., Hein N. M., Kremer H., Lechner P., Marshall M., Neff M., Schramm T., Schreiber M. A., Weigold B., Wiessner M., Winters I. (co-authors): Diagnostyka ultrasonograficzna, Wydawnictwo Medyczne Urban & Patner, 1st Polish edition, edited by Wiesław Jakubowski, Wrocław 1996.
2. OSIRIS, A Medical Image Manipulation System available at: <http://www.sim.hcuge.ch/osiris/Osiris419.pc.zip>
http://www.sim.hcuge.ch/osiris/01_Osiris_Presentation_EN.htm
3. Szczypiński P., Strumiłło P.: Application of an Active Contour Model for Extraction of Fuzzy and Broken Image Edges, Machine GRAPHICS & VISION, vol.5, no.4, 1996, pp. 579-594.
4. Szczypiński P., Materka A.: Object Tracking and Recognition Using Deformable Grid with Geometrical Templates, International Conference on Signals and Electronic Systems, Ustroń-Poland 2000, pp. 169-174.
5. Szczypiński P., Materka A.: Variable-Flexibility Elastic Model for Digital Image Analysis, XXI KKTOiUE, Kiekrz, 1998. Also in: the Bulletin of the Polish Academy of Sciences, Technical Sciences, Vol. 47, No. 3, 1999, pp. 263-269.
6. Bodzioch S.: Porównanie metod analizy obrazów tekstur USG., in: Automatyka, vol. 8, book 3, pp. 189-205, Uczelniane Wydawnictwo Naukowo-Dydaktyczne, Kraków 2004.
7. Abolmaesumia P., Siropour M.R., Ultrasound Image Segmentation Using an Interacting Multiple Model Probabilistic Data Association Filter, Medical Imaging 2004, Image Processing. Fitzpatrick, J. Michael (ed.); Sonka, Milan. Proceedings of the SPIE, Volume 5370, pp. 484-493 (May 2004).
8. Nowicki A.: Diagnostyka ultradźwiękowa. Podstawy fizyczne ultrasonografii i metod dopplerowskich, Wydawnictwo Medyczne MAKmed, Gdańsk 2000.



GRID SYSTEMS AND THEIR APPLICATIONS TO BIOMEDICAL SCIENCE

SYSTEMY GRIDOWE I ICH ZASTOSOWANIA W NAUKACH BIOMEDYCZNYCH

MACIEJ MALAWSKI*, TOMASZ SZEPIENIEC**, IRENA ROTERMAN-KONIECZNA***

* *Institute of Computer Science AGH, Mickiewicza 30, 30-059 Kraków*

** *ACK Cyfronet AGH, Nawojki 11, 30-950 Kraków*

*** *Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College,
Faculty of Medicine, św. Łazarza Str. 16, 31-530 Kraków, www.bit.cm-uj.krakow.pl*

Abstract: In this paper we briefly introduced the current developments in e-science and the potential the Grid technology can bring to the scientific community in terms of sharing computing power, access to scientific data and supporting collaboration. After high energy physics, which was the pioneering science exploiting the benefits of grid systems, now also the biomedical sciences are beginning to join the Grid community. The important issue is that the adaptation of applications to the grid requires some effort; therefore a close collaboration between computer scientists (grid experts) and domain scientists becomes a crucial factor to success. What makes the collaboration between biomedical and grid computing community especially promising and challenging is the fact that the complexity of the biological systems is so high, that even the largest computing infrastructure such as EGEE project is not powerful enough to solve all the problems of biomedical science. This still requires a lot of interesting scientific work in the development of new bio-algorithms and grid technologies to support them.

Key words: grid systems, distributed computing, large-scale applications, in-silico experiments

Streszczenie: Celem artykułu jest przedstawienie podstawowych zagadnień dotyczących systemów gridowych i ich zastosowań w naukach biologicznych i medycznych. Ideą systemów gridowych jest współdzielenie zasobów pomiędzy rozproszonymi ośrodkami komputerowymi na świecie, w celu efektywniejszego wykorzystania ich mocy obliczeniowej i umożliwienia realizacji zadań wymagających współpracy pomiędzy wieloma grupami naukowców. Aplikacje mogące wykorzystać potencjalne możliwości gridu to takie obliczenia, które dają się wykonać w sposób równoległy, czyli dają się podzielić na większą ilość zadań do wykonania. Systemy gridowe umożliwiają automatyczne przydzielanie zadań do dostępnych zasobów obliczeniowych oraz przesyłanie danych w obrębie gridu. Obecnie największa na świecie infrastruktura gridowa projektu EGEE dysponuje 30000 procesorów i 14 PB danych w 180 ośrodkach w 42 krajach świata. Przykładami biomedycznych aplikacji gridowych są symulacje struktury białek, testowanie leków przeciwko wirusom ptasiej grypy, symulacje przepływu w układzie krwionośnym pacjenta a także analizy statystyczne i epidemiologiczne na dużych klinicznych zbiorach danych.

Słowa kluczowe: systemy gridowe, obliczenia rozproszone, aplikacje wielkiej skali, eksperymenty in-silico

E-science and Grid technology

Scientific computing becomes an increasingly important methodology of conducting modern research, complementing traditional theoretical and experimental activities. Computers can be used to perform large-scale simulations and data analysis, while the Internet allows access to huge amount of scientific data and facilitates collaboration among scientists. Such a trend of increasing utilization of information technology for scientific applications is often described under the term e-science. E-science poses several requirements on computer infrastructure, and we will enumerate the most important ones:

- A huge amount of computer power is required to conduct simulations of natural phenomena, e. g. in climate modeling, geophysics, biology or pharmacy
- Current experiments produce a large amount of data, which need to be stored, made accessible to large collaborations, and processed during the analysis stage; prime examples are experiments in high energy physics producing data in the order of Peta Bytes per year.
- Infrastructure belonging to and administered by single institutions should be shared among others, allowing more optimal resource utilization and collaborative access to shared instruments, e. g. telescopes, satellites or spectrometers.

These requirements lead to the development of Grid technology [Grid], which is regarded as perhaps the most suitable solution for providing the information infrastructure for e-science. Since there is no simple or widely agreed definition of a Grid, we will try to briefly introduce the main ideas associated with this term.

Generally, the most typical Grid system consists of resources, which may include computing processors (CPUs), mass storage media (disk systems, tape libraries), and instruments. These resources are in general heterogeneous and geographically distributed, and are owned and maintained by participating institutions, often large computing centers. This means that they belong to separate administrative domains, with various access mechanisms and policies, including security considerations. It is now becoming clear, that sharing such resources between institutions in a collaborative way is not a trivial task. Grid technology aims at facilitating such collaborative resource sharing, by virtualizing the resources in order to make them appear in a uniform way to the users. Many institutions may form a so-called virtual organization and offer some of their resources for collaborative usage, benefiting from better balanced utilization of CPU power or access to shared data, without compromising local security policies.

There are many additional topics closely related to grid systems, which are however sometimes excluded from the definition. One is the utilization of idle CPU power, e. g. from workstations in student labs, unused during nightly hours, or even when a person is not interacting with a keyboard. Such idle CPU cycles can be used for performing some useful computations, and theoretical computing power of such systems can be very large. Similarly, computations may be performed using peer-to-peer networks, which offer even larger potential computing power, however with

smaller reliability and limited accessibility. Also, it is possible to build Grid-like infrastructures hidden inside a company or institution intranet.

Grid environments implementation

Having defined a grid system, now we would like to describe an example of an implementation of this idea. Probably the most popular and the most useful from the perspective of users is grid maintained by the EGEE project [EGEE]. The initial motivation of building this grid was to build infrastructure for processing and storing data produced in high energy physics experiments using Large Hadron Collider at CERN, but later the infrastructure was opened to other scientific disciplines and evolved into the main scientific computational platform in Europe. Currently the EGEE grid integrates the power of more than 30 000 CPUs and enables a possibility to store about 14 PB (1PB = 1000 000 GB) of data. The infrastructure is supported by about 180 installations distributed in 42 countries worldwide.

The EGEE grid infrastructure is based on a set of software packages that enables all necessary grid services. This software is usually named as *Grid middleware*, because this is software layered between *fabric*, which is hardware and the operating system, and the user's applications. These packages, integrated in the frame of the EGEE Project, are named gLite. The services of this middleware, could be put in two groups (see Fig. 1) *basic services* ensuring all functionality on resources connected to the Grid system like security, user access to the resource, ensuring data transfers and providing information about the resource, and *collective services* that manage work flow and data flow through many resources.

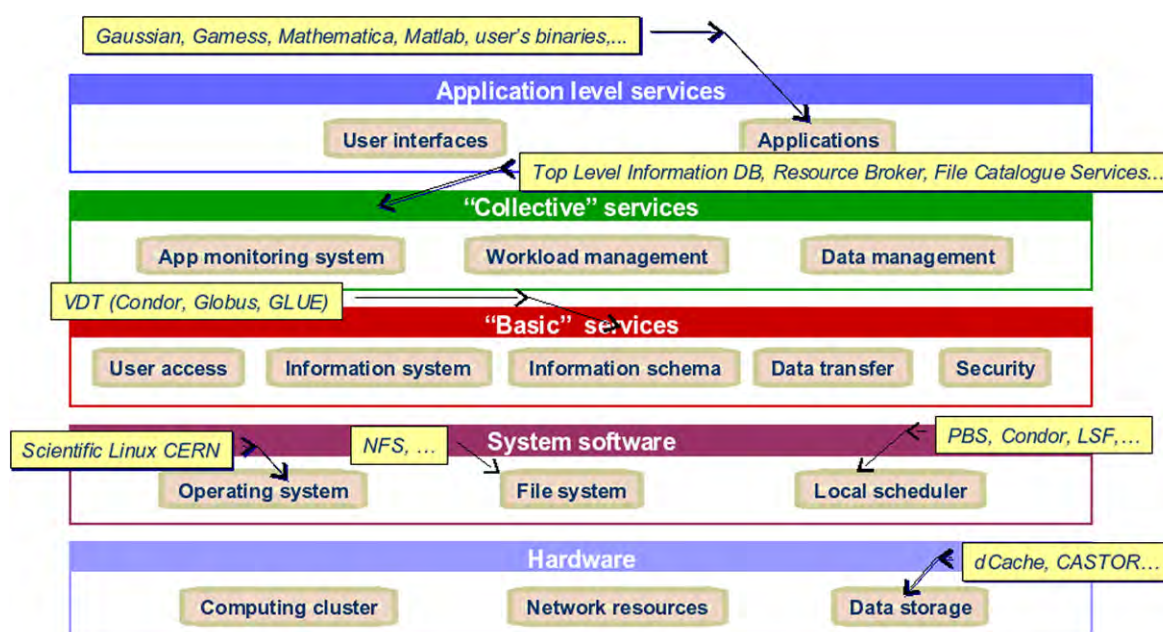


Fig. 1. gLite software stack (source: EGEE Project materials)

Users in Grid

The Grid infrastructure can be treated as a platform for sharing resources, but, of course, rights of using resources are not equal. While the Grid is subject to non-centralized administration, each resource provider could maintain his own policy for using the resources he manages. On the other hand, there are thousands of users that want to ensure resources for their needs. Providing negotiation between single resources providers and every single user would be unfeasible, so the idea of groups of users, named *Virtual Organizations* (VO) was introduced. VO members join the VO to share resources and data for a common goal. VO management, on behalf of all members, negotiates conditions of using the resources with the providers.

In EGEE infrastructure there are several VOs officially supported by the project and much more regional and smaller VOs maintained outside the project, but using the infrastructure. They may support a single experiment (like ATLAS VO – high energy physics project), gather scientists using the same computational package (for example GAMESS VO) or even working in the same scientific domain (like biomed VO or CompChem VO). There are also some initiatives to enable the possibility of grid newcomers to join and test their applications in a Grid environment. Example of such a VO would be VOCE – Virtual Organization for Central Europe.

Grid applications

When discussing Grid systems, we have to answer the question, which kinds of applications are best suitable for running them on the Grid infrastructure and most easily adapted to such an environment. After describing some general remarks we will give examples of applications that are successfully run on current Grids.

General hints for application designers and developers come from the nature of a Grid system, which is a large-scale distributed system, consisting of heterogeneous resources, without a central control and with diversified network performance. Most grid systems offer a batch processing capability, where a single unit of application is a job processing some input files and writing the output to another set of files. A job may consist of multiple processes, possibly distributed between many Grid sites, and communicating using any supported mechanism. Most grid systems do not guarantee the successful execution of a batch job, which may fail unexpectedly. Therefore there is a need to provide some fault tolerance mechanisms on top of existing infrastructures. Another popular approach is based on the service oriented architecture [Foster05], where a grid application may consist of invocations of services offered by others. Such services may provide various computational or data access functionalities, e. g. offering remote access to scientific computing software libraries or databases.

The first class of applications most naturally fitting the Grid environment are the so called embarrassingly parallel problems, which can be solved by running a set of independent copies of an identical program, differing one from

another by some parameter or input data. The processes do not communicate between them at runtime and may be run on many distributed resources in parallel without the need for high-speed network. Such applications may be sometimes naturally fault-tolerant, which means that a failure of some percentage of jobs does not influence the final result. To this class of applications belong e. g. Monte Carlo simulations, where the loss of some jobs has no statistical significance.

Parallel-distributed applications form another class. They consist of multiple processes, which communicate with each other at runtime for synchronization and data exchange purposes. It is well known fact, that the scalability of such applications, which may be very good on a massively parallel machine or a single computing cluster, may decrease substantially when executing on distributed grid resources, which are more loosely coupled. However, such applications may be also ported to a grid, but only when the communication volume between processes is reduced to a minimum.

The third interesting class of applications can be described in terms of a scientific workflow. A workflow is a set of activities (may be jobs, services) executed in a specified order, when subsequent steps depend on the results of previous ones. Such a process may be represented as a graph defining activities and their dependencies. Workflows can be very useful for representing tasks performed repeatedly and may consist of such steps as data access, preprocessing, many simulation runs, and data analysis.

Adaptation of applications to the Grid

There are two main ways of adapting applications to the Grid environment, depending on the selection of batch- or service- oriented architecture.

In the case of batch processing systems [Nabr03], there are a few steps required to transform a standalone program into a job which can be submitted on the Grid. First, the program should be transformed into a single executable or a script, which reads some input files and writes its output also to disk files. The Grid middleware is able to transfer the executable and all input files (often called an “input sandbox”) to the machine on the Grid and execute it. The user can check the status of the job and retrieve output when ready. It is important to consider all dependencies such as libraries or additional data, which need to be either pre-installed on the Grid machines or transferred together with the sandbox. It is also possible to use the files which are stored and registered in a Grid-enabled storage system, so they can be accessed by the computing job. The process of submitting the jobs may be controlled manually by a user by command line or portal interface, or may be automated by tools for managing bulk submission of jobs.

When the Grid is based on service oriented architecture [Foster05], the application should be decomposed into services which can be deployed on the Grid. A service can be defined as software offering some well defined functionality (operations) which are accessible for invocation over the network using a published interface and protocol. Per-

haps most often the Web services standards such as SOAP [SOAP] and WSDL are used to interface the services in Grid systems. These standards allow the implementation of a service in many programming languages, so an existing program or library can be quite easily wrapped into a service. When there are many services installed on Grid sites and available to users, it is possible to orchestrate their operations to form a parallel processing application or a workflow where outputs of previous steps are used as inputs to subsequent services.

Biomedical Application Examples

Starting from parameter study application, we could give the example of drug discovery application. Before the new drug can be applied to standard medical procedures, the complicated experiments must be performed ensuring positive reaction upon its presence in the organism. The large part of these analyzes is performed *in silico*. Simulations of different systems of pharmacological character are able to estimate the positive and/or negative consequences of the presence of potential drug in the human organism. This is why the preliminary analysis *in silico* is important for cost lowering as well as for shortening the time of experimental part of research. Particularly, protein molecules are in focus of the attention of pharmacologists because of their role in functioning of our body. The grid system, assumed to be the result of the project accomplishment, would be used for large scale computing on protein structure in the context of drug design. This kind of research is done in the frame of EUChinaGRID Project [EUChG] by our team.

Another exercise worth mentioning was done using Wisdom Initiative [Wisdom] application adapted to the grid – in April 2006, 300 000 possible drug components was tested against the avian flu virus H5N1 using the EGEE Grid infrastructure. The goal was to find potential compounds that can inhibit the activities of an enzyme on the surface of the influenza virus, the so-called neuraminidase, subtype N1. Using grid in this case helped to mobilize resources that was able to complete in 25 days computations that in case of using a single processor would take almost 19 years.

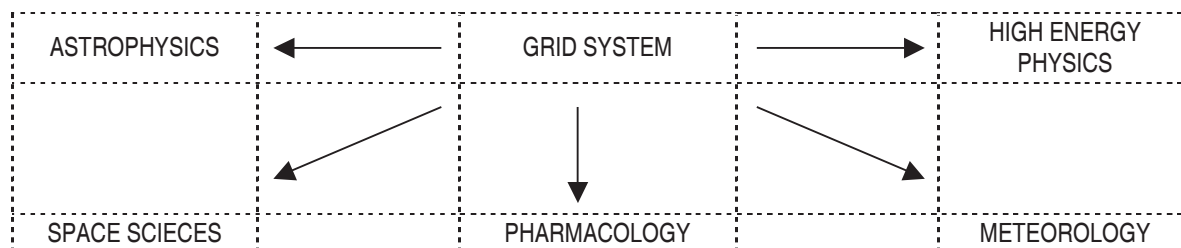
As an example of a medical application which includes parallel programs running on the Grid, may serve the environment for the simulation and visualization of a cardiovascular system [Tirado04]. The system is designed to support a medical doctor in surgery planning. The application con-

sists of a Lattice-Boltzmann blood flow simulation which is a parallelized code running on a high performance cluster on the Grid. It is connected to the visualization system, which allows a user to dynamically interact with the simulation running on remote resources. A doctor can see simulated results of e.g. adding a bypass before the surgery begins. Additional steps like data acquisition from a CT scan, pre-processing and segmentation are also supported.

Many examples of biomedical applications which can be run on a Grid using the workflow approach are developed in the course of myGrid Project [myGrid, Goble03]. The project created a general purpose workflow engine which can orchestrate the services responsible for data access and computation steps of the application on the Grid. Example usages include gene identification and sequence analysis in drug discovery, analysis of clinical records, discovery and analysis in medical imaging systems, and text mining in scientific literature.

References

- [EGEE] EGEE Project: <http://www.eu-egee.org/>
- [Grid] I. Foster, C. Kesselman, S. Tuecke: The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International J. Supercomputer Applications, 15(3), 2001.
- [EUChG] EUChina Project: <http://www.euchinagrid.org/>
- [Tirado04] A. Tirado-Ramos; P.M.A. Slood; A.G. Hoekstra and M. Bubak: An Integrative Approach to High-Performance Biomedical Problem Solving Environments on the Grid, Parallel Computing, (special issue on High-Performance Parallel Bio-computing) vol. 30, nr 9-10 pp. 1037-1055, 2004.
- [mygrid] myGrid Project: <http://www.mygrid.org.uk>
- [gLite] gLite middleware: <http://www.glite.org/>
- [Goble03] C.A. Goble, S. Pettifer, R. Stevens and C. Greenhalgh: Knowledge Integration: In silico Experiments in Bioinformatics in The Grid: Blueprint for a New Computing Infrastructure Second Edition eds. Ian Foster and Carl Kesselman, 2003, Morgan Kaufman, November 2003.
- [Foster05] I. Foster. Service-Oriented Science. Science, vol. 308, May 6, 2005..
- [SOAP] SOAP Version 1.2, W3C Recommendation 24 June 2003, <http://www.w3.org/TR/soap>
- [Nabr03] J. Nabrzyski, J.M. Schopf, J. Weglarz (Eds). Grid Resource Management, Kluwer Publishing, Fall 2003.
- [Wisdom] WISDOM Initiative homepage: <http://wisdom.eu-egee.fr/avianflu/>



STRESS STATE ANALYSIS IN THE FEMUR AFTER THE IMPLANTATION OF THE ANATOMICAL CENTEGA TYPE STEM

ANALIZA STANU NAPRĘŻENIA W KOŚCI UDOWEJ PO IMPLANTACJI ANATOMICZNEGO TRZPIENIA TYPU CENTEGA

WOJCIECH WIĘCKOWSKI

*Department of Biomedical Engineering, Częstochowa University of Technology,
Armii Krajowej Str. 21, 42-200 Częstochowa, wieckowski@iop.pcz.czyst.pl*

Abstract. The relationship between the stiffness of the endoprosthesis stem and the bone stiffness, depending substantially on the geometrical and material features of the implant, influences significantly the long-lasting bio-functionality of the artificial hip joint. In this paper the results of the numerical FEM calculations of this stresses and strain energy density distribution in the bone tissue of the model of a proper femur and a femur after implantation of the modern anatomical stem of the Centega cement endoprosthesis have been presented. The results obtained from the calculations show significant influence of the implantation of the endoprosthesis stem on the change of loading transfer method in the bone region. The observed phenomenon is a main factor determining the adaptive processes occurring in the bone.

Key words: Centega endoprosthesis, hip joint alloplasty, FEM simulation.

Streszczenie. Decydujący wpływ na długotrwałą biofunkcjonalność sztucznego stawu biodrowego ma relacja sztywności trzpienia endoprotezy do sztywności kości, która w znacznej mierze zależy od cech geometrycznych i materiałowych implantu. W pracy przedstawiono wyniki obliczeń numerycznych, przy zastosowaniu MES, rozkładu naprężeń oraz gęstości energii odkształcenia w tkance kostnej w modelu prawidłowej kości udowej oraz w modelu kości udowej po implantacji nowoczesnego trzpienia anatomicznej endoprotezy cementowej Centega. Uzyskane wyniki obliczeń wykazują znaczący wpływ, implantacji trzpienia endoprotezy, na zmianę sposobu przenoszenia obciążeń w obszarze kości, co jest głównym czynnikiem decydującym o przebiegu procesów adaptacyjnych zachodzących w kości.

Słowa kluczowe: endoproteza Centega, protezoplastyka stawu biodrowego, symulacja MES.

1. Introduction

In advanced degenerative changes of hip joint cemented or cementless alloplasty of the hip joint (total or partial) seems to be the only method allowing to restore the functions of the damaged joint [3,6]. The number of realized hip alloplasties grows continuously, simultaneously the number of reimplantations caused by aseptic loosening of the implants increases. The construction of the endoprosthesis is one of the leading factors of the success of hip joint alloplasty. The constructional solutions of the new endoprosthesis types made of modern biomaterials significantly improved the quality of operations however the problems of implant stability remains very current [4,10,18].

Proper selection of the implant dimensions to the patient's femur has essential influence on the result of the hip joint endoprosthesis implantation. It mainly depends on the geometrical adjustment of the shape and dimensions of the endoprosthesis stem to the profile of marrow cavity of the femur. Anatomical cement endoprosthesis Centega (fig. 1) made by AESCULAP is an example of the construction of

the endoprosthesis in which it was tried to better imitate the shape of the marrow cavity of the bone as well as spatial positioning of the head of the implant.



Fig.1. The anatomic endoprosthesis of the hip joint Centega made by Aesculap [2]

The long term analytic examination has resulted in the construction of the anatomical shape of the femoral unit of implant. The distinctive forward bend of the endoprosthesis stem allows the correct imitation of a patient's femur anatomy, whereas the decrease of the neck angle results in enlargement of the movement range of the operated limb. The use of the anatomical endoprosthesis secures the optimum thickness of the cement sheath around the implant (fig. 2).



Fig.2. Radiograph of the hip joint after implantation of the Centega stem [7]

An appropriate matching of the endoprosthesis stiffness to the surrounding bone realized by the selection of suitable stem shape as well as the choice of material of required mechanical properties, influences significantly the manner of loadings transmission from implant to the bone, and therefore determines its long-lasting stability.

Clinical evaluation of the implanted hip joint as well as thorough experimental and numeric examination, deliver essential information regarding the usefulness of a given type of endoprosthesis. They empower the optimum selection of an implant type, aid the forecasting of the results of treatment as well as processes of rehabilitation.

The few results of clinical examinations, considering the comparatively short service life of the Centega type endoprosthesis, forbid precise assessment of the usefulness of that type of stem construction due to its long-lasting bio-functionality. Therefore results of the investigations based on the numeric models may turn out to be helpful, as they allow to predict the effectiveness of the conducted implantation operation [6,14,15].

2. Methodology

The numerical calculations of stress distribution were conducted in the three-dimensional model of the proper femur

as well as after implantation of the stem of cemented anatomical Centega endoprosthesis (fig. 3). The discreet model of femur was worked out on the basis of the data published by *Istituti Ortopedii Rizzoli from Bologna*.

A geometrical model of Centega endoprosthesis stem (L11 - NK 741 K) has been worked out on the basis of the data obtained from the measurements conducted on the coordinating measuring machine. Twenty nodal Solid type units have been used to build the 3D model.

Construction of the "bone - cement - implant" model has required the performing of the simulation of the resection of femur head by removing from the block model the areas occupied by neck and head of femur, specifying mutual position of the endoprosthesis and the femur and isolating around the endoprosthesis surface the layer of cement providing the junction of the implant with the surrounding bone. In the model complete bounding of cement with the implant as well as the cement with the bone all-over the adhesion surface of the adequate areas has been assumed.

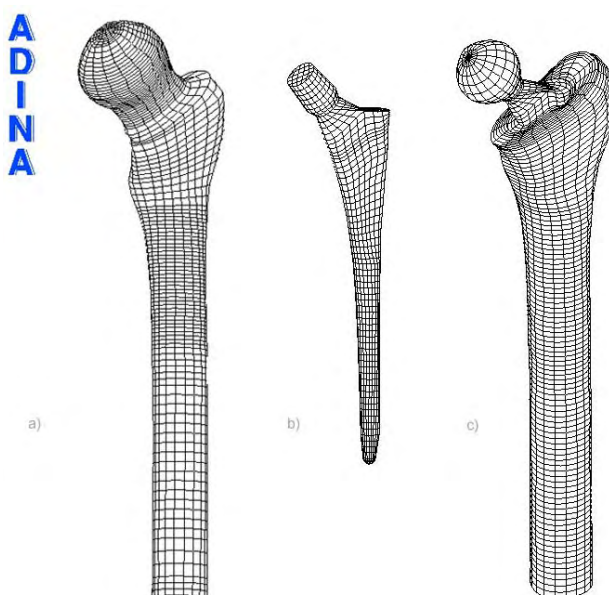


Fig.3. Geometrical models: a) proper femur, b) Centega stem, c) bone - cement - implant system

The basis of the proper construction of the numerical model of implanted joint is the acquainting of type and value of forces active in the surroundings of the hip joint. The adopted loading model should correspond to the actions as well as their phases, which are most unfavourable from a point of view of the implant life. In the discussed system the simplified but commonly used Pauwels's model of hip joint loading, corresponding to standing on one leg has been used: where the external forces applied to the endoprosthesis head $P1 = 2.47 BW$ and to the greater trochanter of the femur $P2 = 1.55 BW$ (fig. 4) [5,6].

A material model of the bone - cement - implant system has been worked out on the basis of available in the literature results of the experimental investigations [8,9,17]. The area of the femur has been modeled as an elastic isotropic

$BW = 600$ [N]
 $P_1 = 2,47 \times BW$
 $P_1 = 1482$ [N]
 $P_{1x} = 1430$ [N]
 $P_{1y} = -257$ [N]
 $P_{1z} = 266$ [N]

$P_2 = 1,55 \times BW$
 $P_2 = 930$ [N]
 $P_{2x} = -871$ [N]
 $P_{2y} = 219$ [N]
 $P_{2z} = -227$ [N]

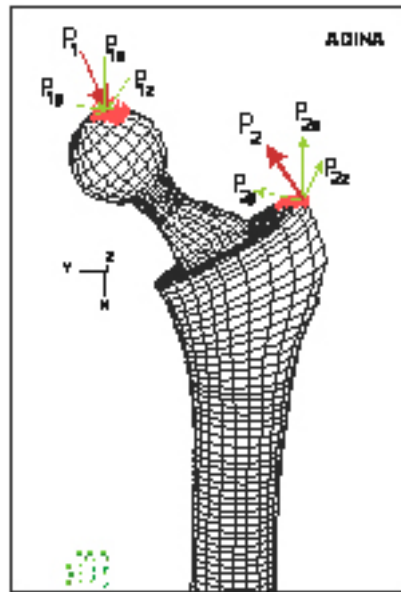


Fig.4. Model of the loading of the system

material, regarding the inhomogeneity of osseous tissue in which two areas occupied by cortical bone and spongy bone were distinguished. The CoCrMo alloy (ISODUR® F) has been taken to the calculations for a stem material. The strength parameters of the stem, osseous cement as well as the bone materials taken into calculations have been presented in table 1 [12,13].

Tab.1. Strength parameters of the individual elements of the bone - cement - endoprosthesis system.

	CoCrMo	PMMA	cancellous bone	cortical bone
Young modulus [Pa]	$2,0 \times 10^{11}$	$2,3 \times 10^9$	$1,0 \times 10^9$	$1,7 \times 10^{10}$
Poisson ratio ν [-]	0,3	0,36	0,4	0,36

In order to provide stability for the examined bone - cement-endoprosthesis system the model has been rigidly fixed in a further part of femur, by removing of all the degree of freedom from the nodes lying in the fixing plane.

Calculations have been conducted using the commercial ADINA software [1].

3. Results

The following variables have been taken to the comparative analysis of the results: the main stresses - $\sigma_1, \sigma_2, \sigma_3$, normal tension components σ_x towards x axis, Huber- Mises reduced stress - σ_{rz} and the strain energy density - U . The evaluation was conducted on the basis of the changes of values of the above-mentioned variables in the area of the osseous tissue surrounding the implant in relation to the state existing before implantation of the endoprosthesis stem (fig. 5).

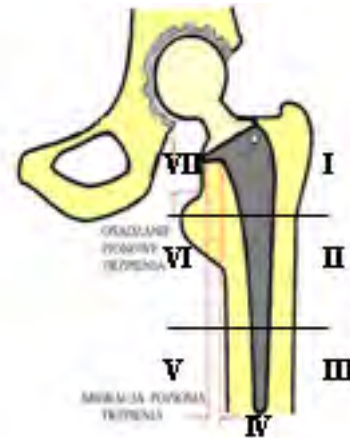


Fig.5. The distribution of Gruen's zones the taken to the evaluation of the numerical calculations results of stress distribution in the model of proper femur as well as in the bone - cement - implant system [21]

Figure 6a presents the reduced stress distribution σ_{rz} in the cortical panniculus for the model of proper femur. The largest values of the reduced stresses $\sigma_{\text{rz}} = 25 \div 46$ [MPa] occur along the whole inner and lateral side of diaphysis.

In the model of femur after Centega type stem implantation (fig. 6b) the largest reduced stresses in the cortical panniculus are respectively $\sigma_{\text{rz}} = 22 \div 31$ [MPa] and similarly to the proper bone model they occur along internal and external side of the diaphysis.

In comparison with the results received from the analysis of the model of the bone before implantation, the largest differences can be observed in the proximal fragment of femur. Maximal reduced stresses in the zone VI comprise 60%, and in zone VII 30% of the stresses values occurring in the proper bone model. In the diaphysis below the implant end these values decreased about 10%.

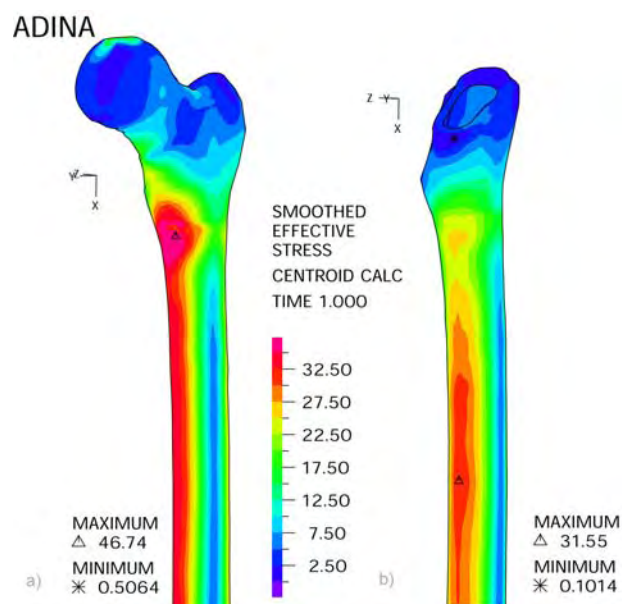


Fig.6. Reduced stress distribution σ_{rz} [MPa] in the region of cortical bone: a) in proper bone model, b) in the model of bone - cement - Centega endoprosthesis

Figure 7a presents the distribution of the normal tensions components s_x towards the x axis - in the area of cortical tissue of the proper femur. The extreme values of s_x stress occur on the surface of the femur in region II and VI of the Gruen's zone and amount respectively $s_{x \max} = 34$ [MPa] and $s_{x \min} = -43$ [MPa]. Because of the assumed boundary condition as well as loading model, compressive stresses occur on the inner side of the femur, however tensile stresses are on the external side. Moreover the assumed loading conditions cause the twisting of the femoral diaphysis, which results in the non-orthogonal arrangement of the isolines of normal stresses in relation to the coordinate system.

In the model of the femur after implantation of Centega type stem (fig. 7b) the extreme values of stresses s_x occur at both the internal as well as exterior side of the femur diaphysis below the end of the implant and amount to $s_{x \min} = -31$ [MPa] and $s_{x \max} = 30$ [MPa]. Compressing stresses in the proximal area of the bone as well as values of tensile stresses in the region of cortical tissue situated outside the system in region II and III of the Gruen's zone decreased to the largest extent.

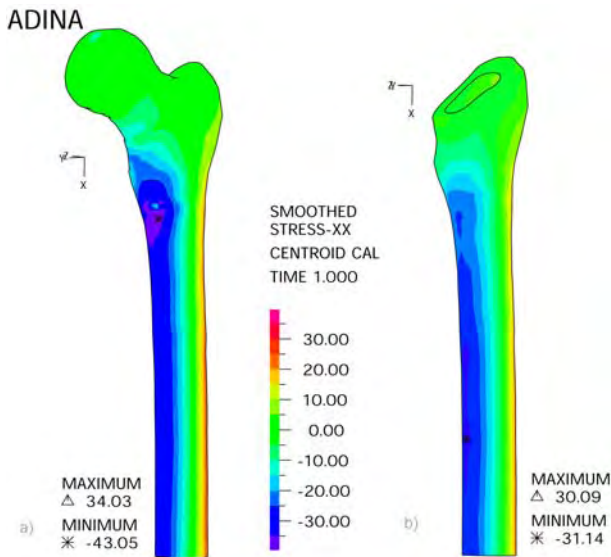


Fig.7. Distribution of normal tensions components σ_x [MPa] in the x axis direction - in the area of the cortical bone: a) in a proper bone model b) in the bone - cement - Centega endoprosthesis model

The largest differences in distributions of principal stresses values $\sigma_1, \sigma_2, \sigma_3$ in the cortical tissue layer are visible in the principal stresses σ_1 , where tensile stresses dominate (fig. 8) as well as in principal stresses σ_3 , where compressive stresses prevail (fig. 9).

Distribution of the strain energy density in the region of cortical tissue in a proper bone model has been shown in figure 10a. The maximum value of $U = 0.067$ [J/m³ $\times 10^6$] appears at the inner side of the bone near a smaller trochanter. Local changes in the strain energy density distribution caused by the endoprosthesis implantation can stimulate the changes occurring in the osseous tissue region. Implantation of the Centega type stem leads to reduction of the strain energy density in the cortical tissue region compared to the analyzed proper bone model (fig. 10b). The largest differences can be noticed in the region of bone

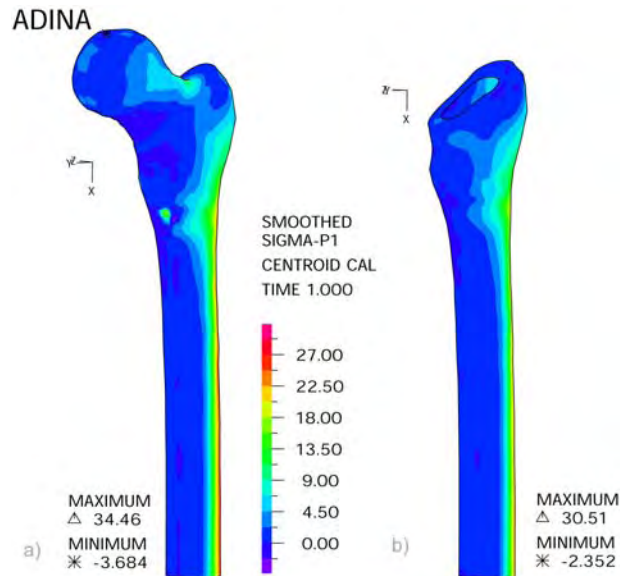


Fig.8. Distribution of the principal stress σ_1 [MPa] in cortical bone region: a) in a proper bone model, b) in a bone - cement - Centega endoprosthesis model

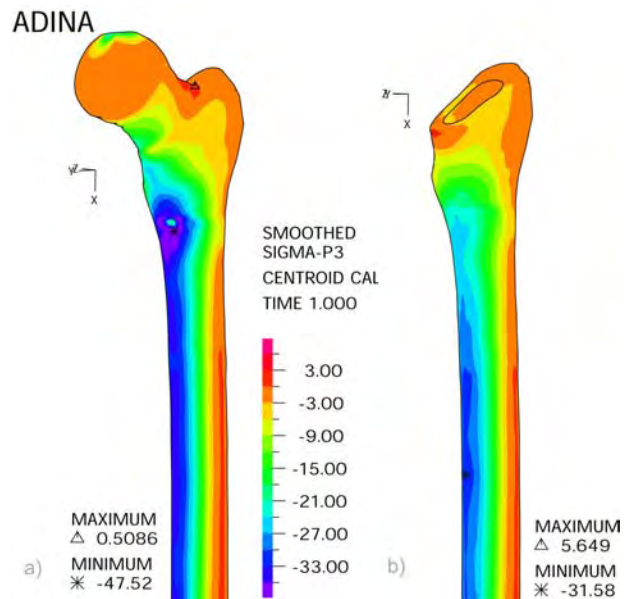


Fig.9. Distribution of the principal stress σ_3 [MPa] in cortical bone region: a) in a proper bone model, b) in a bone - cement - Centega endoprosthesis model

around the proximal part of implant sword as well as in the region of the femur spur.

When discussing the results of the numerical calculations it is essential to take into consideration the stress state in the region of cancellous bone. In a region occupied by the spongy tissue in a proper femur model the maximum value of the reduced stress $s_{zr \max} = 3$ [MPa] is observed inside a head of the femur in a region of the loading application (fig. 11a). In a bone - cement - Centega type implant model, in the cancellous bone region the largest values of the reduced stress occur at the paracentral side of the cancellous tissue region, at its end, as well as at the back side closer to the analyzed area. The maximum value of the reduced stress amounts to $s_{zr \max} = 2.4$ [MPa]. In the

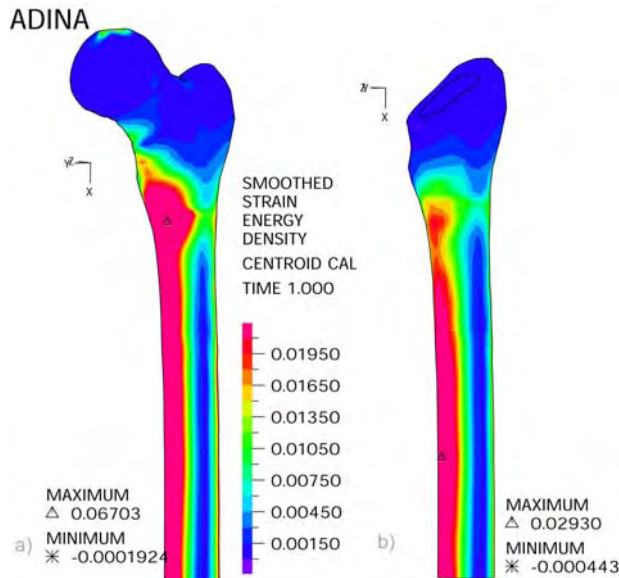


Fig.10. Strain energy density distribution U [$\text{J/m}^3 \times 10^6$] in the cortical bone region: a) in a proper bone model, b) in a bone – cement – Centega endoprosthesis model

VI zone these stresses have doubled in value compared to the stresses occurring in the cancellous tissue in a proper bone model (fig. 11b).

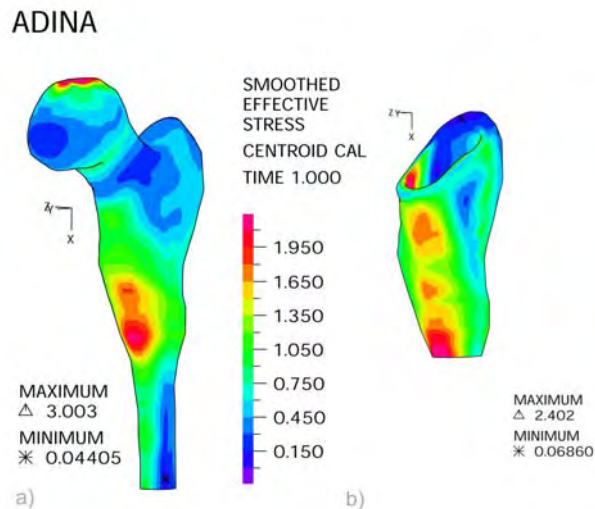


Fig.11. Reduced stress distribution σ_{zr} [MPa] in the cancellous bone region: a) in a proper bone model, b) in a bone – cement – Centega endoprosthesis model

Stress distribution in a stem as well as in a cement layer surrounding the implant significantly affects the durability and a stable fixing of the implant stem. The presence of regions of high concentration of stresses in both: the stem and the cement sheath can lead to a failure of a stem or to formation and expansion of cracks in the cement, being the cause of loosening of the implant.

Reduced stress distribution σ_{zr} occurring in the stem of the Centega type endoprosthesis has been shown in figure 12a. The maximum value of the reduced stress has been observed in the region of the implant neck and amounts to $\sigma_{zr \max} = 141$ [MPa]. The largest reduced stresses in the im-

plant sword occurred in the internal as well as external material layers at the central part of a stem and amounted to $\sigma_{zr} = 60 \div 100$ [MPa]. The extreme values of σ_x stresses have been observed in the external material layers of a stem in both: the internal as well as external side of the system and equal respectively to $\sigma_{x \min} = -121$ [MPa] and $\sigma_{x \max} = 100$ [MPa] (fig. 12b).

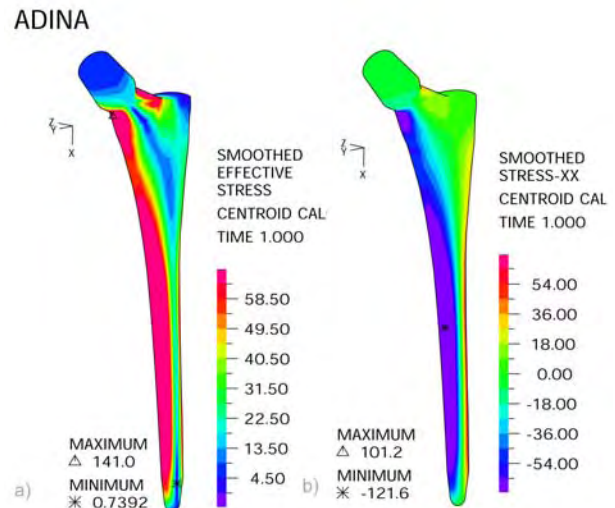


Fig.12. A reduced stresses distribution σ_{zr} [MPa] in Centega stem, b) normal stress components distribution in the $x - \sigma_x$ axis direction [MPa] in the implant stem.

Reduced stress distribution s_{zr} in the region of cement sheath around the Centega type stem has been presented in figure 13. The largest values of the reduced stresses have been noticed in a cement layer in the region of endoprosthesis stem end as well as in the cement area on the inside of the implant. The maximum value of the reduced stress the has been observed on the cement-implant contact surface in a region of a closer beginning of an implant sword at the paracentral side of the femur and amounts to $s_{zr \max} = 4.7$ [MPa].

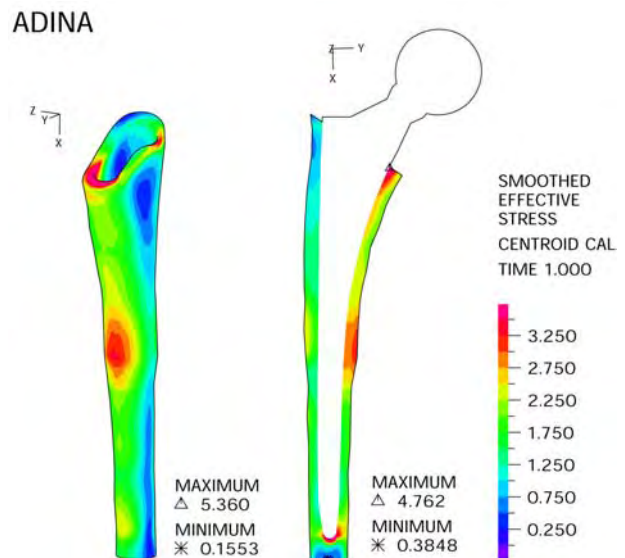


Fig.13. Reduced stress distribution σ_{zr} [MPa] in the cement layer in the bone – cement – endoprosthesis Centega model

4. Conclusions

Numerical calculations performed in the present work confirm the occurrence of the essential changes in a way of loadings transmission in the femur due to the implantation of the hip joint endoprosthesis stem. In a proper hip joint loadings are transmitted by a femur head to the area of spongy tissue, then to the external structure of cortical tissue closer to the femoral epiphysis and further to the subsequent section of diaphysis. The implantation of the endoprosthesis stem causes the disorders of the stresses state predominant in a natural hip joint. It also leads to unloading of the cortical bone region near the implant, depriving the bone of the specific stresses occurring in a natural hip joint. These changes are particularly visible in a metaphysis region of a closer femur.

The results of the calculations prove that the use of Centega type asymmetric stem leads to more profitable loading of the closer femur base as well as to more uniform stress distribution in the cement compared to the straight stems [11,19,20].

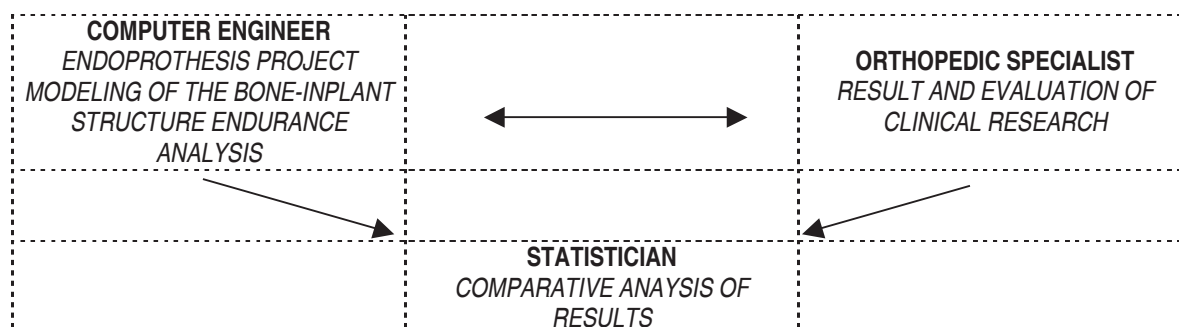
Implantation of the Centega stem causes the unloading of a bone on level VI, VII as well as II and III of the Gruen's zone, whereat the stresses in a region of the femur spur decreased to the largest extent, which can lead to resorption of the osseous tissue in this area. However in none of the analysed regions, places of strong stresses concentration or stress-free zones were observed. Their presence could lead to rapid loosening of the implant.

The clinical applicability of the anatomical Centega stem is also confirmed by the early results of the clinical investigations [16].

The investigations performed with the numerical models make it possible to quickly evaluate the influence of a new constructional solution of the endoprosthesis stems on the manner of the loading transfer in the region of an implanted hip joint, and therefore make it possible to explain the causes of the destructive processes occurring in the area of the implanted hip joint and to foresee the effectiveness of the implantation.

References

- ADINA. Theory and Modeling Guide, Adina R&D, INC. 1997
- Aesculap, CENTEGA – Panta Rhei, Prospekt Nr O-191-01,
- Aluisio F. V., Christensen Ch. P., Urbaniak J. R.: Ortopedia, Wydawnictwo Medyczne Urban&Partner, Wrocław 2000.
- Balin A.: Mechaniczne i materiałowe uwarunkowania stabilności trwałości endoprotezy stawu biodrowego – studium zagadnienia, Inżynieria Materiałowa, nr 2/1997.
- Bergman G., Deuretzbacher G., Heller M. i in.: Hip contact forces and gait patterns from routine activities, Journal of Biomechanics, 34 (2001), 859-871.
- Będziński R.: Biomechanika inżynierska. Zagadnienia wybrane, Wyd. Politechniki Wrocławskiej, Wrocław 1997.
- Funken S.: Änderung des femoralen Offset durch Implantation handelsüblicher Hüftprothesen, Duisburg 2004
- Kowalczyk P.: Numerical Evaluation of sensitivity of stress distribution in bone to geometric parameters of endoprosthesis. Journal of Theoretical and Applied Mechanics, No. 3, vol. 37 (1999), 555-577
- Krzesiński G., Zagrajek T.: Modelowanie własności mechanicznych kości. Biology of Sport, vol. 15, suppl. 8, 1998, 238-243.
- Kusz D.: Biomechaniczne aspekty endoprotezoplastyk stawu biodrowego, Inżynieria Materiałowa, nr 2/1997.
- Kusz D. Okrajni J.: Komputerowa symulacja rozkładów naprężeń w modelach kości udowej prawidłowej oraz po implantacji trzpieni typu Weller, Centrament i Parhofer-Monch. Chirurgia Narządów Ruchu i Ortopedia Polska, 65 (5), 519-527, 2000.
- Łaskawiec J., Michalik R.: Zagadnienia teoretyczne i aplikacyjne w implantach, Wydawnictwo Politechniki Śląskiej, Gliwice 2002.
- Marciniak J.: Biomateriały. Wydawnictwo Politechniki Śląskiej, Gliwice 2002.
- Prendergast P. J.: Finite element models in tissue mechanics and orthopaedic implant design, Clinical Biomechanics, vol. 12, No 6, pp 343-366, 1997.
- Prendergast P. J.: Biomechanical techniques for pre-clinical testing of prostheses and implants. Centre of Excellence for Advanced Materials and Structures, Lecture notes 2, Warszawa 2001.
- Puzicha D.F., Miehke R.K., Sundermann H.H., Strocker P.: Early results after primary hip replacement with the „centega” stem. ERASS Congress in Montpellier 18-20 October 2001: Symposium.
- Reilly D. T., Burstein A. H.: The Mechanical properties of cortical bone. The Journal of Bone and Joint Surgery, vol. 56-A, no. 5 (1974), 1001-1022
- Wall A., Dragan Sz.: Postępy chirurgii ortopedycznej w aspekcie rozwoju biomechaniki, Acta of Bioengineering and Biomechanics, vol 1, suppl. 1, 1999, s. 535-539
- Więckowski W., Włodarski J., Szarek A., Szyprowski J.: Prognozowanie wyników alloplastyki biodra z wykorzystaniem badań numerycznych, Kwartalnik Ortopedyczny, vol.58, nr 2, s. 167-173, 2005
- Wittmann J., Ziółkowski W., Krzemiński M., Omernik M.: Ocena przydatności klinicznej trzech rodzajów trzpieni endoprotez bezcementowych stawu biodrowego w oparciu o materiał Kliniki, Chirurgia Narządów Ruchu i Ortopedia Polska, 1996, LXI, Supl. 3A, s. 31-35
- Ziółkowski W., Kamiński M., Doerr-Winogradow K., Ostojski R.: Radiologiczne i histologiczne zmiany w jądrowym obluzowaniu udowej części endoprotezy stawu biodrowego według Wellera, Chirurgia Narządów Ruchu i Ortopedia Polska, 1996, LXI, suppl. 3A, s. 121-125



PICTURE GRAMMARS IN 3D MODELING OF MEDICAL STRUCTURES

GRAMATYKI OBRAZOWE W PRZESTRZENNYM MODELOWANIU WYBRANYCH STRUKTUR MEDYCZNYCH

MAREK R. OGIELA, RYSZARD TADEUSIEWICZ, MIROSLAW TRZUPEK

*AGH University of Science and Technology, Institute of Automatics, Biocybernetics Laboratory
30 Mickiewicza Ave, 30-059 Krakow, Poland*

Abstract. The paper presents the new opportunity for making semantic descriptions of medical structures with the use of graph-based linguistic formalisms. Discussed in detail are methods of computational intelligence to the development of a syntactic semantic description of spatial visualisations of the heart's coronary vessels. The methodology can be used for other goals related performance of computer-assisted semantic interpretation of selected elements and/or the entire 3D structure of the coronary vascular tree. The obtained semantic information makes it possible to describe the structure focused on the semantics of various morphological forms of the visualised vessels from the point of view of the operation of coronary circulation and blood supply to the heart muscle. Our approach allows a fast and, to a great degree, automated interpretation of semantics of various morphological changes in the coronary vascular tree, and makes it possible to detect the stenoses in the lumen of the vessels that can cause critical decrease of blood supply to extensive important regions of the heart muscle. This is achieved through the use of graph-based image formalisms based on indexed edge-unambiguous graphs generating grammars that allow discovering and automatic semantic interpretation of irregularities visualised on the images of the heart muscle from helical CT scans.

Key words: syntactic pattern analysis, medical image understanding, spatial modeling, coronary arteries

Streszczenie. W pracy omówiono możliwości tworzenia semantycznych opisów wybranych struktur medycznych z wykorzystaniem grafowych formalizmów strukturalnej analizy obrazów. W szczególności zaprezentowano sposób wykorzystania technik inteligencji obliczeniowej do tworzenia syntaktycznych opisów znaczeniowych przestrzennych wizualizacji naczyń wieńcowych serca. Proponowane podejście może być wykorzystane do zadań komputerowego wspomagania diagnostycznej interpretacji zobrażeń oraz analizy przestrzennych struktur unaczynienia wieńcowego. Możliwość otrzymania semantycznego opisu takich struktur umożliwia analizę prawidłowości funkcjonowania krążenia wieńcowego oraz określenia stopnia zaopatrzenia serca w krew utlenowaną. Tak prowadzona analiza w konsekwencji pozwala na przeprowadzenie automatycznej interpretacji znaczeniowej wielu różnych form zmian morfologicznych naczyń wieńcowych, w tym także wykrycie istotnych przewężeń światła tętnic, mogących ograniczać dopływ krwi utlenowanej do najważniejszych części mięśnia sercowego. Taką analizę oparto o grafowe techniki syntaktycznego rozpoznawania obrazów pozwalające na dokonanie semantycznej klasyfikacji nieprawidłowości krążenia wieńcowego uwidacznianych z wykorzystaniem technik spiralnej tomografii komputerowej.

Słowa kluczowe: strukturalna analiza obrazów, rozumienie zobrażeń medycznych, modelowanie przestrzenne, diagnostyka naczyń wieńcowych serca

1. Introduction

The problem of automatic understanding, not just analysis and recognition, of medical images is very important. The reason is that the ultimate goal of interpretation of medical images is always understanding of the essence of the pathological processes affecting the organ examined and design of the appropriate therapy. To attain this goal it is not enough to classify and recognise some pathological symptoms (either automatically or manually) since recognition only assigns names/labels to diseases, while treatment is performed on the diseased organ, or in short on the human subject. It simulates an experts' method of thinking: if one observes a deformation of the organ shown by the medical

image used then one tries to understand the pathological process that was the reason for the appearance of the deformations. One does not perform a mechanic classification for the purpose of pointing out more similar samples on the pathological image.

Current state-of-the-art methods of image diagnostics provide abundant and varied diagnostic and research material since 3D visualisations are available for practically all human body organs and structures. They allow seeing pathological changes to a greater degree and more completely, assuming that the physicians interpreting the image understand well what they see and what it means. Clinical practice shows that in spite of the frequent use of com-

puter-assisted diagnostic equipment the visualisations (including the spatial reconstructions of specific body parts) obtained later undergo only a rough qualitative assessment by the physician with no in-depth semantic analysis.

Practical implementation of the postulate of understanding 3D medical visualisations requires resorting to algorithms and information-theoretic techniques far more advanced than those used in traditional automatic understanding and classification. This is the role of science: it must outpace the potential offered by the methods known and generally applied in pursuit of new approaches that meet practical requirements better. For that reason, in this work, we shall deal with the construction of cognitive systems supporting the semantic analysis of 3D visualisations used in diagnostic systems.

Due to various meanings associated with term *understanding* and its crucial importance for this work, let us consider that a certain computer-assisted interpretation of the 3D image meets the requirements of “automatic understanding” when direct results of this interpretation include the 3D structure of the organ examined, the spatial relations between this structure and other organs and parts of the body, and what the consequences of this specific structure and relations are. Taking into account both diagnostic and therapeutic needs we shall require that the process of automatic understanding points to where, how, and with what consequences the pathological process has interfered with the natural (normal, though differing between patients) 3D structure of the given organ. Using this definition of automatic understanding of 3D medical images, and its practical implementation, one should be able to provide the physician with more useful guidelines and premises for therapeutic decisions than one would usually receive from traditional image recognition, which ends in the delivery of the decision (correct or not) stating the name of the disease entity and making reference to an assumed – usually in an arbitrary manner – taxonomy of the diseases considered.

2. The problem

Characteristic for the coronary vessels in different patients with various pathologic forms observed on flat, i.e. 2D, images, is a certain level of repetitiveness and regularity, however, the additional degree of freedom given by the transition from 2D to 3D visualization makes this question exceptionally intricate. Thanks to use 64-slice CT scanner which delivers optimal image quality in cardiac applications with a spatial resolution of 0,4 mm, the same images registered by a machine rendering 3D images have a much greater number of details. However, results also in the increase of the number of individual differences and differences between different examinations of the same patient. Having considered these points, we decided to use methods of automatic image understanding for the interpretation of the images, which requires their semantic descriptions. Application of linguistic description, using special languages such as graph image languages based on expansive graph grammar reaching into the semantic field of the examined patterns (also as images) opens the possibility

of its understanding with methods derived from parsing of appropriate grammars.

The parsers suggested in the article shall allow the conversion of image description built around the suggested language into a description in semantic terms and making it possible to understand all significant medical facts related to the image. The linguistic and cognitive methodology postulated in this article shall also allow creation of hybrid records at further stages of the analysis; such records will contain both the considered 3D images and their automatically generated cognitive descriptions containing a definition of their semantic properties. Finding appropriate languages, necessary for description of the semantic aspects of the 3D medical visualization, is a difficult and complex task. Description of important semantic aspects of the image should not be burdened with the details that – from the point of view of understanding the content of the image – are of secondary importance. This is why, importance for image understanding is the application of pre-processing methods and analysis of 3D medical visualizations adjusted to the specific features of this task. Such methods are developed besides the development of linguistic description of 3D images and intelligent methods of using the experts' experience and knowledge for acquiring the effect of cognitive resonance. Hence, this work also finds such methods of extraction and description of medical images features that ignore individual distinctive features characteristic for a given patient. At the same time they are focused on the exposition and appropriate representation of morphological properties crucial for the understanding of pathologies present in the visualizations. Only well-selected elements of the description at the level of automatically identified elements of the visualization and their spatial relations may become the basis for the development of a language description that would, at the parsing stage, allow for analyzing the semantics and detection of symptoms of the pathological process troubling the organ. The elements of the spatial description mentioned above, treated as letters of some alphabet symbols and groundwork for the construction of specific language formulas, must be oriented towards discovering pathologic changes: this will allow not only to localize such changes but also to perform an interpretation of their nature and, later, to establish their medical significance.

The goal of this article is to present the ways of taking the methodology [13] that has been developed earlier for the needs of automatic understanding of 2D medical visualisations to the level of 3D visualisations. It will be examined possibilities of defining, in these visualisations, elements describing the morphology of pathologic changes and the possibility of applying the linguistic approach, as well as cognitive image analysis for the description and modeling of spatial biomedical structures. We will in particular focus on the 3D structure of the coronary vascular tree that will be described with the use of graph-based languages of shape feature description [8, 10, 12, 14].

The selection of a linguistic technique results from the fact that – thanks to the use of advanced formalisms referring, among other things, to the knowledge and experience of expert physicians – these languages make it possible to

perform automatic identification of regions in the analysed image that are crucial from the point of view of medical diagnostics. After the definition of appropriate structural relations allow also for the location and definition of the degree of advancement of pathologic changes. Using methods of computational intelligence to a great degree improves the possibility of automatic interpretation of complex physiological and pathological processes continuing in the visualised organ, at the same time allowing for the performance of semantic analysis that makes computer-assisted understanding of images possible. The methods proposed in the article are oriented on these types of tasks.

3. Functional characteristics of coronary arteries

To understand morphological changes in an organ one must first become acquainted with the organ. We very briefly introduce some information from the vast repository of medical knowledge on coronary arteries. Let us recall here that the oxygen and nutrient supply for the heart muscle as well as the removal of final products of metabolism and carbon dioxide is performed through coronary circulation [2, 6, 9] is important for correct functioning of the coronary system. It confirms the advisability of analysis of these anatomic structures in a system of 3D presentation – for the well-being of patients. The individual branches of the coronary arteries are situated coil-like on the heart's surface so as not to cause increased flow resistance during the contraction and relaxation phases; they supply blood to defined areas of the heart, taking only minor branches to the aorta and the pulmonary trunk [2, 6, 9]. As the layout of the arteries on the diaphragmatic surface of the heart muscle may differ in its form, three basic types of distribution of these arteries may be distinguished. The first type is most widespread, and in 60–70% of cases it features a balanced distribution of arteries, with none being dominant. In the second type, the circumflex branch of the left coronary artery is highly retarded (poorly developed) while the right coronary artery is highly developed and is the main supplier of blood to the posterior surface of the left chamber. Such a situation is present in on average 20–24% of cases, with a variety of intermediate forms possible. In the third type, in only 10–14% cases, the right coronary artery is strongly retarded, while the posterior surface of the right chamber is vascularised by the left coronary artery, whose anterior interventricular branch may additionally reach the diaphragmatic surface [2, 9]. This shows what difficulties resulting from the lack of universal standards that would allow unambiguous representations of such many kinds of vessels may arise. In the case of the sternocostal surface, there are no significant variations of coronary arteries, with the exception being the variations in the anterior interventricular branches, which may be divided into two branches of similar thickness, both in its initial and final section [2, 6].

The key element of the diagnostic examinations of the heart performed with methods including the angiography of coronary arteries, helical CT scanning, and IVUS are a number of complex algorithms focussed on acquisition and pre-

processing of such images [1, 4, 8]. All these examinations are complementary and provide the physician with vast amounts of information that allows formulating appropriate diagnosis and reconstruction of information concerning the structure of coronary vessels. As concerns computer-assisted methods of 3D structure rendering it is possible to perform a visualisation of the 3D spatial structure [1, 3, 5]. Application of such methods makes it possible to obtain diagnostic material containing crucial information about the presence and/or advancement of pathologic changes characteristic for certain medical conditions. Unfortunately, no computer-assisted systems offering automatic interpretation that is comparable to that conducted by the physician's eye have been built. For that reason, the article describes ways of developing linguistic semantic representation of spatial structure of coronary vessels showing certain pathological changes and their analysis with the use of cognitive methods of image understanding. Thanks to such a linguistic semantic representation, it will later be possible to support interpretative tasks and define the semantics of actual pathologies.

4. Graph-based formalisms in the semantic description of coronary vessels

To introduce linguistic formalisms in the form of graph-based image grammar that allows finding stenoses in main arteries or their ramifications and, at the same time, quantification of the pathological changes, several visualization sequences were performed for the arteries in the DICOM format. Such sequences were obtained for various patients during diagnostic examinations of the heart conducted with tomographic angiography. Such visualizations show in a very clear manner all morphologic changes of individual sections of arteries in any plane. Furthermore, they allow performing a spatial rendering of coronary vessels. Shown in Fig. 1 is an example of a spatial reconstruction of the heart's coronary vessels, presenting the vessels without the accompanying muscle tissue of the heart. This reconstruction was obtained from the coronarographic images acquired during the diagnostic examination with a helical CT scanner with 64 detectors. It shows the main arteries along with minor vessels and a stenosis in the initial section of the right coronary artery. The structures made visible in this manner will be described with the use of graph-based grammars – constituents of the graph-based language defining their proper, spatial topology. Proper graph-based grammar describing these structures was defined so that the individual branches of the graph in the description provided identify all start and end points of coronary vessels and all bifurcations or transitions of main vessels into lower-level ones. Thus developed graph-based structure will constitute the elements of the language for defining spatial topology and correct vascularization of the heart muscle together with the potential morphological changes, e.g. in the form of stenoses in their lumen (Fig. 1).

Mutual spatial relations that may occur between elements of the vascular structure represented in the form of a

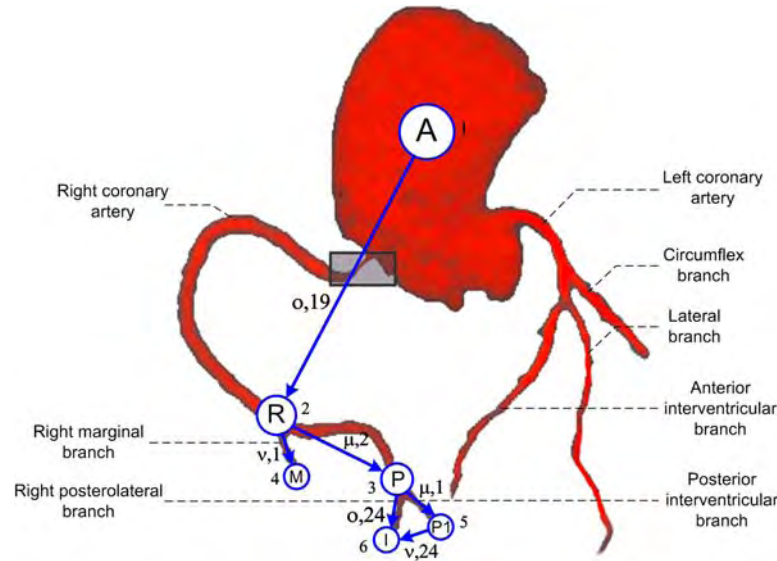


Fig. 1. Spatial labelling of coronary arteries and defining relations occurring between them for the right coronary artery. The grey rectangle marks the place of a major stenosis in a coronary artery.

graph describing the set of edge labels. Elements of the set are defined by the introduction of appropriate spatial relations: vertical, defined by the $\{\alpha, \beta, \dots, \lambda\}$ set of labels, and horizontal, defined by the $\{1, 2, \dots, 24\}$ set of labels which designate individual sections, each with the angular span of 15° on the hypothetical sphere surrounding the heart muscle. This hypothetical sphere represents the 'geographical' spatial relations of the graph representing the vascular structure.

Next, for labels defined in this manner, edge terminal labels are assigned to each of the bifurcations marked by the start and end points of appropriate coronary artery sections, depending on the situation of sections of the arteries. Once these relations are taken into account and node labels of the artery-spanning graph are entered, it is possible to develop a linguistic description for defining the semantics of the changes sought after and recognition of pathologic stenoses. Such creating the graph model of the 3D structure of considered vessels and its linguistic description among other things allows for conducting a computer-supported analysis of the rendered structure, oriented towards the automatic detection of places of stenosis, the degree of narrowing, and its type (concentric or eccentric). Such a representation provides a concise and unambiguous description of all the elements of the vascular structure in the form of a semantic record. Places where pathologic changes are present are tracked during the analysis of the graph-based representations, made with the edNLC type grammar. Such an analysis refers to the spatial representation of coronary vessels acquired thanks to the description developed with the application of IE graphs (indexed edge-unambiguous), a subclass of directed and edge-labeled directed graphs (EDG).

Below we repeat formal definition of the edNLC type grammar, while emphasizing the fact that this grammar class comes with highly efficient syntax analyzers for parsing in multinomial time [12].

Def. The name of G_{edNLC} (edge-labeled directed Node-Label Controlled graph grammar) grammar is applied to the set of five elements defined as follows [12, 13]:

$$G_{edNLC} = (\Sigma, \Delta, \Gamma, P, Z), \text{ where:}$$

Σ is a finite, non-empty set of node labels. These labels denote additional nodes in the graphs that are entered as temporary during the syntax analysis or while deriving the representation of a new coronary vascularisation.

Δ (a subset of Σ) is a set of terminal (i.e. end-marking) node labels. These labels describe the terminal elements of the modelled structure of the heart's vessels.

Γ is a finite, non-empty set of edge labels. This set makes it possible to define spatial relations occurring between the individual sections approximating the recognised vessels. P is a finite production set in the form of (l, D, C) , where $l \in \Sigma$, and D is an EDG graph. These productions define the graph-based language describing potential diagnostic cases.

$C : \Gamma \times \{\text{in}, \text{out}\} \rightarrow 2^{\Sigma \times \Sigma \times \Gamma \times \{\text{in}, \text{out}\}}$ is an embedding transformation defining the manner of connecting old and new nodes during the syntactic analysis conducted.

Z is the start graph, which is used to begin the process of generation of all graph-based forms for coronary arteries.

To define the location where the vessel pathology is present in the graph of coronary arteries in Fig. 1, the following grammar is proposed.

$$G_{edNLC} = (\Sigma, \Delta, \Gamma, P, Z)$$

where:

$\Sigma = \Sigma_N \cup \Sigma_T$ and is a set of both terminal and non-terminal node labels describing the examined graphs and defined as follows.

$\Delta = \Sigma_T = \{\text{right coronary artery, right marginal branch, right}$

posterolateral branch, posterior interventricular branch} – a set of terminal node labels (presented also in Fig. 1).

$\Sigma_N = \{A, B, C, I, M, P, P1, R\}$ is a set of non-terminal node labels.

$\Gamma = \{\mu, 1; \mu, 2; v, 1; v, 24; o, 19; o, 24\}$ is a set of labels describing edges of the graph.

$Z = \{A\}$ is the original starting graph.

P – is a finite set of productions recorded in the graph-based form and presented in Fig. 3

As the edNLC-class grammar derived above allows for spatial analysis of the structure described with the IE graphs, defined below as follows [13, 14].

Def. The name of an IE graph (indexed edge-unambiguous) is applied to the system:

$$H = (V, E, \Sigma, \Gamma, \varphi)$$

where:

V is a finite, non-empty set of graph node with unambiguously assigned indices,

Σ is a finite, non-empty set of node labels,

Γ is a finite, non-empty set of edge labels,

E is a set of graph edges in the form of (v, λ, w) , where $v, w \in V, \lambda \in \Gamma$ and the index v is smaller than the index w ,

$\varphi : V \rightarrow \Sigma$ – function of node labelling.

The IE graph presented in Fig. 2 together with the characteristic description of the nodes present therein may be used to describe the selected sections of the coronary vessels presented in the Fig. 1.

For the graph-based grammar introduced earlier, the set of deriving rules in the form of IE graphs, which allow for deriving spatial representation for the right coronary artery described in Fig. 1. is presented in Fig. 3.

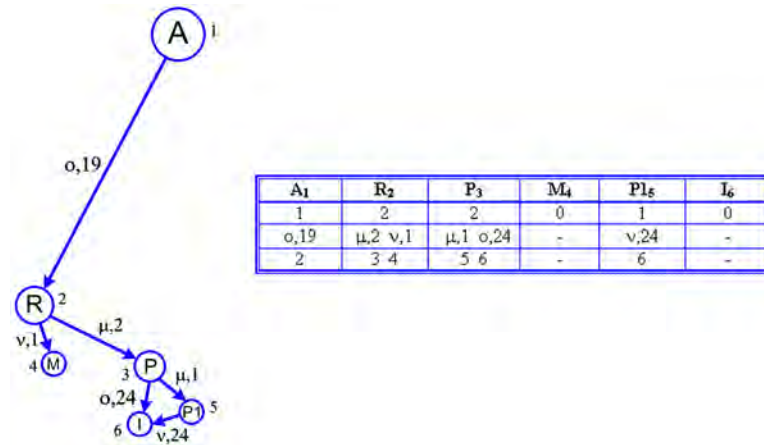


Fig. 2. Representation of the spatial reconstruction of the coronary vessels with IE graph together with the vertex characteristic description.

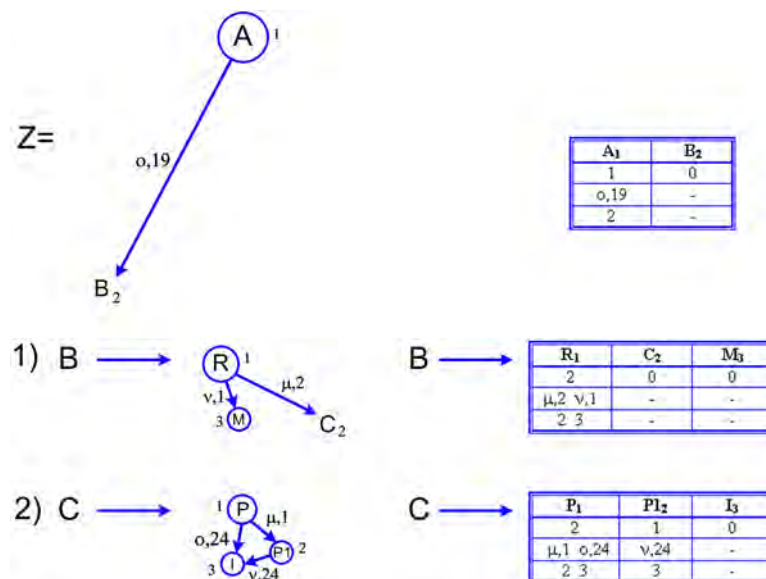


Fig. 3. The set of productions in the form of IE graphs deriving the structure of coronary vessels together with the vertex characteristic description.

5. Results

We have shown, thorough the application of the formalisms of the graph-based grammars presented in this work, that it is possible to develop semantic descriptions and spatially model morphology of the coronary vessels, as well as discover ischemic regions in the arteries. The discovery of such changes (in the form of stenoses in their lumen) is performed by determining the place of their occurrence i.e. the definition of the position of stenosis in the vessel and the degree of stenosis of the artery's lumen, and by designation of morphometric parameters defining its scope and type (eccentric or concentric). The important element of the syntactic and semantic analysis is parsing that, in the case of structural analysis of the representations developed, provides automatically (thanks to the use of semantic procedures) full information defining the spatial topology of the graph, describing the coronary vessels as well as discover in them the symptoms of ischemia. Subsequently, such information is used for the structural and semantic analysis of individual sections of the arteries that are defined by terminal node and edge labels. This allows for precise definition of morphological irregularities in the vessels examined that may be present in the form of various stenoses [11, 13]. All of the above makes it possible to develop a semantic record containing a semantic description of the structure in question. Such descriptions may provide a certain type of indexing keys that allow archiving and finding data (in the form of 3D visualisations) based on information concerning the pathological conditions present. Due to the small set of test data (covering several elements) the results achieved are highly promising. Such efficiency, however, is also the result of the high generalisation properties of the formalisms used.

The shown results show that thanks to the application of the described type of grammar, allowing for a semantic interpretation of spatial reconstruction of coronary vessels, it is possible to correctly analyse morphology of individual sections of arteries and their branches. This broadens greatly the scope of use of CI methodologies towards the computer-assisted methods of understanding diagnostic images and definition of medical significance of pathologies present.

6. Conclusion

The innovative cognitive interpretation of spatial visualisations of coronary vessels presented in this work contributes to the area of computer-assisted medical diagnostics. The application of graph-based image languages in spatial modelling of coronary structures helps to define the semantics of pathological changes. Application of such formalisms provides opportunities to focus interpretation tasks on the performance of automatic analysis and detection of actual morphological irregularities in a perceptive manner, which inspires the development of cognitive information systems based on graph-based image analysis languages.

Such methods are important as they allow for the discovery of pathologic changes and for definition of their se-

mantics, which – in the case of medical diagnostic images – may lead to computer-assisted ‘understanding’ of their medical significance. In the future it may also help to determine optimal therapeutic options. It is worth to note the vast progress of biomedical engineering in the field of intelligent information systems. It can be achieved on the base of analogy between the operations performed by the presented structural model of automatic analysis of the images and cognitive procedures taking place in the human mind. Although the methods of modeling of medical structures described in this paper are not precise models of cognitive processes taking place in the human mind, one can show deep analogy between a physician's thinking and computer processing. The analogy lies in the fact that between the stream expectation (knowledge collected in the form of graph-based grammatical rules) and the data stream that is obtained by means of analysis of the currently considered image, there must be an interference. Both in our computer programs and in a physician's mind two information processes take place. The first starts on the image analysis and comes to the internal description of the image features. The second starts from the knowledge about medical problems investigated by means of image analysis and leads to the expectations, which are some predictions in the form: “If some merit interpretation of the image is true, on the image must be found specific parameter or specific feature”. Therefore a very important part of analysis and semantic interpretation of the image is based on coincidences between expectations, obtained on the base of previous knowledge, and the features found on the image during visual inspection performed by the man or during automatic analysis performed by the computer. Sometime features and parameters localised on the image are pursuant with the theoretical expectations and this fact gain on importance of some specific semantic interpretation while others lose their importance. This interference leads to a cognitive resonance, which confirms one of certain possible hypotheses in the case of an image whose content can be understood or makes it possible to determine that there is a discordance, which cannot be removed. Despite the often complex processes of grammar-based reasoning, the syntactic methods of image recognition using graph-based image languages in the analysis of the spatial structure of heart vascularization provide nearly all the information for the description of morphological changes in coronary vessels. Such information may be found useful in tasks of diagnostic and therapeutic nature as well as in the tasks of context-sensitive semantic searches for specific visualisation cases.

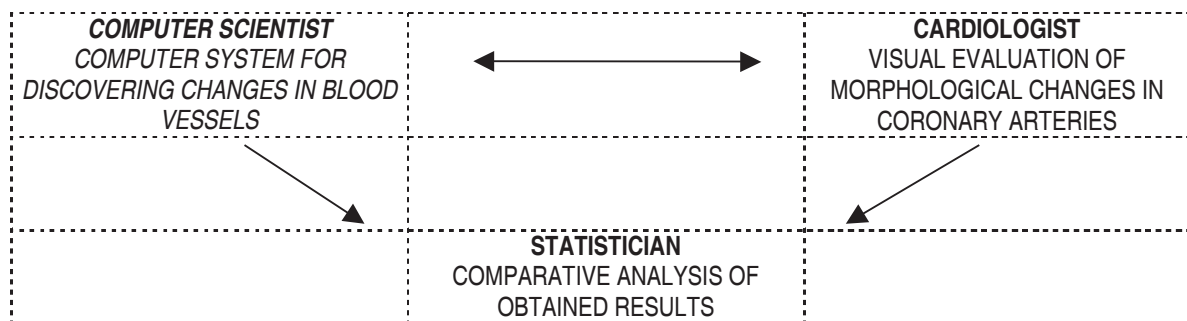
Acknowledgement

This work was supported by the AGH University of Science and Technology under Grant No. 10.10.120.39.

Bibliography

1. Bankman I. (ed.): *Handbook of Medical Imaging: Processing and Analysis*. Academic Press 2002.

2. Bochenek A., Reicher M.: Human Anatomy, vol. III. Cardiovascular system (in Polish), PZWL, Warszawa 1998.
3. Higgins W. E., Karwoski R. A., Ritman E. L., Spyra W. J. T.: System for analyzing true three-dimensional angiograms. *IEEE Trans. Med. Imag.*, 15: 377-385, 1996.
4. Higgins W. E., Reinhardt J. M.: Cardiac image processing, 789-804, in *Handbook of Video and Image Processing*, ed. A. Bovik, Academic Press, 2000.
5. Hoffman K. R., Sen A., Lan Li, Chua Kok-Gee, Esthappan J., Mazzucco M.: A system for determination of 3D vessel tree centerlines from biplane images, *The International Journal of Cardiac Imaging*, 16: 315-330, 2000.
6. Khan M. G.: *Heart Disease Diagnosis and Therapy*, Williams & Wilkins, Baltimore 1996.
7. Kurgan L., Cios K. J., Tadeusiewicz R., Ogiela M., Goodenday L. S.: Knowledge Discovery Approach to Automated Cardiac SPECT Diagnosis, *Artificial Intelligence in Medicine*, 23(2): 149-169, 2001.
8. Meyer-Baese A.: *Pattern Recognition in Medical Imaging*, Elsevier-Academic Press 2003.
9. Netter F.H., Colacino S.: *Atlas of Human Anatomy*, Novartis Medical Education, 1998.
10. Ogiela M. R., Tadeusiewicz R.: Picture Languages in Intelligent Retrieval of Visual Data Semantic Information, *LNAI* 3336, 389-396.
11. Ogiela M. R., Tadeusiewicz R.: Nonlinear Processing and Semantic Content Analysis in Medical Imaging – A Cognitive Approach. *IEEE Transactions on Instrumentation and Measurement*, 54(6): 2149-2155, 2005.
12. Skomorowski M.: *A Syntactic-Statistical Approach to Recognition of Distorted Patterns*. Jagiellonian University, Kraków 2000.
13. Tadeusiewicz R., Ogiela M. R.: *Medical Image Understanding Technology*. Springer Verlag, Berlin-Heidelberg 2004.
14. Tadeusiewicz R., Flasiński M.: *Pattern Recognition*, PWN, Warszawa 1991.



E-LEARNING AS A NEW TOOL IN MEDICINE TEACHING ON THE EXAMPLE OF AN INTRODUCTORY STATISTICS COURSE

E-NAUCZANIE NOWYM NARZĘDZIEM W DYDAKTYCE MEDYCZNEJ

ANDRZEJ A. KONONOWICZ, IRENA ROTERMAN-KONIECZNA

*Department of Bioinformatics and Telemedicine, Medical College, Jagiellonian University,
31-530 Kraków, Św. Łazarza 16, Poland, www.bit.cm-uj.krakow.pl*

Abstract: In this paper we present a short e-learning course concerning basic statistical concepts which represent quite an unattractive branch of knowledge for the medicinal students. Applets presented in this paper have been worked out in a manner linking the physician's intuition with methods found in statistical tools. This effect has been gained by showing entities from a sample (the language of statistics) in a form of individualized images of particular patients (the language of medicine). The demonstrated e-learning model allows the translation of the physician's experience in contact with many patients into a generalized analysis of a set with a great amount of elements. The sophisticated mathematical mechanism is hidden from the student, however the results have an significant influence in the analysis and interpretation of the data. The learning module presented in this paper has been judged by students.

Key words: e-learning, biostatistics teaching, Java applets

Streszczenie: W pracy przedstawiono krótki kurs e-learningowy dotyczący podstawowych pojęć statystycznych, które reprezentują dyscyplinę mało atrakcyjną dla studentów medycyny. Przedstawione w pracy aplety opracowane zostały w sposób łączący intuicję lekarza z metodami implementowanymi w narzędziach statystycznych. Uzyskano to dzięki prezentacji elementów zbioru (język statystyki) w postaci zindywidualizowanych obrazów poszczególnych pacjentów (język medycyny). Zaprezentowany model e-learningu pozwala przełożyć doświadczenie lekarza kontaktów z wieloma pacjentami na uogólnioną analizę zbioru o dużej liczbie elementów. Skomplikowany aparat matematyczny jest dla studenta niewidoczny aczkolwiek wyniki uzyskane mają istotne znaczenie dla analizy i interpretacji wyników. Przedstawiony w pracy moduł dydaktyczny poddany został ocenie studentów.

Słowa kluczowe: e-nauczanie, biometria, aplety Javy

Introduction

The dissemination of computers, the increase of their multimedia performance, as well as the fast growth of the Internet in recent years, leads to a great interest in the application of new technologies in education. E-learning contributes to an increase in efficiency of teaching due to a better adaptation of learning material to the actual knowledge and interests of a student. The multimedia setting of e-learning materials extends their attractiveness and suggestiveness. It also introduces dynamic elements which allow a continuous observation of calculation results. E-learning also enables a greater degree of mobility to the student, who is now able to reach learning materials from almost anywhere and at anytime depending only on the network availability.

The need of building up an e-learning infrastructure has also been perceived in Kraków. In the summer of 2005, the Blackboard Academic Suite E-learning Platform [6] had been installed in the Academic Computer Center Cyfronet AGH. The platform is jointly used by the University of Science and Technology AGH and the Jagiellonian University.

Statistics at a medical college

Statistics teaching at a medical college is nowadays more justifiable than anytime before. The information resources accessible by the Internet are so huge that their complete analysis is not possible with the usage of current tools. Facing the enormous dimension of databases waiting for their interpretation it becomes necessary to prepare appropriately the medical circles to consciously use statistics and proper software tools. The difficulty is to find a common language linking such seemingly different scientific fields as medicine and mathematics. E-learning seems to tie these fields by offering methods of learning based on experience. The ability to operate fast on large datasets combined with the possibility of data verification and dynamic generation gives a great opportunity to improve teaching methods without the requirement for medical students to know sophisticated mathematical operations which seems to be a barrier in teaching statistics to them. These methods may also bring statistics closer to the every day experience of physicians.

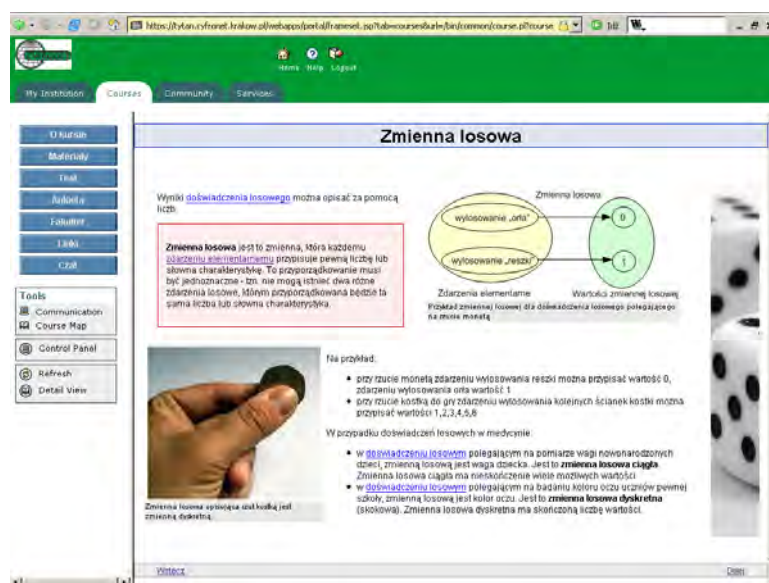


Fig. 1. Statistics e-learning module on the Blackboard

E-learning at the Department of Bioinformatics and Telemedicine

The Department of Bioinformatics and Telemedicine has actively joined the development of educational services since the very beginnings of Blackboard in Kraków. The first experience in e-learning was the creation of a short learning module dedicated to basic statistical concepts. The unit was presented to first and third year students of Medicine within the Telemedicine course.

Founding on the previous experience it is planned to organize regular courses in the summer's semester using the Blackboard platform.

E-Learning module introduction to statistics

The above mentioned e-learning unit (Fig.1) consists of six web pages dealing with topics, such as random variable, frequency distribution, probability distribution or the Gaussian function. The web pages have been made more appealing by using many photographs and figures. Color schemes have been applied to improve the legibility of the

module (e.g. definitions in red boxes, tasks in yellow). Three interactive Java applets have been worked out for the course showing the discussed topics (a comparison of the discrete probability function with the cumulative probability function, frequency distribution study for a normally distributed variable and an example showing the influence of μ, σ parameters on the Gaussian function's curve). The applets allow the students to experiment with the knowledge introduced in the unit.

The applets have been written in the Java programming language using the Anstat Beans library [4] developed at the Technical University Clausthal. Besides the web pages and applets, a multi-answer quiz was created giving the students the possibility to test their knowledge after reading the unit. The test was constructed using tools from the Blackboard platform.

Contact with an individual patient

Random samples are experienced by university students as highly anonymous, therefore from the physician's viewpoint as not very interesting. The physician wishes to contact the patient individually. Within the statistics course we have



Fig. 2. An applet presenting the measurement of a normally distributed variable

created an applet (Fig.2), which shows to a student the transition between the properties regarding the individual and the collectivity. From a mathematical viewpoint this applet is an illustration of frequency measurement of a normal distributed characteristic. We have chosen the newborns' weights as the examined characteristic. Clicking on the "next"-button shows a new infant picture, as well as the child's weight. The picture is unimportant from the mathematical viewpoint, but it helps the student to fall into the role of a doctor. All multimedia elements (e.g. patients' pictures, videos from procedures carried out) are always welcomed by the students – unfortunately such data is not easily obtainable.

In the above described applet the newborns' weights are divided into classes and then expressed as a histogram. Students may see how the histogram's shape transforms more and more into a form similar to the Gaussian curve as the number of weight measurements increases. Furthermore, they should notice that for small samples it is difficult to predict the actual distribution of the examined characteristics.

Knowledge gained by experiment

Taking into account the difficulty of some medical students in understanding the mathematical notation, we have created an interactive example comparing a discrete probability distribution function with the cumulative probability distribution function. Like the previous example, this one also deals with the weights of newborns. Five weight intervals are presented to students. Dragging the slider increases or decreases the number of elements in intervals. The applet allows the student to see, in an intuitive manner, the relation between the probability, distribution function, and the cumulative probability distribution function. We believe that students experimenting on their own with the concepts presented in the theoretical part of the unit, will consolidate and increase their knowledge in the discussed topic.

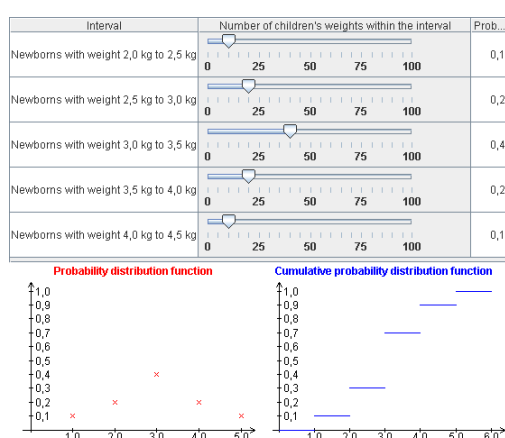


Fig. 3. An applet comparing the discrete probability distribution function with the cumulative probability distribution function

The students' opinion

The educational unit described in this paper was shown to first and third year students of medicine and dentistry

within the Telemedicine course conducted in the Department of Bioinformatics and Telemedicine. The students' reaction was mostly positive however there were exceptions. Third year students were asked to write a short free-text opinion on the topic of the presented e-learning module. Among first year students a more detailed inquiry was conducted, which aimed to find out students expectations and opinions about e-learning in the context of the presented statistics learning unit. All first and third year students present at the class about e-learning had the possibility to take part in the poll. The questionnaire for first year students was computer-based and anonymous. Questions about e-learning for third year students were part of the report after class. We received almost 60 third year students' opinions and 47 inquiries from first year students. The total number of medical students in a year is approximately 250.

Do you like the idea of e-learning?

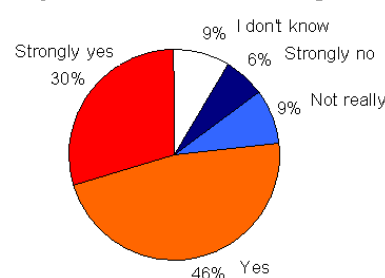


Fig. 4. Results of an inquiry among first year students into the popularity of e-learning

A great number of students think that multimedia techniques, interactive examples and many colorful images are a very suggestive learning method. This method allows getting away from the monotonic learning with traditional textbooks. A great advantage in the opinion of students is also the fact that e-learning allows to use e-learning materials from home, practically round-the-clock. Students also liked the possibility of adapting the learning scope and speed to the actual learner's needs. In the students' view e-learning is a good method for those who prefer gaining knowledge single-handedly.

Would you like to have more e-learning services at the college?

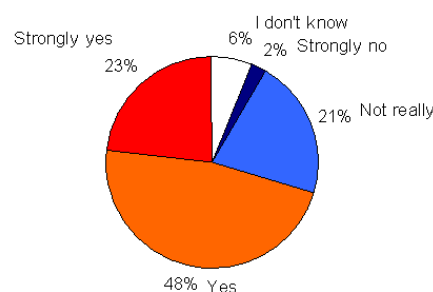


Fig. 5. Results of an inquiry among first year students into the expansion of e-learning

The biggest disadvantage of e-learning in the students' view is the lack of personal contact with the teacher and

schoolmates. Answers to detailed questions are often difficult to find in the e-learning course – it's easier to ask directly the teacher during the classes than to search by oneself in the learning unit. In the presented statistics learning module the way of explaining the term cumulative distribution function was not clear for some students.

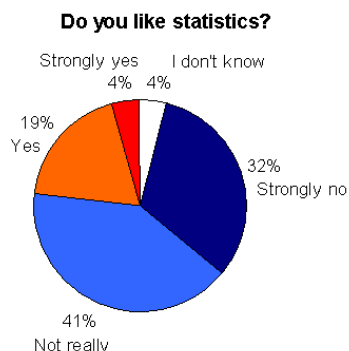


Fig. 6. Results of an inquiry among first year students into the popularity of statistics

The students are afraid that learning in front of a computer screen isolates them socially and has a bad influence on the health condition. The sound of the computer's power supply system and the temptation to access other internet pages during the learning time could be a distraction. According to the students, technical difficulties in the use of multimedia often occur, which is a barrier to the popularity of e-learning. Sometimes due to a complicated course structure they get lost in the e-learning unit. A disadvantage of e-learning is the high cost of Internet access in Poland.

Do you think that statistics will be useful in your future work?

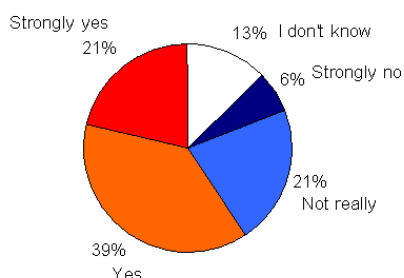


Fig. 7. Results of an inquiry among first year students into the awareness of the necessity of statistics

Giving their opinions about e-learning in general the students complained about the lack of access to complex and sophisticated e-learning services. There were opinions that e-learning helped only in gaining knowledge which is fragmentary and taken out of the context. In the students' opinion learning to use a computer is not taken seriously by high school teachers. Learning materials presented within e-learning courses do not correspond to the study curricula, therefore are of little use. Answering the question about the further development of e-learning services at the Jagiellonian University the students stated that they expected courses with a great quantity of multimedia content (movies; interactive and colorful multimedia illustrations).

According to students a perfect course should contain minimum words and maximum content; should have many summaries and quizzes testing the presented knowledge. An intuitive and not complicated learning module's structure is also very important.

The detailed inquiry was conducted among first year medical students. 47 students (approx. 20%) took part into the inquiry. Most of the polled students stated that they liked the idea of e-learning (76% answered yes or strongly yes compared to 15% strongly no or not really – Fig. 4). The majority of students want more e-learning services at the university (71% answered yes or strongly yes compared to 23% strongly no or not really – Fig. 5). Statistics is not popular in the opinion of medical students (to the question if they liked statistics 73% answers were strongly no or not really compared to 23% yes or strongly yes – Fig. 6). However, many students think statistics will be useful for them in their practice (regarding the question if they would need statistics in their future work 60% answered yes or strongly yes, but 27% answered strongly no or not really – Fig. 7).

Do you prefer the e-learning of statistics to the traditional statistics learning method?

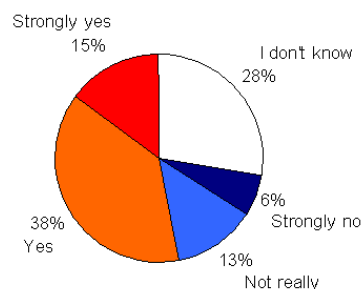


Fig. 8. Results of an inquiry among first year students into the e-learning method in preference to traditional learning method

Regarding the question if e-learning of statistics was more likeable than traditional learning 53% of polled students answered yes or strongly yes, 19% thought it was strongly no or not really; as many as 28% of students could not decide (Fig. 8). We think the great students' indecision in this question results from the little experience in learning statistics by e-learning methods. We have not done research yet to find out how effective the e-learning method is in comparison to the traditional method.

Summary

E-Learning encounters much resistance in the Polish education system. The apprehensions result mainly from lack of knowledge about this method of learning and their broad possibilities. It should be emphasized that e-learning will never totally replace the traditional learning form and personal contact with the teacher. In our paper we have presented an exemplary e-learning course introducing the field of statistical analysis. It provides the possibility of presenting dynamically the computation results. The peculiarities of the medical study should be considered in the statistics' teaching at the medical colleges. The medical educa-

tion mainly deals with the individual contact with the patient, therefore the medical statistics' e-learning should concentrate on the physician's practice. The presented mathematical dependencies ought to be illustrated by as many interactive examples as possible. They help the medical students to experience the gained information, without the knowledge of sophisticated mathematical calculations. However, in designing an e-learning unit we should consider the substantial time requirements needed for building high-quality interactive examples.

Most students like learning with computers. There is a great demand for the extension of e-learning services. The students questioned also recognized many of the e-learning drawbacks. Statistics is not very popular among our students, although the students are aware of the usefulness of this science. In our opinion, the lack of experience in using statistics e-learning units disallowed the students to answer which learning method they prefer – the traditional or e-learning.

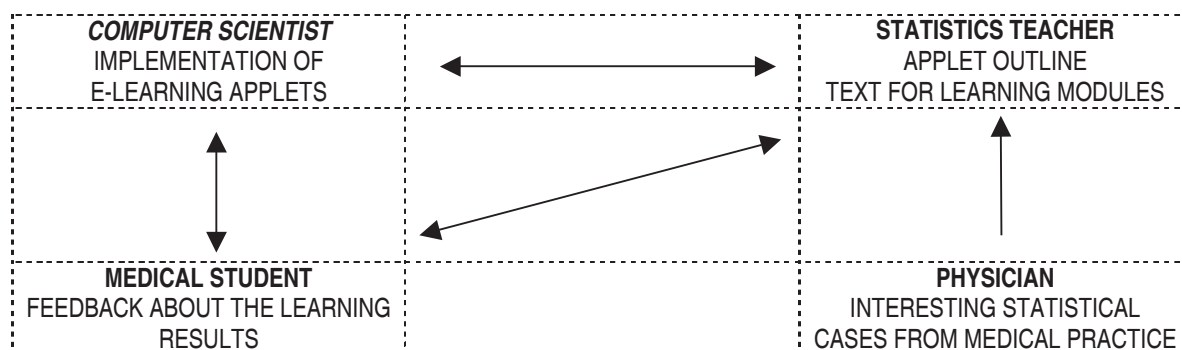
The whole replacement of traditional statistics courses by e-learning is in our opinion impossible. However, we believe that well designed and implemented multimedia illustrations and interactive examples accessible by the Internet will turn out to be a very valuable extension of the traditional statistics teaching style at medical colleges.

Acknowledgement

We would like to thank Prof. Michael Kolonko from the Technical University Clausthal for allowing us to use the Anstat Beans library. We also are grateful to M.Sc. Justyna Stefaniak for helping to write the text for the learning unit.

References:

1. Aydinli G, Härdle W., Rönz B.: E-learning/e-teaching of statistics: a new challenge, Proc. of the IASE Conference, Korea 2003.
2. Cramer E., Cramer K., Janzing P., Kamps U.: EMILeA-stat: A web-based learning environment in applied statistics with a focus on learning and teaching in secondary schools, Proc. of the IASE Satellite Conference on Statistics Education and the Internet, 2003.
3. Genschel U., Gather U., Busch A., EMILeA-stat: Structural and Didactic Aspects of Teaching Statistics through an Internet-based, Multi-medial Environment [in:] Härdle W., Rönz B., (ed.) Computational Statistics, Physica-Verlag, Heidelberg, New York, pp. 340–342.
4. Kononowicz A.: Anstat Beans – Visuelle Komponenten zur Erstellung von didaktischen Applets für die Statistiklehre, Student Research Project, Clausthal, 2004.
5. Saporta G.: St@Tnet, An Internet Based Software For Teaching Introductory Statistics, Proc. of the 6th International Conference on Teaching Statistics, Cape Town 2002.
6. Blackboard Academic Suite – <http://www.blackboard.com>



A TOOL FOR RESUSCITATION TRAINING

NARZĘDZIE DO NAUCZANIA RESUSCYTACJI

www.prc.krakow.pl

PAWEŁ KRAWCZYK*, GRZEGORZ CEBULA*, WIESŁAW PYRCZAK**, JANUSZ ANDRES*

* Chair and Department of Anaesthesiology and Intensive Care Medicine, Jagiellonian University
– Medical College, Faculty of Medicine, 31-501 Kraków, Kopernika Str. 17, Poland,
www.katedraanest.cm-uj.krakow.pl, katanest@cm-uj.krakow.pl

** Department of Bioinformatics and Telemedicine, Jagiellonian University
– Medical College, Faculty of Medicine, 31-034 Kraków, Kopernika Str. 7e, Poland,
www.bit.cm-uj.krakow.pl, bit@cm-uj.krakow.pl

Streszczenie: Resuscytacja jest dynamicznie rozwijającą się dziedziną medycyny. Aktualnie poszukiwane są coraz lepsze i gwarantujące większą dostępność do odbiorców techniki edukacyjne. Jedną z nich jest e-learning, czyli edukacja za pośrednictwem nowoczesnych mediów, między innymi komputerów, Internetu czy telefonii komórkowej.

Słowa kluczowe: e-learning, edukacja, resuscytacja, podstawowe czynności resuscytacyjne, zaawansowane czynności resuscytacyjne, Polska Rada Resuscytacji, Europejska Rada Resuscytacji.

Podstawowe i zaawansowane czynności resuscytacyjne warunkują przeżycie poszkodowanego, u którego doszło do zatrzymania krążenia. Wcześniej wdrożone i prowadzone we właściwy sposób mogą znacznie poprawić wyniki resuscytacji¹ (Fig.1.). Edukacja w tym zakresie rozwija się dynamicznie i poszukiwane są coraz wydajniejsze metody przekazywania i upowszechniania wiedzy o resuscytacji. Jednym z elementów tego procesu jest e-learning. Pozwala on na dotarcie do szerokiej rzeszy odbiorców z materiałami umożliwiającymi potencjalnemu ratownikowi posiadać wiedzę pozwalającą skutecznie wykonywać czynności resuscytacyjne, a co za tym idzie ratować życie ludzkie. Przykładem takiego działania jest strona internetowa Polskiej Rady Resuscytacji (PRR) /www.prc.krakow.pl/.

Polska Rada Resuscytacji została powołana w 2001 roku. Od listopada 2002 roku funkcjonuje jej strona internetowa www.prc.krakow.pl posiadająca aktualnie wielu sympatyków – liczba odwiedzających około 1000 osób dziennie.

Początkowo strona miała charakter informacyjno-sprawozdawczy i dotyczyła głównie działalności PRR. Z czasem pojawiły się tam coraz liczniejsze doniesienia przedstawiające aktualne osiągnięcia dotyczące resuscytacji. Należy tu wspomnieć m.in. stwierdzenia doradcze ILCOR (International Liaison Committee on Resuscitation) dotyczące hipotermii w opiece poresuscytacyjnej czy edukacji w resuscytacji.

Abstract: Resuscitation is a dynamically developing branch of medicine. Better and more accessible for receivers educational techniques are currently being sought. One of them is e-learning that is education by means of computers, the Internet, mobile phones etc.

Key words: e-learning, education, resuscitation, basic life support, advanced life support, Polish Resuscitation Council, European Resuscitation Council.

Basic and advanced life support determine the outcome after cardiac arrest. Early and properly conducted it may significantly improve the results of resuscitation¹ (Fig.1). Education in resuscitation dynamically develops and more efficient ways of transferring and disseminating such information are sought. One of the components of this process is e-learning. E-learning enables the wide range of receivers to reach materials making it possible to acquire knowledge in order to save people's life. An example of the actions mentioned above is the Polish Resuscitation Council (PRC) website /www.prc.krakow.pl/.

The Polish Resuscitation Council was founded in 2001. Its website www.prc.krakow.pl was established in November 2002. Currently it has many sympathizers with number of about 1000 daily visitors.



Fig.1. Steps of treatment algorithm.

Strona funkcjonuje w języku polskim (w przygotowaniu wersja angielska) i stanowi jedno z nielicznych tego typu źródeł informacji w naszym kraju, gdzie jeszcze wciąż znajomość języków obcych i dostęp do literatury fachowej są ograniczone.

Ważnym elementem umożliwiającym zdobycie wiedzy i umiejętności w zakresie resuscytacji krążeniowo-oddechowej (RKO) było przedstawienie propozycji kursów certyfikowanych przez Polską i Europejską Radę Resuscytacji. Są to: Podstawowe zabiegi resuscytacyjne oraz automatyczna defibrylacja zewnętrzna (Basic Life Support – Automated External Defibrillation), Natychmiastowa pomoc w stanach zagrożenia życia (Immediate Life Support), Specjalistyczne zabiegi resuscytacyjne u osób dorosłych (Advanced Life Support), kurs Zaawansowanych zabiegów resuscytacyjnych u dzieci (European Paediatric Life Support) oraz kurs instruktorski – Generic Instructor Course. Wprowadzenie za pomocą strony internetowej możliwości komunikacji pomiędzy potencjalnymi odbiorcami szkolenia a jego organizatorami pomogło usprawnić system rekrutacji uczestników szkoleń. Obecnie prowadzone są prace mające na celu zaadaptowanie systemu zarządzania kursami Europejskiej Rady Resuscytacji (Course Management System). Dostępny będzie on w dwóch wersjach językowych (polskiej i angielskiej) umożliwiając prowadzenie rekrutacji on-line jak i certyfikacji poszczególnych szkoleń. Planowany termin wprowadzenia w życie tego projektu to II kwartał br.

Przełomem w funkcjonowaniu strony było umieszczenie na niej polskiej wersji językowej „Wytycznych 2005 Resuscytacji Krążeniowo-Oddechowej” w grudniu 2005 roku. W połączeniu z innymi środkami przekazu (prasa, telewizja) doprowadziło to do znacznego zwiększenia liczby sympatyków strony.

Prowadzona statystyka wejść potwierdza duże zainteresowanie stroną głównie w ośrodkach edukacyjnych (m.in. województwa: mazowieckie, dolnośląskie, małopolskie, śląskie, pomorskie), wśród grup zawodowo związanych z medycyną i ratownictwem jak również indywidualnych internautów z kraju i zagranicy.

Poprzez stronę internetową prowadzona jest współpraca z organizacjami związanymi z propagowaniem idei resuscytacji jak i prowadzona jest ich promocja – projekt „przyjaciele” mający na celu cykliczne, cotygodniowe spoty prezentujące poszczególne portale (np. Ratownictwo, Fundacja Świat Dzieciom). W przygotowaniu są również Forum Dyskusyjne PRR jak i czat z osobami z grup roboczych PRR mające na celu wyjaśnianie i prezentacje ważnych zagadnień dotyczących resuscytacji. Obecnie takie działania prowadzone są za pomocą zapytań składanych drogą e-mailową przez skrzynkę PRR prc@prc.krakow.pl

Oferta edukacyjna kierowana jest także do najmłodszych odbiorców – we współpracy z Fundacją „Świat Dzieciom” przygotowywane są materiały edukacyjne (gry, rebusy, puzzle) pomagające małym ratownikom właściwie zachować się w stanie zagrożenia życia lub zdrowia.

Udostępnienie materiałów edukacyjnych przygotowanych na podstawie Wytycznych Resuscytacji Krążeniowo-Oddechowej (RKO) nastąpi po opracowaniu i zatwierdzeniu ich ostatecznego formatu przez Europejską Radę Re-

Originally the page was created in order to inform and report about events connected with PRC activity. With time, more and more up to date resuscitation reports have been presented there. Not to mention the ILCOR (International Liaison Committee on Resuscitation) advisory statements concerning hypothermia as a part of post resuscitation care or education in resuscitation.

This Polish website /English version to be prepared/ is a unique source of this kind of information in Poland where both knowledge of foreign languages and access to publications for specialists are still limited.

An important element in the process of acquiring knowledge and skills in resuscitation has been the introduction of wide range of courses certified by Polish and European Resuscitation Councils. These are Basic Life Support – Automated External Defibrillation Course, Immediate Life Support Course, Advanced Life Support Course, European Paediatric Life Support Course and Generic Instructor Course. Thanks to the possibility of communication via the website between potential candidates and course organizers the enrolment system has been improved. This idea has been developed and there is much effort put in order to adopt a European Resuscitation Council Course Management System (CMS). It will be working in both the Polish and English languages, giving a possibility to enroll on-line and to improve the certification system. Planned time of the implementation of the CMS is the second quarter of 2006.

A crucial moment in the page history was presenting the full Polish version of the “European Resuscitation Council Guidelines for Resuscitation 2005” in December 2005. In addition to information spread by the press and TV it has led to a great increase of the website visitors number.

The statistics show great interest in the page especially from main Polish educational centers (provinces: mazowieckie, dolnośląskie, małopolskie, śląskie, pomorskie), groups of healthcare providers as well as individuals from Poland and foreign countries.

The PRC collaborates with the organizations involved in disseminating the idea of resuscitation. Thanks to the website the PRC also promotes these organizations – project “friends” created in order to present resuscitation portals (e.g. Ratownictwo, Świat Dzieciom); each of them every week. Debating forum and chat, with head persons from the PRC working groups to be prepared. The aim of it all, is to explain and present important issues connected with resuscitation. Until the mentioned above is ready there is a possibility of asking questions and sending them to the PRC e-mail box: prc@prc.krakow.pl.

There will be educational materials (games, puzzles, rebus) for the youngest visitors of the page, to be prepared in collaboration with “Świat Dzieciom” foundation. It will help young rescuers to find the best way to manage when facing life or health threats. The number of educational films, slide shows and posters will be increased and updated according to the 2005 guidelines. Via the website there is possibility of ordering materials published by the PRC on-line.

The primary objective of the PRC is education in resuscitation. It has been executed using many ways. The website described above is one of the well functioning ways of knowl-

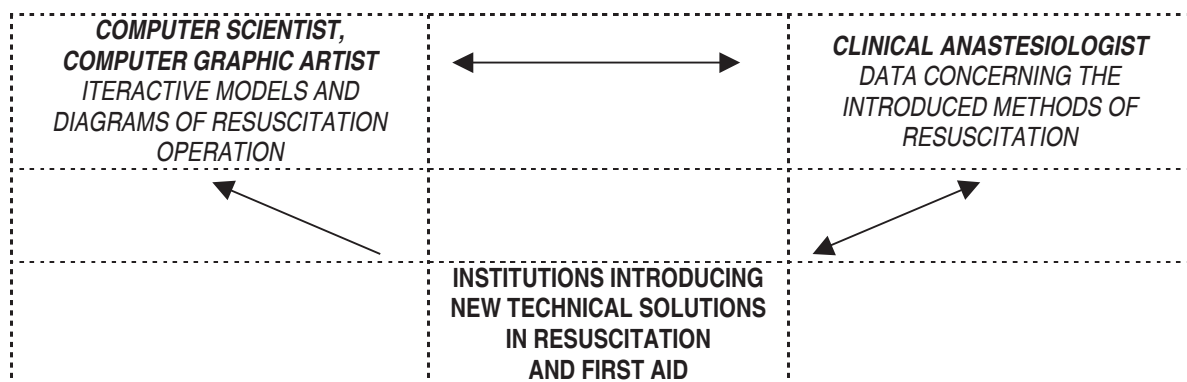
suscytacji. Planowane jest również rozszerzenie oferty filmów edukacyjnych, pokazów i schematów prezentujących aktualne wytyczne RKO. Poprzez stronę internetową dostępne jest również wydawnictwo PRR.

Podstawowym celem PRR jest edukacja w resuscytacji. Realizowany jest on na wiele sposobów. Przedstawiona powyżej strona internetowa jest jednym z dobrze funkcjonujących sposobów przekazywania tej wiedzy, a rosnące wciąż zainteresowanie jest najlepszym dowodem na potrzebę rozwoju tego środka edukacji.

edge transfer. Increasing the number of page visitors constitutes proof of the need to develop such a means of education.

1. European Resuscitation Council Guidelines for Resuscitation 2005.

1. Wytyczne 2005 resuscytacji krążeniowo-oddechowej, Kraków 2005.



BREAST CANCER – E-LEARNING GUIDE

E-SUTEK – SERWIS EDUKACYJNY

www.telemedycyna.cm-uj.krakow.pl/e-sutek

BARTŁOMIEJ ŁASOCHA, PRZEMYSŁAW MATUSZEWSKI, WIESŁAW PYRCZAK

*Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College,
Faculty of Medicine, 31-034 Kraków, Kopernika Str. 7e, Poland,
www.bit.cm-uj.krakow.pl, bit@cm-uj.krakow.pl*

Streszczenie: Praca prezentuje serwis internetowy gromadzący zasoby edukacyjne związane z nowotworem gruczołu sutkowego.

Zamieszczone w nim materiały koncentrują się wokół głównego zagadnienia, jakim jest rak piersi i pochodzą z wielu gałęzi medycyny min. patologii, radiologii, epidemiologii. Przy opracowaniu serwisu „E-sutek” wykorzystano nową technikę edukacyjną, jaką jest e-nauczanie. Witryna jest adresowana nie tylko do studentów medycyny i nauk przyrodniczych, a także do wszystkich innych zainteresowanych profilaktyką i leczeniem schorzeń piersi.

Słowa kluczowe: e-nauczanie, edukacja, rak piersi, mammografia

Abstract: In this paper a website that collects education resources concerned with breast cancer is presented. Materials are focused on the main problem (breast cancer) and come from many brands of medicine e.g. pathology, radiology, epidemiology.

The “E-sutek” service was created using a new educational technique – e-learning.

The website is addressed not only to medical and life science students, but to everyone who is interested in prophylaxis and the treatment illnesses of the breast.

Key words: e-learning, education, breast cancer, mammography

Prace nad serwisem internetowym zatytułowanym „E-learning w zakresie problematyki raka sutka” (www.telemedycyna.cm-uj.krakow.pl) rozpoczęły się w styczniu 2004. Strona ta od początku tworzenia była pomyślana w celu pomocy nauki kolegom z młodszych lat. Dlatego też zostały w niej przedstawione działy medycyny związane tematycznie z piersią (anatomia, histologia) jak i z rakiem piersi (genetyka, epidemiologia, patomorfologia, problem diagnostyki i terapii, rehabilitacja).

Podczas tworzenia poszczególnych podstron staraliśmy się, aby informacje zawarte w nich były jak najnowsze i jak najbardziej dokładne (oczywiście w zakresie wymaganym od studenta). Dlatego też na poszczególnych etapach pracy prosiliśmy o pomoc i konsultacje wielu pracowników Collegium Medicum Uniwersytetu Jagiellońskiego oraz Instytutu Onkologii im. Marii Curie-Skłodowskiej w Krakowie.

W serwisie znalazł się film nagrany specjalnie na potrzeby serwisu z badania gruczołów piersiowych (oraz węzłów chłonnych pachowych) w poradni schorzeń sutka przy I Katedrze Chirurgii CMUJ. W tejże poradni został także zarejestrowany dźwięk z przykładowego wywiadu z pacjentką w trakcie kontroli. Nagranie to pokazuje przybliżony schemat jak powinien wyglądać wywiad z osobą z guzkiem piersi. Należy zaznaczyć, że wywiad ten został ukierunkowany na schorzenie danej pacjentki.

The work over an internet site titled “E-learning – breast cancer” (www.telemedycyna.cm-uj.krakow.pl) has begun in January 2004. The web page since the very beginning was thought to help colleagues to study this problem. The teaching process has been selected to show the spectrum of disciplines starting from anatomy and histology through the genetic base of cancer, epidemiology, pathology, the issues of diagnostics and therapy, and rehabilitation.

During the creation of particular sub-sites we took care to provide the most accurate and up to date information (of course in the range expected from a student to master). That is why over certain stages of the framework we asked many doctors and scientists employed by the Collegium Medicum at the Jagiellonian University and The Maria Curie-Skłodowska Institute of Oncology, for help and advice.

There is a film found in the website presenting an examination of lactic glands (and axillary lymph nodes), recorded to full fill our needs in The Regional Out-Patient Clinic for Breast Illnesses, at The First Clinic of Surgery JUMC. Also a tape recording of the sound during an example anamnesis of a patient with breast cancer took place there. Those samples show the interview and the examination of a patient with a breast lump. It is to be stressed that the questions were focused on the disease under consideration.

Zdjęcia mammograficzne zostały wykonane w w/w poradni i zeskanowane w Zakładzie Radiologii CMUJ (Fig. 1).

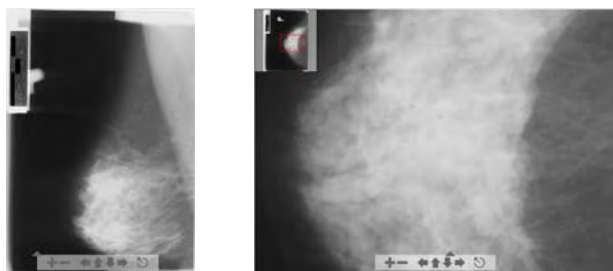


Fig. 1 Obraz mammograficzny (67 lat) bez zmian nowotworowych.

Obrazy histologiczne prawidłowej piersi zostały zamieszczone dzięki uprzejmości Zakładu Histologii (Fig. 2).

Wszystkie rysunki i zdjęcia zawarte w serwisie zostały wykonane przez autorów.

Krótki opis serwisu przedstawia się następująco:

EPIDEMIOLOGIA – podane tu zostały najważniejsze informacje epidemiologiczne jak zapadalność w populacji polskiej, czynniki ryzyka.

Dział **BUDOWA PRAWIDŁOWA** dzieli się na:

- **ANATOMIA** – tu zawarliśmy krótki opis gruczołu sutkowego poparty odpowiednimi rysunkami. Szczególną uwagę zwróciliśmy na spływ chłonki z gruczołu.
- **HISTOLOGIA** – opis budowy histologicznej gruczołu, ze zwróceniem uwagi na rodzaj tkanek. Wszystko jest podparte odpowiednimi zdjęciami preparatów histologicznych

BADANIE – w części tej opisano badanie podmiotowe i przedmiotowe. Wskazano na co należy zwrócić uwagę podczas badania piersi. W załączeniu jest tu film z badania piersi oraz przykładowy wywiad z pacjentką pochodzący z badania kontrolnego.

GENETYKA – w dziale tym zawarliśmy informacje o mutacjach wpływających na rozwój raka piersi. Mutacje te odszukaliśmy w bazie NCBI, oraz zostały one przedstawione graficznie. Zamieściliśmy tutaj też krótkie ćwiczenie przedstawiające jak poruszać się po bazie NCBI.

RADIOLOGIA – tutaj zostały opisane podstawowe metody diagnostyki opisowej, wraz z projekcjami oraz przykładowe zdjęcia diagnostyczne.

PATOMORFOLOGIA – w dziale tym podzielonym na cztery części zostały opisane najczęstsze typy raka sutka, zostały opisane również metody pobrania materiału do badania histopatologicznego, podział TNM nowotworów piersi (Fig. 3).

Dział **LECZENIE** – przedstawia metody stosowane w leczeniu raka piersi, kryteria doboru do wybranych metod terapeutycznych, oraz kilka słów o rehabilitacji w raku piersi.

DONIESIENIA – jest działem przedstawiającym najnowsze doniesienia ze świata dotyczące problematyki raka piersi. Są one przedstawione w formie abstraktu, podzielone na każdy miesiąc.

Całość materiału jest bogata w odniesienia do licznych stron w Internecie.

Na stronie zliczane są odwiedziny wraz z monitorowaniem adresów odwiedzających. Zamieściliśmy również linki dla zgłaszania uwag i korekt.

Original mammograms collected in the clinic mentioned above were scanned in the Radiology Plant of the JUMC (Fig. 1).

Histological samples of a normal breast could be published courtesy of the Histology Plant (Fig. 2).

All drawings and pictures on the website were produced by the authors.

The following is a quick review of the site:

EPIDEMIOLOGY – the most important information was provided here, such as the incidence rate in the population of Poland, risk factors or preventive measures.

NORMAL STRUCTURE is divided into:

- **ANATOMY** – a short description of the mammary gland supported by a series of drawings. Special attention is focused on the lymph drainage from the gland.
- **HISTOLOGY** – a presentation of histological structure of the gland, with drawings and microphotographs included.

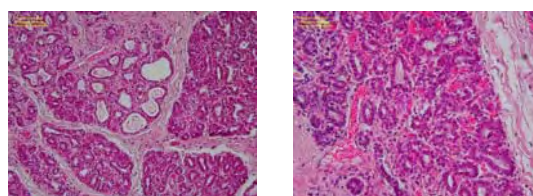


Fig.2. Histological structure of the normal gland in lactation period

EXAMINATION – the subjective and objective part of examination is described. It is pointed out what to be careful about, and pay attention to. An exemplary film of the examination and the anamnesis of the patient are attached.

GENETICS – this part contains information concerning mutations that influence the development of breast cancer. Most of them were found in the NCBI databases and their pictorial representation follows. A short exercise was published also, to encourage those who read, to use the NCBI.

RADIOLOGY – Basic methodology of imaging as well as enumeration of mammographic views and exemplary X ray films were presented.

PATHOLOGY – this part, subdivided into four units begins with the major histologic types of breast cancer. The methods of tissue sample acquisition for microscopic examination are described in details together with the TNM classification of breast tumors (Fig. 3).

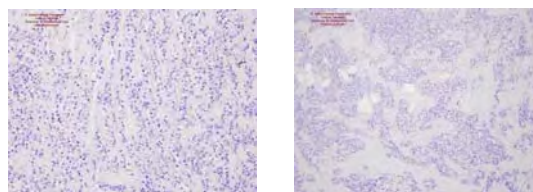


Fig. 3. Glycogen-Rich Carcinoma (Glycogen-Rich, Clear Cell Carcinoma)

The **THERAPY** part – methods used in breast cancer treatment, qualification criteria for selected treatment methods and a few statements about rehabilitation in breast cancer.

Stosujemy system oznaczeń jak specjalistyczny jest materiał niektórych podstron, lub jak drastyczne jest zdjęcie, w postaci oznaczeń barwnych. I tak:

ZIELONY – oznacza informacje ogólne i podstawowe

ŻÓŁTY – zakres materiału, który jeszcze może zainteresować przypadkowego użytkownika, a przeznaczony głównie dla osób częściowo związanych z medycyną

CZERWONY – to informacje potrzebne studentom medycyny, jak również przydatne dla lekarza, zdjęcia z zabiegów itp.

Tworzeniu witryny od początku towarzyszyła idea załączenia testu służącego utrwaleniu, nie sprawdzeniu tego, czego odwiedzający się dowiedział. Opracowywany test będzie również dostępny na różnych poziomach zaawansowania, z podziałem na dziedziny: osobny test dla epidemiologii, osobny dla patomorfologii oraz jeden zbiorczy adresowany do studentów wydziałów lekarskich.

The *REPORTS* section reflects the most up to date messages connected with the matter of mammary gland cancer. We publish abstracts every month.

Most of the text is rich in links to numerous website in the Internet.

The visits on the page are counted with notice about the location of visitors' computers. We have also provided a link for feedback correction.

The reader on the webpage can represent different preparation to the problem. Thus the color scale is introduced to inform the reader as to the character of the information eliminating these pictures, which may be very stressful for unprepared readers:

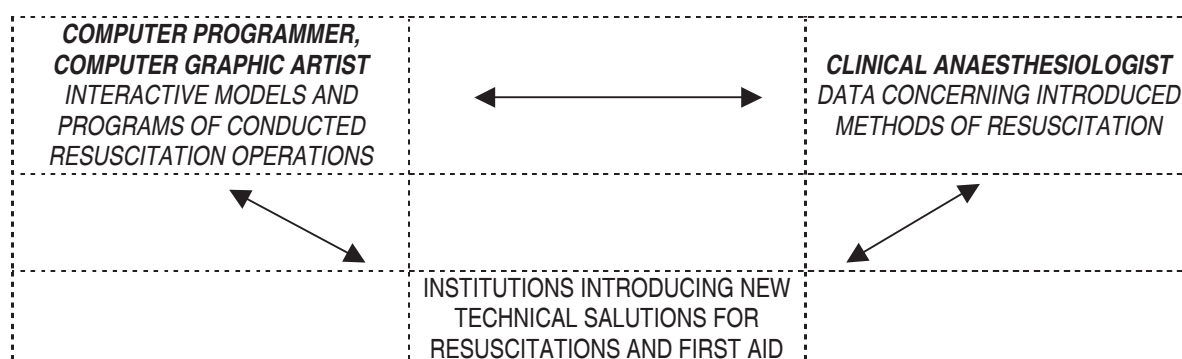
GREEN – open for the general public

YELLOW – for mature and aware visitors

RED – open for medical students and doctors and teachers.

Since the very beginning of the project, we have remembered in mind to provide a test to consolidate rather than check what the visitor learnt. The test is not ready yet, probably it would also be accessible in different levels. We do not exclude that it could also be divided into disciplines: separate for epidemiology, separate for pathology. One collective, a summary for students of medical faculties.

We use a system of marking how specific a text is on the page you are reading (or a sub page you are about to read) and how drastic are the pictures by using color tabs. As such: green - means basic and general information, Yellow a range of knowledge that might appear interesting to a random user, but is intended mainly for people at least partially connected with medicine. Red eventually presents information needed by medical students, as well as doctors and operating theater photos.



COMPUTER SCIENCE FOR PHARMACY STUDENTS. VERIFICATION OF EARNED ABILITIES BY USING THE INTERNET

INFORMATYKA DLA STUDENTÓW FARMACJI. WYKORZYSTANIE INTERNETU DO SPRAWDZANIA ZDOBYTYCH UMIEJĘTNOŚCI

www.farmacja.cm-uj.krakow.pl/~mysmieta/studenci/

JACEK ŚMIETAŃSKI

*Department of Pharmacokinetics and Physical Pharmacy, Medical College, Jagiellonian University,
Medyczna Str. 9, 30-688 Kraków, Poland; mysmieta@farmacja.cm-uj.krakow.pl*

Streszczenie: Artykuł prezentuje e-learningową stronę internetową, będącą pomocą w nauczaniu podstaw informatyki oraz praktycznego wykorzystania dostępnego oprogramowania, przede wszystkim przeznaczoną dla studentów Wydziału Farmacji CM UJ. Oprócz ogólnych informacji oraz konkretnych wskazówek dotyczących wykonania poszczególnych zadań, witryna wyposażona została w system weryfikacji nabytych przez studentów umiejętności. Nie jest to jednak test wyboru (jak w większości tego typu rozwiązań), a określone zadania do wykonania, co pozwala na sprawdzenie praktycznych umiejętności. Rozwiązania przesyłane są bezpośrednio za pomocą omawianej strony.

Słowa kluczowe: e-learning, informatyka, farmacja, test

Abstract: The following article presents an e-learning website which has been created as a support in teaching of basic computer science and practical use of available software, mainly for the students of the Pharmacy Department CM UJ. Beside general information and concrete tips for particular tasks, the site has been equipped in a system for verifying the acquired skills. However it is not a multiple-choice test (like in most solutions of this kind), but tasks are given to solve. This enables us to verify practical skills. The answers are sent directly by the website.

Key words: e-learning, computer science, pharmacy, test

Umiejętność korzystania z komputerów w życiu codziennym oraz znajomość najpopularniejszych programów użytkowych staje się koniecznością. Trudno dziś wyobrazić sobie osobę z wyższym wykształceniem, nie znającą choćby elementarnych umiejętności z zakresu obsługi komputera. Nic więc dziwnego, że podstawy informatyki znalazły się w programie niemal wszystkich kierunków studiów.

Prowadząc zajęcia z informatyki dla studentów Wydziału Farmacji UJ, miałem okazję przekonać się, jak trudno jest zrozumieć treść przedmiotu, korzystając jedynie ze skryptu wyjaśniającego wykonanie poszczególnych ćwiczeń. Trudności te zainspirowały mnie do stworzenia strony internetowej, gromadzącej w jednym miejscu przydatne informacje i materiały.

Strona została podzielona na sześć działów. W pierwszym (*Trochę formalności*) przedstawiono informacje dotyczące warunków uzyskania zaliczenia. Drugi dział (*Pracownia*) wyjaśnia problemy związane z logowaniem się i korzystaniem z dostępnego w pracowni komputerowej sprzętu. Bardzo przydatną informacją jest instrukcja łączenia się ze swoim kontem uczelnianym z dowolnego komputera na

The ability to use a computer in everyday-life and the familiarity with the most popular programs becomes a necessity. It is difficult to imagine a well educated person who does not possess even the basic knowledge about using a computer. No wonder that computer science has been placed in the programs of the majority of faculties.

I teach computer science at the Faculty of Pharmacy at the Jagiellonian University. I could convince myself how difficult it is to understand the subject only with the aid of a script with explanation of each single exercise. These difficulties have inspired me to create a website which contains all useful information and materials.

The site has been divided in six sections. The first one (*Trochę formalności*) presents information concerning the terms of attestation. The second section (*Pracownia*) explains the problems of logging in and using the available equipment. A useful information is the instruction how to connect with the student's account from any computer in the world. The rule of testing, examples of tasks and an exemplary test has been presented in the next part (*Sprawdziany*). In the section *Ćwiczenia* can be found a list of all tasks, questions

świecie. W kolejnej części (Sprawdziany) przedstawione zostały zasady sprawdzania nabytych umiejętności, przykładowe zadania oraz przykładowy sprawdzian. W dziale Ćwiczenia znaleźć można listę wszystkich ćwiczeń, zagadnienia do przygotowania oraz obszerną bazę odnośników do stron zewnętrznych, dogłębnie opisujących poszczególne zagadnienia. Można stąd także ściągnąć skrypt do ćwiczeń oraz niektóre z programów wykorzystywanych na zajęciach (z powodu ograniczeń licencyjnych programy zostały udostępnione jedynie studentom Wydziału Farmacji). Kolejny dział (Wskazówki do ćwiczeń) zawiera dodatkowe, nieobecne w skrypcie, wskazówki do wykonania poszczególnych zadań. Wyjaśnione są tu też sprawiające najczęściej kłopotów zagadnienia ogólne. Ostatni dział (Takie inne różne) zawiera luźne strony nie związane bezpośrednio z nauczaniem informatyki. Są tu też informacje kontaktowe.

Zgromadzenie w jednym miejscu i usystematyzowanie informacji oraz odnośników przydatnych w nauczaniu informatyki studentów analityki medycznej i farmacji z pewnością ułatwia znalezienie potrzebnych materiałów oraz zrozumienie omawianych problemów i wykorzystanie w praktyce nabytych umiejętności. Umieszczenie tych informacji w Internecie umożliwia korzystanie z nich nie tylko w wydziałowej pracowni komputerowej, ale w dowolnym miejscu na Ziemi.

Nieodłączną częścią edukacji jest weryfikacja zdobytych umiejętności w formie różnego rodzaju testów i sprawdzianów. Trudno jednak sprawdzać umiejętność pracy z komputerem podczas teoretycznego testu pisanego na sali wykładowej (nie neguję zasadności takiego testu, gdyż teoretyczne przygotowanie z podstawowych zagadnień informatyki może bardzo ułatwić zrozumienie praktycznego wykorzystania omawianych programów). Oczywiście, sprawdziany pisane na komputerze to żadna nowość, są to jednak albo testy wyboru sprawdzające bardziej wiedzę niż praktyczne umiejętności, albo sposób przekazywania rozwiązań nie jest najlepszy.

Można oczywiście przekazywać rozwiązania za pomocą tradycyjnych nośników informacji, czyli dyskietek. Te jednak bywają bardzo zawodne, aby więc mieć względną pewność, że zadania będą odczytane poprawnie, trzeba je zapisywać na kilku dyskietkach naraz. Również przesyłanie odpowiedzi pocztą elektroniczną bywa dość ryzykowne. Wiadomości nie dochodzą do adresata w czasie rzeczywistym. I choć na ogół czas ten nie przekracza kilku minut, zdarza się, że e-mail krąży po sieci przez kilka godzin nim trafi do właściwej skrzynki. Jest też spore ryzyko pomyłki ze strony studenta. Drobną błąd przy wpisywaniu adresu e-mail może spowodować, że wiadomość w ogóle nie dojdzie. Stwarza to też możliwość różnego rodzaju oszustw, np. wysyłanie zadań już po zakończeniu sprawdzianu albo wysyłanie rozwiązań przez inną osobę, podszywającą się pod studenta.

Zaproponowany przeze mnie system sprawdzania nabytych umiejętności wykorzystuje stronę internetową. Losowanie i wyświetlanie zadań, a także przesyłanie rozwiązań odbywa się za pomocą strony www.

to prepare and a large base of useful external links to the sites which describe the single questions in details. It is also possible to download the script and some of the programs used during the classes (due to the licence limits the programs are available for the Faculty of Pharmacy students only). The next section (*Wskazówki do ćwiczeń*) contains additional information, not present in the script, how to do single tasks. It explains also the general questions which bring the most problems. The last part (*Takie inne różne*) presents additional sites which are not directly related to the subject. Here is the contact information, too.

Gathering in one place and systematizing of information and links which are helpful for teaching computer science to pharmacy and medical analytics students, surely makes it easier to find the needed materials and to comprehend the discussed problems and practical use of the acquired skills. By placing of this information on the Internet it has become possible to reach them not only from the faculty's computer laboratory, but also from any place on the Earth.

An inherent part of the education is the verification of the acquired skills in the form of different tests. However, it is difficult to test the ability of computer use during a test written in the lecture room (I do not deny the sense of such a test because the theoretical preparation of the basic computer science questions can make comprehension of the practical use of the discussed programs much easier). Of course, testing with the aid of a computer is nothing new, but either it is rather a testing of knowledge than practical skills, or a way to transfer solutions is not the best one.

It is certainly possible to deliver answers on the traditional carriers of information, this means diskettes. However, these are often very unreliable. It is necessary to save copies on several diskettes to be relatively sure that the tasks will be read correctly. The sending of the answers by e-mail can be risky, too. Messages do not come to the recipient in real time. Though it generally does not take more than a few minutes, it can happen that the e-mail circulates several hours in the network before it reaches the proper address. There is also a risk that the student makes a mistake while typing the address so the message will not reach its target. This method makes swindles possible, too. For example the answer can be sent after the end of the test or sent by someone who lines the student.

The system which I propose, uses a website to verify the acquired skills. The drawing and displaying of tasks, as well as the sending of answers proceeds with the aid of the website.

The advantages of the projected system are the following:

- each student can get another, previously drawn tasks set;
- time can be measured for each student individually;
- there is a continuously active clock which shows exactly the remaining time on the site with the tasks;
- the reminder about the coming end of the work is sent automatically;

- Zalety zaprojektowanego systemu są następujące:
- każdy student może otrzymać inny, uprzednio wylosowany zestaw zadań;
 - czas pracy może być mierzony indywidualnie dla każdego studenta;
 - na stronie z zadaniami znajduje się wciąż aktywny zegar, dokładnie pokazujący ilość pozostałego czasu;
 - przypomnienie o zbliżającym się końcu pracy jest wysyłane automatycznie;
 - rozwiązania dostarczane są w czasie rzeczywistym – docierają do prowadzącego bezpośrednio po wysłaniu;
 - istnieje możliwość przesłania częściowo wykonanego zadania oraz wielokrotnego wysyłania rozwiązań poprawionych; nie powoduje to jednak bałaganu i utrudnienia dla sprawdzającego, gdyż ostatecznie pozostaje u niego tylko najnowsza wersja rozwiązania;
 - istnieje możliwość wysłania pojedynczego, całkowicie wykonanego zadania jeszcze przed zakończeniem całego sprawdzianu; prowadzący może je od razu sprawdzić
 - aby zapobiec podszywaniu się pod inne osoby, na jedno nazwisko może zalogować się tylko jedna osoba
 - hasło potrzebne do zalogowania jest podawane tuż przed sprawdzianem, nie ma więc możliwości wcześniejszego obejrzenia uprzednio wylosowanych zadań.
- the answers are transferred in real time;
 - it is possible to send a partially done task and to send corrected answer many times; it does not cause any difficulties for the checking person because only the latest version is shown to him;
 - it is possible to send a single, completely done task before the end of the test; the checking person can check it at once;
 - to prevent line other persons, there is only one student who can log in using the concret name;
 - the password needed to log in is given just before the test, so it is impossible to see previously drawn tasks.

There have also been elaborated a script for the checking person. It makes the verification of the sent answers easier. The received results are on currently saved on the server. Thanks to that the student can see his or her result immediately after the answers were checked.

The website is on currently, completed and updated. In case its functionality is extended it is provided for placing the site on the e-learning platform VBoard. The site is addressed to the Jagiellonian University's Faculty of Pharmacy students mainly. However, the materials can be used by everyone. It is available in Polish only.

Opracowany został także skrypt dla prowadzącego, ułatwiający sprawdzanie nadesłanych rozwiązań. Otrzymane wyniki są na bieżąco zapamiętywane na serwerze, dzięki czemu student może zobaczyć swój rezultat niemal natychmiast po sprawdzeniu pracy.

Strona internetowa jest na bieżąco uzupełniana i aktualizowana. W przypadku rozszerzenia jej funkcjonalności przewiduje się umieszczenie jej na platformie zdalnego nauczania VBoard. Adresowana jest głównie do studentów Wydziału Farmacji UJ, jako pomoc w procesie uczenia się obsługi komputera. Z umieszczonych tu materiałów może jednak skorzystać każdy. Witryna dostępna jest wyłącznie w języku polskim.

VIRTUAL REALITY – E-EDUCATION

PROF. JAN TRĄBKA

Collegium Medicum UJ

Streszczenie. Gdyby nie opracowano elektronicznych procedur edukacyjnych, nie można by było wykorzystać czterech głównych odkryć, które gruntownie zmieniły aktualny pogląd na świat: 1) Filozofia procesowa Whiteheada, 2) Biofizyka w postaci Topologicznej Geometro-Dynamiki (TGD) Matti Pitkanena, 3) Wielopokładowa hierarchiczna budowa kosmosu, która wprowadziła ujemną energię i ujemny czas oraz swoiste prawo „handlowe”: „bierz teraz, a zapłaci później ktoś inny”, 4) Wykorzystanie nowej p-adycznej (opartej na prymach) matematyki, która zaoferowała ciągi nieskończone, ale o skończonej liczbie wyrazów w ciągu.

Due to the process-philosophy of A. N. Whitehead there was created an ontology of the possibilities that have been estimated as the subjective kind of virtual reality. In order to manage this virtuality by scientific methods three other, additional presumptions, should be made previously. First of all, the classical physical view had to be substituted by the biophysics in the form of “Topological Geometro-Dynamics” (TGD) worked out by M. Pitkanen. Moreover instead of the algebraic, p-adic (based on the primes) mathematics should be accepted, that brought into real practice

special infinite series, but with finite number of elements. Thus all psycho-physiological states of consciousness can be described without the danger of ambiguity: genes, memes, qualia and semitrance. New aspects of TGD offered the hierarchy, many sheeted space-time multidimensional model of the universe, equipped by negative energy, negative time and a strange law “take now, pay later”. All these new scientific achievements might not be taken advantage of, if they were not supported by e-Education, i.e. by aids of powerful computing and effective modeling a new world-view has been rebuilt.

1. Whitehead A.N.: Science and Modern World. Macmillan, New York 1926.
2. Pitkanen M.: Topological Geometrodynamics and p-Adic Numbers. Internal Report HU-TFT-IR-95-5, (1995) (Helsinki University). Report on application of p-adic numbers in attempts to understand quantum field theory limit of TGD. <http://www.physics.helsinki.fi/~matpitka/>.
3. Pyrczak W., Lasoń W., Sarapata K., Walecki P., Roterman-Konieczna I.: E-Learning i telemedycyna – problemy strukturyzacji wiedzy, W: Komputer w edukacji, red. J. Morbitzer, Wyd. Naukowe Akademii Pedagogicznej, Kraków 2006.

MAŁOPOLSKA INFORMATION TECHNOLOGIES CLUSTER – SUPPORTING REGIONAL DEVELOPMENT

SŁAWOMIR KOPEĆ

*Governor's Office of the Małopolska Region, The Information Society Department,
Basztowa 22, 31-156 Kraków*

According to the Ernst&Young European Attractiveness Survey 2006 [1] Poland is ranked the fifth most attractive country in the world for foreign direct investments (FDI). In one year Poland jumped from 12th to 5th place in the "FDI Confidence Index" published by A. T. Kearney [2]. U.S. and European investors rank Poland their 7th and 3rd most attractive location, respectively.

In recent years Poland has generated significant interest as an off- and nearshore destination. On the neoIT Off-shore Attractiveness Index [3] Poland is ranked forth in the world in Information Technology Outsourcing (ITO) and even third in Business Process Outsourcing (BPO).

Another reason for optimism is that the Polish IT industry expands conspicuously – the growth rate in 2005 exceeded 30% with rather promising perspectives for the future [4].

Małopolska holds a strong position among Polish regions in terms of ITO, shared services outsourcing and BPO as well as R&D centres. In particular, neoIT recognizes Kraków as the Tier 1 – highly attractive - ITO and BPO city of excellence [5].

In this context the question appears: what steps should be taken by major stakeholders in Małopolska to benefit from such propitious circumstances and not only retain its strong position, but make the region the IT hub of Central Europe?

In the research "The truth about foreign direct investment in emerging markets" The McKinsey Global Institute [6] presents crucial factors affecting the choice of an off-shore location. The most important are high-quality infrastructure, availability of trained workers and regulatory environment. It is worth to note that financial incentives are the least important factor for a potential investor. We can make a safe assumption that similar factors are of vital importance also for already established undertakings.

Determination and invention of regional stakeholders may result in improvement of the quality of the infrastructure as well as the manpower. Although one may argue that changing the regulatory environment rests in hands of central authorities, regional actors undoubtedly have many ways to support the change by diagnosing problems and suggesting ways of overcoming them. By all means, the crucial and critical factor is a close cooperation of industry,

academia and regional authorities. And precisely this particular idea lies behind the concept of establishing the Małopolska Information Technologies Cluster (MITC).

Among the founders of MITC are renowned higher education and research institutions (such as the AGH University of Science and Technology and the Jagiellonian University), the largest IT enterprises and the authorities of the Małopolska region. The mission of the MITC [7] is enhancing growth of the IT industry as a sector of regional opportunity (cf. "Małopolska Region Development Strategy 2007-2013" [8] and "Małopolska Regional Innovation Strategy 2005 - 2013" [9]). The main objective of the cluster is to establish an efficiency-enhancing, value-adding environment, facilitating collaborative research relationships, knowledge and technology transfer, raising competitiveness of existing IT companies and effective support for new IT investments, optimal distribution of public funds on boosting research and innovation, and, last but not least, securing the steady inflow of highly qualified manpower in the region.

In general, the term "cluster" is rather fluid. It may denote either formal or rather loose structure of either exclusively firms from one sector or firms and other institutions. MITC may be considered as a regional cluster in Porter's sense, ie. as a *geographic concentration of interconnected companies and institutions in a particular field* [10] or even as an attempt to establish the regional innovation system, allowing for cooperation of firms and various institutions for knowledge development and diffusion (cf. [11]). MITC is not a legal entity - its role is to be rather a facilitator and coordinator of specific undertakings (such as the IT park or competence and training centre), for which formal project consortia of some cluster's partners will be established.

The long-term challenge for MITC may be clearly derived from the global tendencies on the labour market. As Deloitte Research point out, by 2008 the first members of the Baby Boom generation will turn retirement age and a wealth of skills and experience will begin to disappear from the job market in the large developed economies of North America, Europe and Asia. Over the next 15 years 80 percent of their workforce growth will occur among people 50 years or older [12]. The shortage of new workers in developed economies is especially striking in science and engi-

neering. The reason is self-evident: only 5 percent of U.S. students earn undergraduate degrees in science and engineering, compared to 42 percent in China. In Germany the number of engineering graduates has declined by almost a third within 10 years to one-tenth the number of China. Facing such figures one can not be surprised that Siemens decided to develop its new cell phones in Beijing.

It is clear that if Poland and Małopolska in particular want to benefit from their current positive image and long-term global talent market tendencies, the issue of building up size and competency of the IT labour pool is of critical importance. Therefore development of the training system suiting the needs of the IT industry should be the first MITC task. Fortunately, the starting point is rather good. Fifty percent of engineers in Poland could work for multinational companies (MNC), compared to 25 percent in India and 10 percent in China or Russia. As a result Poland has nearly as many suitable young professional engineers as Russia (and Central and Eastern Europe – as many as India) [13]. The data for Małopolska are also encouraging. The Marshall's Office of the Małopolska Region estimates that within the next 5 years the number of graduates able to work in multinational companies will exceed 50 000 [14]. A substantial fraction of them, however, will need additional training to match the market needs, and this will require an active cooperation of regional stakeholders in the triangle: academia, industry and administration. Their engagement should result in creating a complete, coherent and ready for immediate implementation training system.

When we look from the long-term perspective, we arrive at the conclusion that the system should be supplemented by a wide range of measures raising interest in science and technology, especially among young generation. Such steps are taken by numerous companies throughout the world. CEO of SAS Jim Goodnight explains the reason of the company's involvement in the US schools: *All corporations should get involved in the school system. The future of our country is in producing highly educated people. Otherwise we'll lose our high-tech jobs to India and China* [12]. The company introduced the program SAS in School Teach to prepare students for college-level learning [15]. Motorola introduced their Diversity program, well known also in Małopolska schools. The aim of the program is *to give students a hands-on learning experience that can excite them about business and technology, show them the possibilities of a career in a high-tech field with a company like Motorola, and perhaps help them on the way to college* [16]. This is the model for the partners of MITC to follow. The message should be simple: *the future is in science and engineering*. Moreover, such activities should be spread throughout the region – it is clear that the cities, where higher education institutions are located (eg. Nowy Sącz, Tarnów, Oświęcim, Sucha Beskidzka, Nowy Targ), may serve as the second tier locations for ITO and BPO investments. Parallel actions should be taken to raise awareness of importance of elevating the language skills among young people.

Needless to say, what really matters is not only the size, but also the quality and competence of our talent pool. Again, this can be achieved only by close cooperation between

universities and industry supported by public administration. This is intimately connected to another crucial activity – differentiating capabilities of the Małopolska talents from other regions (cf [3]).

Another question frequently raised in this context is: what will happen if the prosperity is over and global companies focus their attention on other low-wage countries? The answer is simple: we should concentrate on building unique competencies of our manpower, or, in other words, try to move up the value chain from BPO to Knowledge Process Outsourcing, and R&D in particular (this is the way the Indian government has chosen in mid-1990s with conspicuous results in recent years). While multinational companies tend to match their demand for innovation with supply of available talents and ideas by building global "innovation networks" [17], Małopolska should try to make its presence as visible as possible in the R&D part of the global value chain. This is an important task for forward-looking policy makers; analysing recommendations provided for other countries (eg. [18]) may be instructive here.

In the field of R&D outsourcing standings of Poland are good – nearly one third of global investors are planning R&D-oriented FDI in Eastern Europe over the next three years [2]. The majority of investors choose R&D investment locations based on three attributes: lower R&D costs, availability and quality of local R&D labour and intellectual property protection. Other important attributes are the quality of universities and research centres as well as IT and local infrastructure. Małopolska is currently the leading R&D centre in Poland, but we should be aware that there are strong and in some areas more advanced competitors, such as the Czech Republic, in our part of Europe.

Important actions elevating growth and competitiveness of R&D in the region, in which MITC may play a key role, include also supporting commercialisation of research by establishing incubators and matchmaking of inventors with industrial partners.

As the objectives of the MITC as well as its planned activities and initiatives clearly coincide with the EU policy of building a knowledge society in Europe, they will be eligible for support from structural funds and European programs. Detailed recommendations of the European Commission may be found, for example, in its integrated innovation and research action plan [19]. Some of them such as putting research and innovation at the heart of business (by eg. *promotion of innovation poles and knowledge-driven and industrial clusters* by the European Commission and corresponding task for member states: *take full advantage of Structural Funds for the development of innovation poles and participate in EU cluster initiatives*) directly address supporting clustering initiatives. The Commission is proposing more use of Structural Funds to support activities that enhance the region's economic growth and competitiveness [20]. In practice the source of financial support of MITC's initiatives embodied in specific projects will be potentially found in national operational programmes and Małopolska Regional Operational Programme 2007-2013.

Another important source of financing activities of MITC in an international perspective will be the Seventh Frame-

work Programme for Research and Development 2007-2013. One of its proposed actions, Regions of Knowledge, aims to strengthen the research potential of the regions by bringing together regional authorities, universities, research centres, enterprises and other interested parties [20]. It is clear that funds for supporting research, development, innovation and competitiveness will be ample and it will depend on the regional actors how much of them will be used in Małopolska for ensuring its future growth.

The pie of global IT services is getting bigger and bigger. The role of regional stakeholders, and MITC in particular, is to secure a fair piece of it for Małopolska.

REFERENCES

1. Globalisation Act II: Team Europe Defends its Goals, Ernst&Young 2006. Retrieved July 24, 2006 at http://www.ey.com/global/content.nsf/Switzerland_E/HOME
2. FDI Confidence Index, Global Business Policy Council 2005, vol. 8, A.T. Kearney, Inc. Retrieved July 24, 2006 at http://www.atkearney.com/shared_res/pdf/FDICI_2005.pdf
3. Mapping Offshore Markets Update 2005, Offshore Insights Market Report Series, vol. 3, issue 8, neoIT 2005. Retrieved July 24, 2006 at http://www.neoit.com/pdfs/whitepapers/Olv3i08_0905_Mapping-Offshore-Markets.pdf
4. Computerworld Top 200. Retrieved July 24, 2006 at <http://www.computerworld.pl/top200/>
5. Mapping Offshore Markets Update 2004, Offshore Insights Market Report Series, vol. 2, issue 6, neoIT 2004. Retrieved July 24, 2006 at http://www.neoit.com/pdfs/whitepapers/Olv2i6_June_04_Mapping_Offshore_Markets.pdf
6. The truth about foreign direct investment in emerging markets, The McKinsey Quarterly, 2004 number 1, McKinsey & Company, Inc. Retrieved July 24, 2006 at http://www.mckinseyquarterly.com/article_print.aspx?L2=7&L3=10&ar=1386
7. Porozumienie o współpracy w ramach Małopolskiego Klastra Technologii Informacyjnych, Kraków 2006, unpublished
8. Strategia Rozwoju Województwa Małopolskiego 2007-2013, Kraków 2006. Retrieved July 24, 2006 at <http://um.wrotamalopolski.pl/RozwojRegionalny/Strategia2013/>
9. Regionalna Strategia Innowacji Województwa Małopolskiego 2005-2013, Kraków 2005. Retrieved July 24, 2006 at http://www.wrotamalopolski.pl/root_Przedsiębiorczosc/Przedsiębiorcza+Małopolska/Perspektywy/innowacja0905.htm
10. Clusters and the new economics of competition, Porter M., Harvard Business Review, November-December 1998
11. Regional Clusters in Europe, Observatory of European SMEs 2002, No. 3, European Communities 2002. Retrieved July 24, 2006 at <http://www.competitiveness.org/article/articleview/458/1/50>
12. It's 2008: Do You Know Where Your Talent Is?, A Deloitte Research Study, Deloitte Development LLC 2004. Retrieved July 24, 2006 at <http://deloitte.net/dtt/research/0,1015,sid%253D44426%2526cid%253D71444,00.html>
13. Sizing the emerging global market, The McKinsey Quarterly, 2005 number 3, McKinsey & Company, Inc. Retrieved July 24, 2006 at http://www.mckinseyquarterly.com/article_abstract_visitor.aspx?ar=1635&L2=1&L3=106
14. Marshall's Office of the Małopolska Region, 2006, unpublished
15. <http://www.sasinschool.com/>, retrieved July 24, 2006
16. http://www.diversitycareers.com/articles/pro/04-decjan/dia_motorola.htm, retrieved July 24, 2006
17. Offshore R&D Takes Off, by Bob Violino, CMP – CyberMedia LLC, 2005. Retrieved July 24, 2006 at <http://www.globalservicesmedia.com/sections/ito/showArticle.jhtml;jsessionid=TGYZHAZPOPS5W0QSNLRCCKH0CJUNN2JVN?articleID=180207315>
18. The Internationalization of Corporate R&D, Swedish Institute for Growth Policy Studies (ITPS), 2006. Retrieved July 24, 2006 at www.itps.se/Archive/Documents/Swedish/Publikationer/Rapporter/Allmänna/A2006/A2006_007.pdf
19. An action plan to boost research and innovation, The European Commission Memo/05/366, 2005. Retrieved July 24, 2006 at <http://europa.eu.int/rapid/pressReleasesAction.do?reference=MEMO/05/366&format=HTML&aged=1&language=EN&guiLanguage=fr>
19. Using Europe's regions to generate knowledge, The European Commission Memo/06/196, 2006. Retrieved July 24, 2006 at <http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/06/196&format=HTML&aged=0&language=EN&guiLanguage=en>

EDITORIAL BOARD

EDITOR-IN-CHIEF

Professor IRENA ROTERMAN-KONIECZNA
Medical College – Jagiellonian University, Krakow, Kopernika 17

HONORARY ADVISOR

Professor RYSZARD TADEUSIEWICZ
AGH – University of Science and Technology
Professor JAN TRĄBKA
Medical College – Jagiellonian University

MANAGING EDITORS

BIOCYBERNETICS – Professor PIOTR AUGUSTYNIAK
AGH – University of Science and Technology, Krakow, al. Mickiewicza 30

BIOLOGICAL DISCIPLINES – Professor LESZEK KONIECZNY
Medical College – Jagiellonian University, Krakow, Kopernika 17

MEDICINE – Professor KALINA KAWECKA-JASZCZ
Medical College – Jagiellonian University, Krakow, Kopernika 17

PHYSICS – Professor STANISŁAW MICEK
Faculty of Physics – Jagiellonian University, Krakow, Reymonta 4

MEDICAL INFORMATICS AND COMPUTER SCIENCE – Professor MAREK OGIELA
AGH – University of Science and Technology, Krakow, al. Mickiewicza 30

TELEMEDICINE – Professor ROBERT RUDOWSKI
Medical Academy, Warsaw, Banacha 1a

LAW (and contacts with business) – Dr SYBILLA STANISŁAWSKA-KLOC
Law Faculty – Jagiellonian University, Krakow, Kanonicza 4

ASSOCIATE EDITORS

Medical College – Jagiellonian University, Krakow, Kopernika 7e

EDITOR-IN-CHARGE – PIOTR WALECKI
E-LEARNING (project-related) – ANDRZEJ KONONOWICZ
E-LEARNING (general) – WIESŁAW PYRCZAK
DISCUSSION FORUMS – WOJCIECH LASOŃ
ENCRYPTION – KRZYSZTOF SARAPATA

TECHNICAL SUPPORT

Medical College – Jagiellonian University, Krakow, Kopernika 17

ZDZISŁAW WIŚNIEWSKI – in charge
WOJCIECH ZIAJKA
ANNA ZAREMBA-ŚMIETAŃSKA

© COPYRIGHT BY INDIVIDUAL AUTHORS AND MEDICAL COLLEGE – JAGIELLONIAN UNIVERSITY

ISSN 1895-9091 (print version)
ISSN 1896-530X (electronic version)

<http://www.bams.cm-uj.krakow.pl>

INSTRUCTION FOR AUTHORS

1. The original and one Photostat copy of the manuscript should be mailed to: Managing Editor Zdzislaw Wisniewski. Authors are strongly urged to submit a CD or floppies containing the manuscript in Word for PCs format along with the hard copies. Articles already published or those under consideration for publication in other journals or periodicals should not be submitted.
2. Manuscripts should be in English. Usage of correct language is the responsibility of the author.
3. Submission of manuscripts to us implies that the copyright of the entire material except those already copyrighted by other publications like tables, figures etc and those permission the author(s) have already obtained, belongs to the publisher.

PRESENTATION

4. The text may be prepared from a good quality laser printer or electronic typewriter. It is to be typed (one side only) on a regular sheet of paper A4 size preferred with single line spacing. New Times Roman font of 12 point size is recommended. Authors are requested to paste figures and tables (if possible) at appropriate places in the text. If this is not possible, they may be given separately, in which case their exact locations in the text may be marked by pencil.
5. The article should begin with a short ABSTRACT followed by a short INTRODUCTION. The rest of the article may be titled and arranged as per the wishes of the author.
6. REFERENCES should continue the last section of the article. In the text, references to other papers or books should be cited using consecutive numbers in parenthesis (e.g. [1,2]) and they should be listed numerically in the last section.

Examples:

I. Pyrczak W., Sarapata K.: Instructions for authors: references. *Bio-Algorithms and Med-Systems*, 1, 386, 2005

II. Tadeusiewicz R., Ogiela M.R.: *Medical Image Understanding Technology*. Springer Verlag, Heidelberg 2004

Authors are to make sure that each reference appearing in the text appears in the list of references at the end and vice versa.

7. TITLE PAGE AND FIRST PAGE – The title page of the manuscript, the name(s) of author(s) and institution(s) where the work has been carried out should be typed out on a separate sheet of paper.
8. Section headings (ABSTRACT, INTRODUCTION etc.) are to be typed in upper case letters and placed on a separate line
 1. Scientific names should be typed in italics or underlined
 2. Tables should be numbered consecutively using Arabic numerals.
 3. Illustrations should be limited to materials essential for the text. All figures should be numbered consecutively and should be submitted as sharp, high-quality prints. Figures that are to appear as a group should be photographed and mounted together. All figures and groups of figures should be trimmed at right angles and pasted directly on the typescript in the appropriate places. They should be of a size permitting photographic reduction by 25% together with text (not more than 15 cm across for page width and not higher than 20 cm). Legends to figures should appear directly below the respective figures.
 4. In line drawings, all lines should be of uniform thickness; letters and numbers should be of professional quality and proper dimensions, approximately 1mm high when reproduced.
 5. RUNNING TITLE – Please make sure to mention the running title (short title) of the article separately in the first page where the title, names of authors, their affiliations, etc. are given.
 6. PREFERRED LENGTH – The length of the manuscript may be limited to 20-25 A4 size pages (20 cm width – 26 cms height) through this is not binding.
 7. PERMISSION – Authors must obtain permission from other publishers if they use already published and copyrighted materials.
 8. COLOUR FIGURES – Cost of printing color figures will be charged to the author at the rate of U.S. \$ 125 per plate of four color pages.
 9. REPRINTS – The order for reprints should be placed at the time of submission of the article itself, so that it may be out for printing without loss of time, if found suitable for order for not less than 100 reprints is to be made.

MANUSCRIPT SUBMISSION:

1. Electronic – the www.bams.cm-uj.krakow.pl address is available for all authors to submit manuscript. The instruction is available at each step of submission process.
2. The CD with pdf or doc file with figures and tables incorporated into the manuscript body can be sent by email – address: bams@cm-uj.krakow.pl
3. The CD (as above) can be also sent by post. Address: *Bio-Algorithms and Med-Systems* , 31-501 Krakow, Kopernika 17 POLAND