



EJBI 2008

ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 4 (2008)

www.ejbi.eu



Aims and Scope

The European Journal for Biomedical Informatics reacts on the great European need to share the information in the multilingual and multicultural European area. The journal publishes peer-reviewed papers in English and other European languages simultaneously. This opens new possibilities for faster transfer of scientific-research pieces of knowledge to large international community of biomedical researchers, physicians, other health personnel and citizens.

The generally accepted translations of the English version of the paper to the official European languages and other European languages.

Cooperating journals

Methods of Information in Medicine
Lékař a technika

(language)

(English)
(Czech)

List of European languages

ISO 639-1 code

Albanian	sg
Armenian	hy
Azerbaijani	az
Belarusian	be
Bosnian	bs
Bulgarian	bg
Catalan	ca
Croatian	hr
Czech	cs
Danish	da
Dutch	nl
English	en
Estonian	et
Finnish	fi
French	fr
Georgian	ka
German	de
Greek	el
Hungarian	hu
Icelandic	is
Irish	ga
Italian	it
Kazakh	kk
Latvian	lv
Lithuanian	lt
Luxembourgish	lb
Macedonian	mk
Maltese	mt
Norwegian	no
Polish	pl
Portuguese	pt
Romanian, Moldavian, Moldovan	ro
Romansh	rm
Russian	ru
Serbian	sr
Slovak	sk
Slovenian	sl
Spanish	es
Swedish	sv
Turkish	tr
Ukrainian	uk



Editors and Management

Editor in Chief: Jana Zvárová, Czech Republic
Managing Editor: Petra Přečková, Czech Republic
Design PDF Version: Dana Vynikarová, Czech Republic
Sales and Marketing Manager: Libor Seidl, Czech Republic

Editorial Board

Ammenwerth, Elske (de)	Austria
Blobel, Bernd (de)	Germany
Bobrowski, Leon (pl)	Poland
Bureš, Vít (cs)	Czech Republic
Degoulet, Patrice (fr)	France
Dostálová, Taťjana (cs)	Czech Republic
Eryilmaz, Esat Nadir (tr)	Turkey
Hanzlíček, Petr (cs)	Czech Republic
Iversen, Irma (no)	Norway
Kern, Josipa (hr)	Croatia
Lukosevicius, Arunas (lt)	Lithuania
Mansmann, Ulrich (de)	Germany
Martin-Sanchez, Fernando (es)	Spain
Masic, Izet (bs)	Bosnia and Herzegovina
Mazura, Ivan (cs)	Czech Republic
McCullagh, Paul (en)	United Kingdom
Mihalas, George (ro)	Romania
Naszlady, Attila (hu)	Hungary
Nykänen, Pirkko (fi)	Finland
Paralič, Ján (sk)	Slovakia
Pisanelli, Domenico M. (it)	Italy
Sharp, Mary (ga)	Ireland
Sousa Pereira, Antonio (pt)	Portugal
Svačina, Štěpán	Czech Republic
Valenta, Zdeněk (cs)	Czech Republic
Vinarova, Jivka (bg)	Bulgaria
de Lusignan, Simon (en)	United Kingdom

Publisher

EuroMISE s.r.o.
 Paprskova 330/15
 CZ-14000 Praha 4
 Czech Republic
 EU VAT ID: Cz25666011

Office

EuroMISE s.r.o.
 Paprskova 330/15
 CZ-14000 Praha 4
 Czech Republic

Contact

Karel Zvára
 zvara@euromise.com,
 Tel: +420 226 228 904
 Fax: +420 241 712 990

Instructions to Authors

MANUSCRIPT SUBMISSIONS

European Journal of Biomedical Informatics (EJBI) is an international, peer-reviewed journal that publishes papers in the broad field of biomedical informatics. Manuscripts accepted for the electronic publication in EJBI are original contributions, reviews, brief reports, special communications, commentaries, and many other categories of papers. Due to special multilingual features of EJBI, these types of papers are published in English, but national language versions can be published simultaneously (see below).

Manuscripts should be sent electronically to the e-mail address: manuscripts@ejbi.org. Each manuscript should be submitted in MS Word, plain-text, HTML, TeX or LaTeX format and accompanied by image attachments and

- Identification Form: the name of the corresponding author with his contact address, phone number, fax number and e-mail address.
- Submission Requirement: the statement that the submitted paper has not been published in, nor has been submitted to, any other journal.
- Authorship Form: the covering letter signed by corresponding author, by which the author approves publication of the paper in the European Journal for Biomedical Informatics on behalf of all authors.
- Transfer of Copyright: All authors of the manuscript must have agreed to its publication and are responsible for its content and must also have agreed that the corresponding author has the authority to act on their behalf in all matters pertaining to publication of the manuscript. The corresponding author is responsible for informing the co-authors of the manuscript status throughout the submission, review, and publication process.

Submission Requirement, Authorship Form and Transfer of Copyright should be sent by fax +420 241 712 990 and by the surface mail to EuroMISE s.r.o., EJBI Editorial Office, Paprskova 330/15, 140 00 Prague 4, Czech Republic simultaneously.

Authors may supply a PDF file of the manuscript additionally. Submission of the PDF file only is not sufficient.

PUBLISHER

The European Journal of Biomedical Informatics is published by EuroMISE s.r.o., Prague, Czech Republic (tax identification number: CZ25666011). Papers published in the Internet version of EJBI may be republished in printed digest of EJBI.

LANGUAGE

All manuscripts should be written in English in an easily readable style. English version of the manuscript will be peer-reviewed. However, it is possible to submit another language version of the same paper additionally. The Internet version of the EJBI makes available all language versions submitted. It is also possible to publish the English version of the paper that has been already

published in the national journal, when the agreement of national publisher is obtained (copyright agreement). The English version follows a standard peer-review procedure.

ELECTRONIC FORM OF A MANUSCRIPT

Recommended formats of text files are: .DOC (Word version 6.0, 7.0, 97, 2000 or 2003), .RTF (Rich Text Format). The name of the text file should be identical with the first author's surname, e.g. NOVAK.DOC. Names of attached images, drawings or tables (i.e. figures and tables of the manuscript) should be composed of author's surname and figure or table number, e.g. NOVAKTAB3.BMP. Manuscripts along with attached files should be sent by e-mail to address manuscripts@ejbi.org.

ELECTRONIC IMAGE DOCUMENTATION

Figures and tables (bitmaps) should be submitted in BMP, GIF, TIF, PNG, EPS or JPEG format. JPEG files should have best compressive rate (10-20) for a good quality. All images (although included in the manuscript file) must be also submitted as separate files because the resolution of embedded images is often insufficient. Scanned images should have the resolution of at least 600x600dpi. Submissions that do not meet the Instructions for Authors will be returned.

ABBREVIATIONS AND NOMENCLATURE

Generally known abbreviations do not need to be explained. Abbreviations for symbols and expressions for terms should be spelled and they should be comprehensibly explained in brackets. Every such abbreviation should be explained only after its first occurrence in the text.

Measurements of length, height, weight and volume should be reported in metric units or their decimal multiples. All haematological and clinical chemistry measurements should be reported in the metric system in term of International System of Units (SI). Chemical substances should be described by their own systematic name or expression, medicines by general names. Commercial names of chemicals, medicines or technical innovations may be used after they have been defined by their scientific names. Radionuclides (radioisotopes) will be symbolised by the atomic number.

REVISED MANUSCRIPT SUBMISSION

When revision of a manuscript is requested, authors should return the revised version of their manuscript as soon as possible. The prompt action may ensure fast publication if a paper is finally accepted for publication in EJBI.

FINAL PROOFREADING

The Publisher will send the accepted paper to its author for final proofreading in the PDF format. The author may then correct printing errors only. No other changes or additions will be accepted. Author should send corrected and signed paper back to the Publisher by fax to +420 241 712 990 or by surface mail to: EuroMISE s.r.o., EJBI Editorial Office, Paprskova 330/15, 140 00 Prague, Czech Republic.

ORGANIZATION OF THE MANUSCRIPT

Title page. The first (title) page should contain the title of the paper, names and workplaces of all authors. Individual workplaces are necessary to be graphically differentiated (preferably by numeral as the upper index).

Abstracts and keywords. At the beginning the author puts an abstract and keywords. The abstract should be in the extent of 250-300 words. There should be 4 to 7 keywords, according to author's consideration, preferably from MeSH index.

Main text of the paper. General rules for writing manuscripts recommend use of simple and declarative sentences; avoid long sentences, in which meaning may be lost by complicated construction. All acronyms and abbreviations should be explained when they first appear in the text. The main text of the paper should follow the style of selected type of paper.

Acknowledgement. Acknowledgements, if any, should be given at the end of the paper, before bibliographic references.

References. References should be cited in the text by their index number according to the order of appearance in the manuscript. Each reference should be marked by its index number in square bracket corresponding to bibliography section. It is possible to include references to dissertation works and technical reports. It is obligatory to include information sufficient to look up referenced text.

Examples of references in bibliography section:

[1] Knaup P., Ammenwerth E., Brandner R., Brigl B., Fischer G., Garde S., Lang E., Pilgram R., Ruderich F., Singer R., Wolff A. C., Haux R., Kulikowski C.: Towards Clinical Bioinformatics: Advancing Genomic Medicine with Informatics Methods and Tools. *Methods Inf Med* 2004; 43, pp. 302-307

[2] Blobel B., Pharow P.: A Model-Driven Approach for the German Health Telematics Architectural Framework and the Related Security Infrastructure. In: *Connecting Medical Informatics and Bio-Informatics. Proceedings of MIE2005* (Eds. R. Engelbrecht, A. Geissbuhler, C. Lovis, G. Mihalas), Vol. 116, Amsterdam, IOS Press, 2005, pp. 391-396

[3] <http://www.infobiomed.org/>

Tables and Figures. Authors should use tables only to achieve concise presentation, or where the information cannot be given satisfactory in another way. Tables should be numbered consecutively using Arabic numerals and should be referred to in the text by numbers. Each table should have an explanatory caption that should be as concise as possible. Figures should be clear, easy to read and of a good quality. Styles and fonts should match those in the main body of the paper. All figures must be mentioned in the text in consecutive order and should be numbered with Arabic numerals.

Authors should indicate precisely in the main text where tables and figures should be inserted, if these elements are given only separately or at the end in the original version of the manuscript.

Content

English version

- en 1 Editorial
- en 2 - 6 The Use of Object-Oriented Technologies for Medical Data Storing and Retrieving
Vita Speckauskiene, Arunas Lukosevicius
- en 7 - 11 Wavelet Contour Based Detector for Arrhythmia Analysis
Dina Kičmerová, Ivo Provazník
- en 12 - 16 Coronary Heart Disease Mortality in Czech Men, 1980-2004
Jindra Reissigová, Marie Tomečková
- en 17 - 23 Guideline Knowledge Representation Model (GLIKREM)
David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský

Czech version

- cs 1 Předmluva
- cs 2 - 6 Konturový vlnkový detektor pro analýzu arytmií
Dina Kičmerová, Ivo Provazník
- cs 7 - 10 Ischemická choroba srdeční u mužů v České republice, 1980-2004
Jindra Reissigová, Marie Tomečková
- cs 11 -17 Model reprezentace znalostí v doporučeních
David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský

Lithuanian version

- lt 2 - 7 Objektinių Technologijų Naudojimas Medicininių Duomenų Saugojimui Ir Paieškai
Vita Speckauskiene, Arunas Lukosevicius



EJBI 2008 ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 4 (2008), Issue 1

English version

Cooperating journal – Methods of Information in Medicine
Editor-in-Chief: R. Haux (Germany)
Publisher: Schattauer
www.methods-online.com

www.ejbi.eu



Content

English version

- | | |
|------------|---|
| en 1 | Editorial |
| en 2 - 6 | The Use of Object-Oriented Technologies for Medical Data Storing and Retrieving
Vita Speckauskiene, Arunas Lukosevicius |
| en 7 - 11 | Wavelet Contour Based Detector for Arrhythmia Analysis
Dina Kičmerová, Ivo Provazník |
| en 12 - 16 | Coronary Heart Disease Mortality in Czech Men, 1980-2004
Jindra Reissigová, Marie Tomečková |
| en 17 - 23 | Guideline Knowledge Representation Model (GLIKREM)
David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský |

Editorial

Jana Zvárová
Editor-in-Chief

Biomedical informatics is a burgeoning field, with important applications and implications throughout the biomedical world and healthcare delivery. The European Journal of Biomedical Informatics (EJBI) is reacting on the great European need to share the information in the multilingual and multicultural European area.

EJBI opens for the field of biomedical informatics a new model of electronic publishing. EJBI is publishing accepted peer-reviewed papers in English and other languages simultaneously. This opens new possibilities for faster transfer of

scientific-research pieces of knowledge of many European countries to a large international community of biomedical researchers, physicians, other health personnel and citizens. Moreover, the journal now enables to make results of scientific-research work and practical experiences of foreign specialists accessible to wider health public in a more comprehensible way in each European country.

The aim of the editorial board is to reach the highest scientific level of the journal and show the best practices of biomedical informatics applications to wide

readership. The European editorial board is composed from outstanding specialists in the field of biomedical informatics. We believe that their activities for EJBI will contribute to propagation of journal's good credit. The editorial board also presumes that presentation of English versions of scientific papers with their professional translations to other languages will significantly contribute to unification of applied scientific terminology.

EJBI is now a Schattauer related journal. You may find more information in Editorial of Methods of Information in Medicine, Issue 2 2008.

The Use of Object-Oriented Technologies for Medical Data Storing and Retrieving

Vita Speckauskiene, Arunas Lukosevicius

Biomedical Engineering Institute at Kaunas University of Technology, Lithuania

Summary: The article focuses on object oriented technologies and the usage of these technologies, how to create an information system (IS) for applications in medicine, specifically in ophthalmology. The rapid evolution of medical data standards, concepts and technologies has created a problem – how to develop a system that would be flexible enough to work with evolving data. After the analysis of object-oriented technologies: Java 2 Platform, Enterprise Edition (J2EE) and Microsoft .NET and open source object-oriented data base management systems (OODBMS) we have proposed and created a system based on .NET and OODBMS db4o. Results include the developed system with proposed architecture and used techniques, also the advantages of object-oriented data bases (OODB), the systems' possible interoperability with the National Electronic Health Record system. Developed IS should help medical specialists to reach important information from any computer in any hospital.

Keywords: object-oriented technology, object-oriented data base, information system, data storing standards

Introduction

Information systems (IS) currently used in the medical field are one of the most evolving components for e-health. The main features of IS in medicine are 1) storage of multimedia; 2) information connections in complicated and evolving semantic relations; 3) big amounts of information involving all lifelong patients' records; 4) wide list of users, spanning between all three health care levels (they create a need of secure and convenient communication between these levels and IS); 5) compatibility of IS with the European Union (EU) [1] and international standards [2, 6]; and 6) usage of data for clinical decision support.

These factors raise specific requirements for IS and their architectures. Therefore, a search for new technologies in order to develop such systems is in demand [3, 4, 5, 7, 16, 17]. A number of standards has been created to store and transmit all data of Electronic Health Records (EHR) including medical images. The problem is how to make the best benefit of this data for users. The data has to be stored in such a way that it would be easy to generalize and retrieve it. Therefore, standards are under development for interchanging of medical data. Each country has its own way of solving this problem, so one will stumble over a number of different standards describing various IS and data storing aspects: ASTM, ASC X12, IEEE/MEDIX, NCPDP, OpenEHR, HL7 (Health Level 7) [6], DICOM [2] and others. An organization developing a standard usually specializes in one particular area. Unfortunately, those standards optimize and facilitate the work with medical data, but data processing and storage is still a technological problem. The technologies that are used for a long time are not flexible enough to work with such data. Medical data is stored in relational data bases (RDB); though studies [18, 19] show that object-oriented data bases (OODB) bring a lot of new advantages. Also the need of secure and flexible Internet forces to look for effective programming tools and technologies.

The aim of this article is to show the advantages of OODBMS, also to present the architecture and develop a system for medicine which in certain extent satisfies the above mentioned features. To obtain this goal, object-oriented data bases, compatible programming techniques, data storing methods, and main features of IS for medicine are analyzed and applied. Also, this system is being prepared for integration into the National EHR System (Lithuanian abbreviation NESS), which is now being developed (at the national

level). Unfortunately, we are not able to show the interoperability between NESS and the developed system, because NESS is not yet implemented. But still we have made a few observations.

Methods

The analysis of object-oriented data bases OODBMS is a new quickly developing technology which is increasingly applied by many different companies. The benefits of OODBMS bring a rise in users as well as demands. These DB have distinctive data storing methods. The data is stored in tables, which can have tables in themselves – they form a hierarchy. The tables are filed not only with concrete data, but also with objects. This technology can be implemented effectively for medical data storing. Because of the evolution of medical requirements OODBMS makes the systems more flexible and stable [18, 19]. It also enables interaction between objects in the data base.

RDBMS are convenient in storing text and numerical data; but they cannot work with undefined records which are formed in object-oriented programming languages. Moreover, these data base management systems (DBMS) are not fitted for storing methods. IS requires to work not only with a text, but also with spreadsheets, images, diagrams, sounds and other multimedia. Besides, all this information has to be organized as a united structure, handy for users. The better decision might be to implement OODBMS. We will see that this technology is able to vouch a new approach of data storing and mining.

OODBMS stores composite objects which compound out of other objects. These objects are stored with references to lower level objects. The number of levels is unlimited. Each object can be brought into many other objects as components. This forms a net structure data model. OODBMS allows construction of new data

types; this feature enables to store multimedia: images, audio and video files. There are two types of OODB: passive and active. Passive OODB save the object structure but do not control behaviour of objects. While active OODB allow communication between objects within the DB.

A lot of companies offer the products of OODBMS, but most of them are payable. The most known are listed in Table 1 [8].

Table 1. Commercial OODBMS.

OODBMS	Vendor
Versant Developer Suite and FastObjects	Versant Corporation
TITANIUM	Micro Data Base Systems, Inc.
GemStone/S and GemStone Facets	GemStone Systems, Inc.
Cache'	InterSystems Corporation
JADE	Jade Software Corporation
JYD Object Database	JYD Software Engineering Pty Ltd.
VOSS	Logic Arts, Ltd.
Objectivity/DB	Objectivity, Inc.
ObjectStore Enterprise and PSE Pro	Progress Software Corporation

platforms used are J2EE and .NET. The companies that offer these products are one of the biggest competitors. So, the developer has a dilemma – which one to choose. Sun Microsystems offers to use Java 2 Platform, Enterprise Edition (J2EE) and Microsoft – .NET technology. Both technologies allow implementation of web services. The main advantage of J2EE is the fact that it is a multiplatform – it is possible to use J2EE in different operating systems (OS). Microsoft .NET works only in Windows OS, although the Microsoft company offers an equivalent tool for Linux OS. Both technologies use object-oriented programming languages (.NET – Visual Basic and others, J2EE – JAVA). The advantage of .NET is its capability to render pages in various HTML formats, which abstracts developers from version-specific HTML.

Servlets in Java can achieve the same task, though with more manual coding [7]. Another difference and advantage of .NET is its multilanguage; it works with Visual Basic, C#, FORTRAN and COBOL; J2EE works in Java programming language. The main features of both technologies are presented in Table 3 [3, 4, 16].

As we can see both technologies are very similar. Most important is that both of them implement web services and are able to map data on the Internet. A .NET Web service could serve as a consumer to a J2EE Web service and vice versa because of the same Web services standards used [7]. Moreover, web services can be reused no matter what programming language they were coded in. Most authors do not point out particular reasons why one should choose either technology [5, 16, 17]. Still we can note some advantages of .NET in its development and handling capabilities. That's why this technology was selected for implementation.

Table 2. Open source OODBMS.

OODBMS	Programming environment	Limitations
db4o	J2EE .NET	Handling of default and null values (noticed during implementation).
SOD2	J2EE C++	Does not have a query mechanism.
ozone	J2EE	Implemented in JAVA and works only in the J2EE environment. Does not have a query mechanism.
ObjectDB	J2EE	Implemented in JAVA and works only in the J2EE environment. Free DB edition does not support a client-server architecture (unless the DBMS is in the same server).
GigaBASE	J2EE C++	Does not support null values.
XL2	J2EE	Does not have data recovery functions. Does not have a query mechanism.
Orient	J2EE, C++	Available for C++ (ODMG 3.0 interface) and Java with the support of Sun JDO 1.0 technology.

System architecture and functionality IS for medicine data mining could be implemented using both above analyzed technologies. No matter which object-oriented technology is chosen, the system architecture is analogical, it is presented in Figure 1.

Still there is a part (and much bigger) of software developers who propose free products. In this article we have selected to review seven open source OODBMS; they are listed in Table 2 [10, 11, 12, 13, 14, 15]. We can see that open source data bases are implemented for J2EE, and not all of them correspond to the standards, as well as overall OODB flexibility. The standards of OODB are handled by the Object Database Management Group (ODMG) [9]. Since (at the analysis stage) we did not penetrate any limitations of OODBMS db4o, we chose it for the development of the final result.

Comparison between J2EE and .NET technologies

Web service technologies are very widely used because they provide to programmers a lot of capabilities and their implementation is not difficult. The main

Table 3. Comparison of main features of J2EE and .NET.

Feature	J2EE	.NET
Relational DB Access	Java Database Connectivity (JDBC) or L/J	Active Data Objects (ADO.NET)
Web Client	JAVA server pages (JSP) and servlets	ASP.NET
Graphical user interface (GUI)	AWT/Swing	Windows forms
Messaging	Java Messaging Service (JMS 1.0)	Microsoft Messaging Queuing (MSMQ)
Web Services Support	Java Web Services Developer Pack (JWS DP)	Built directly into .NET and Visual Studio
Infrastructure	Enterprise Java Beans (EJB)	COM+
Naming and Directory Service	Java Naming and Directory Interface (JNDI)	Active Directory Services Interface (ADSI)
Database storage	-	SQLServer (the official .NET platform database technology)

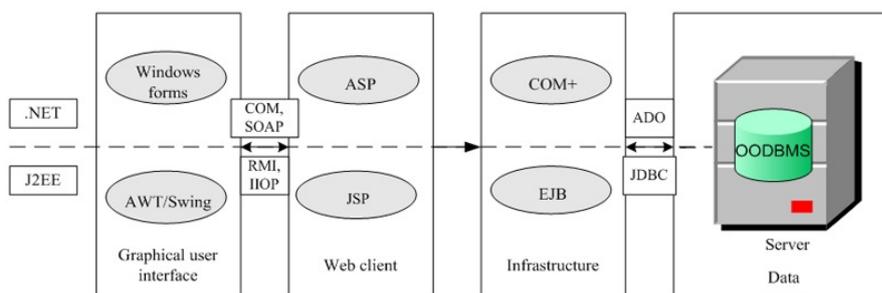


Figure 1. Architecture of medical data IS.

The advantage of similarity of those technologies is that both systems can communicate with the same DBMS, if there is a necessity to communicate. As mentioned in the above sections, for implementation we chose OODBMS db4o [10] and .NET technology to create a medical data storing, retrieving and mining system. In Figure 2 there is presented the architecture of the system we have developed.

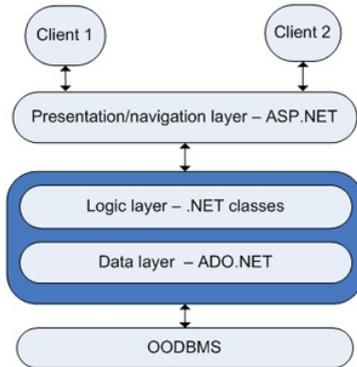


Figure 2. Architecture of implemented IS.

As we can see the system consists of three layers: presentation, logic and data. The communication between presentation and data layers is achieved through the logic layer. In Figure 3 there is presented a use case diagram of implemented IS.

The use case diagram shows the functionality of the system (the functions that a user and administrator can perform).

Results

The system has been developed using the above mentioned technologies. The integration between .NET and db4o did not cause any bigger problems. The communication among the application and the data base itself is achieved through the standard data base query mechanism, included in the standard database libraries. The examples of the source code, showing OODB queries and programming capabilities, used in the developed system, are presented in Table 4.

The data base is a single .yap type file; a blob type file can be any file with no limitation in size or file extension (if desired). So data and file storing and updating mechanisms are very flexible and convenient for a software engineer.

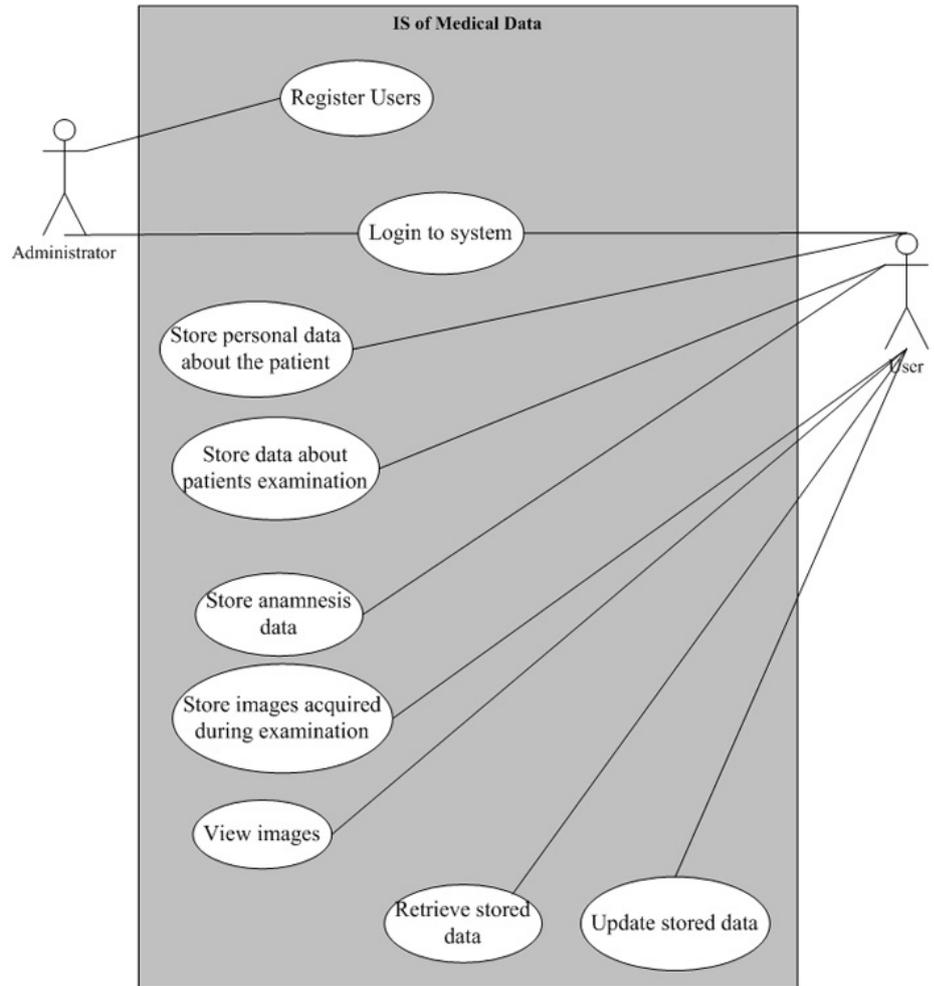


Figure 3. Use case diagram.

Table 4. Parts of the source code of OODB queries and programming capabilities used in the developed system.

<code>using com.db4o;</code> <code>using com.db4o.query;</code>	Standard libraries (mainly used)
Connecting to the data base and making a query	
<code>ObjectServer db4oServer = Db4o.openServer("db.yap", 4488);</code> <code>db4oServer.grantAccess("...");</code> <code>ObjectContainer client = Db4o.openClient("...");</code> <code>Query query1 = client.query();</code> <code>query1.constrain(typeof(Vartotinfo));</code> <code>query1.descend("_vartotojas").descend("_pavarde").constrain(p.Text);</code> <code>ObjectSet result1 = query1.execute();</code> <code>client.close();</code> <code>db4oServer.close();</code>	Connection to the server Create server client Make a query Set the class to look into Find a particular data Execute query Closing the client Closing the server
Setting and updating new data	
<code>Vartotojas vartotojasn = new Vartotojas(P1.Value,p.Text,'N');</code> <code>client.set(vartotojasn);</code> <code>Vartotinfo vartotinfo = new Vartotinfo(v.Text,el.Text,g.Text,k.Text,s.Text);</code> <code>vartotinfo.Vartotojas = vartotojasn;</code> <code>client.set(vartotinfo);</code>	Generating new data for class "Vartotojas" Storing the data Generating new data for class "Vartotinfo" Creating a reference Storing the data
Working with "blob" type files	
<code>Db4o.configure().setBlobPath("blobs");</code> <code>Byla snap = new Byla("...");</code> <code>client.set(snap);</code> <code>snap.Bylapav.readFrom(new File(File1.Value));</code> <code>double status = snap.Bylapav.getStatus();</code> <code>while(status > Status.COMPLETED)</code> <code>{</code> <code> status = snap.Bylapav.getStatus();</code> <code>}</code> <code>Byla found = (Byla)result.next();</code> <code>string pav = found.Bylapav.GetFileName();</code>	Setting the blob path Generating new data for class "Byla" Storing the data (also the text data about the file: the file name, extension and size) Preparing to store the blob type file itself Storing the file itself Querying for the file Retrieving the file

Testing of the system: method.

After the development we tested the system and made an experiment with various system changes. The purpose of the experiment was to show the flexibility of used technologies in making changes in all three system layers – data base, user interface and logical layer. To accomplish this task, we made an experiment in which we dealt with an evolving data problem. We created a scenario in which the functional requirements are complemented. The new data containing parameters of images are to be added to the system. First of all we tested the systems, then made changes according to the scenario and then again tested the system including the changes we made. The following four sections describe the testing procedure, changes in data base, presentation and logic layers.

The system was tested using the data that was collected during patients' examinations in Clinics of Kaunas University of Medicine. We used a limited amount of data, because our main task was to see how the data behaves in OODBMS, and not to measure any rates yet. We tested user's functions (e.g. store/retrieve patient information, analyze patient's image), administrator's functions (e.g. user registration), and user interface (e.g. navigation links, data mining results). All the procedures were performed successfully, as well as the overall testing.

Changes in the data base.

To show the flexibility of OODBMS a similar RDB model (to OODB classes) was created. The changes were compared using visual diagrams. A simplified example of one part of the visual diagram is shown in Figure 4.

The figure shows that there is less information stored in OODB; there is no need to store identifiers because the connection between classes is made by an operation ("EyeImageData()") in Figure 4 a). The field "Filename" (seen in Figure 4 b)) is needless in OODB because the blob type (any multimedia or other type) file information is generated automatically in the DB engine. So there is no need to create extra fields between tables, and the data is not doubled, also the multimedia data management is more convenient.

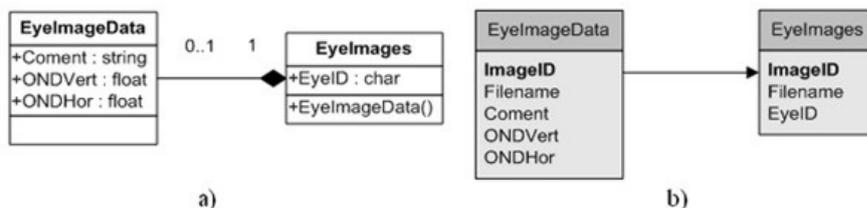


Figure 4. Data base models: a) OODB b) RDB.



Figure 5. Changes in user interface.

Changes in user interface.

The changes in the presentation layer are equivalent in both (OODB and RDB) cases. Figure 5 shows the changes in the created system with OODB.

Changes in the logic layer.

The logic layer is more complicated. For both cases it consists of the following changes:

- programming code for new user interface elements,
- performing test for these elements using the remaining data,
- storing all the remaining data.

We will not go into detail how the OODBMS behaves in the case of these changes, but using the OODBMS db4o data management software „Object Manager”, we could see that DB structures are updated automatically, when the DB is activated. For the data store and update in the case showed in Figure 6 we need the same source code – only one statement is needed. However, in the RDB case we

need two SQL sentences (Insert Into and Update).

It obviously reduces the amount of the programming code. We can see this form in Figure 6, where the reference and new data are presented.

Also note that the reference (in the case showed above) spreads the new data, and avoids data duplication.

There should not be a problem to integrate the system into NESS, because the Integrating Healthcare Enterprise (IHE) is used for this purpose in NESS. Since db4o is a single file convertible to XML, which together with HL7 CDA Level 2 are applied in NESS, there should not be any trouble for systems to communicate.

Further observations and advantages of OODBMS are presented in the next section.

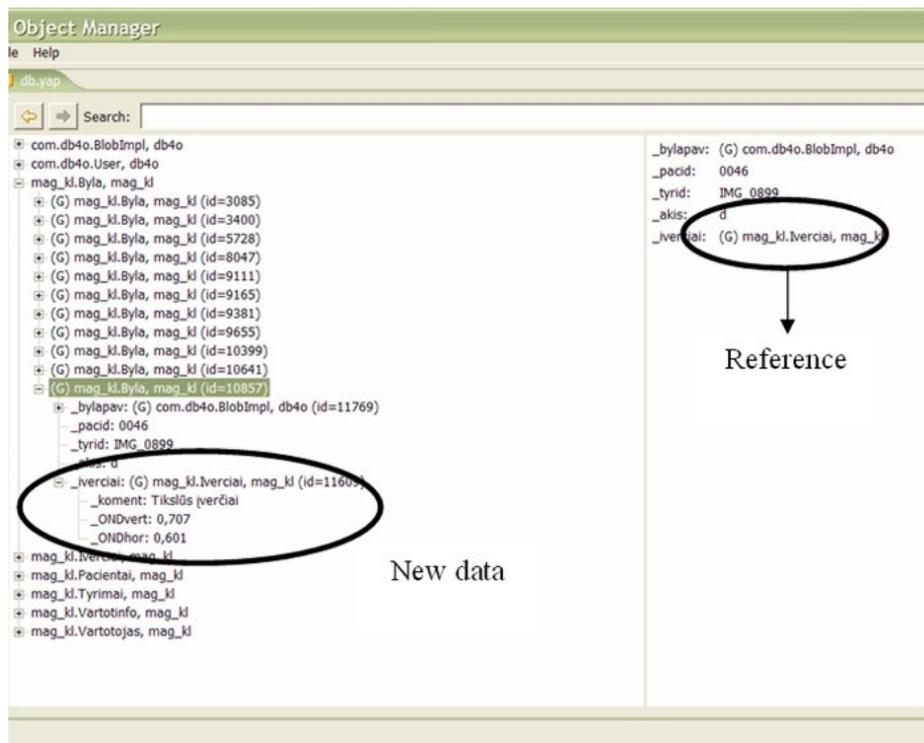


Figure 6. OODB reference and the new data.

Discussion and conclusions

Extensive international and collaborative work with standards and concepts gradually leads to the better quality of developed products, as well as software – the systems are able to work with different data types, to ensure compatibility and flexibility. In future we intend to integrate the system into NESS.

The analysis of OODBMS shows that they are very convenient for storing of semantically and quantitatively evolving medical data. It is possible to implement remote DB access using object-oriented technologies easily.

The developed IS had shown that such a system is flexible and evolutionistic because of the usage of appropriate technologies which are at the same time compatible with conceptual medical informatics standards. Although the system could be implemented by using different technologies, we prefer using .NET technology because of its development and handling capabilities, also the multilanguage ability and the user-friendly environment of Visual Studio .NET software.

Our experiment showed that OODMS has the following advantages:

- possibility to store and retrieve data with the same statement;
- absence of necessity to create additional fields between tables, because of the refeces OODBMS uses;
- absence of duplicated data;
- more flexible storage of multimedia data;
- absence of creating tables and links in the data level in the DB changing process.

For this purpose we used a small amount of data. However, we intend to store a large amount of data and rate systems behaviour.

It is rational, however, to mention that results presented do not imply refusal of RDBMS usage, rather we propose OODBMS as a flexible solution for solving specific problems related with inevitable and continuous evolution of eHealth systems.

References

- [1] CEN/TC 251 Standard Development Workgroup site, <http://www.centc251.org/>
- [2] DICOM standard site, <http://medical.nema.org/>

- [3] Estes B. T., Maxime O.: J2EE vs .Net: The Choice Depends On Your Needs. 2003; <http://www.computerworld.com/printthis/2003/0,4814,84155,00.html>
- [4] Farley J.: Microsoft .NET vs. J2EE: How Do They Stack Up?. 2000; http://java.oreilly.com/news/farley_0800.html
- [5] Hanson J. J.: .NET versus J2EE Web Services. A Comparison of Approaches. <http://www.webservicesarchitect.com/content/articles/hanson01.asp>
- [6] HL7 Standards Developing Organization site, <http://www.hl7.org/>
- [7] Lurie J., Belanger R. J.: Does One Web Services Platform Dominate the Other? 2002; <http://www.javaworld.com/javaworld/jw-03-2002/jw-0308-j2eenet.html>
- [8] Object-oriented database product vendors, http://www.service-architecture.com/products/object-oriented_databases.html
- [9] Object Database Management Group site, <http://www.odmg.org>
- [10] OODBMS db4o site, <http://www.db4o.com/>
- [11] OODBMS ObjectDB site, <http://www.objectdb.com/>
- [12] OODBMS ozone site, <http://ozone-db.org/frames/home/what.html>
- [13] OODBMS SOD2 site, <http://www.tneumann.de/sod/index.html>
- [14] OODBMS XL2 site, <http://www.xl2.net/>
- [15] O-RDB GigaBASE site, <http://www.garret.ru/~knizhnik/gigabase.html>
- [16] Sessions R.: Java 2 Enterprise Edition (J2EE) versus The .NET Platform Two Visions for eBusiness. 2001; <http://www.objectwatch.com/FinalJ2EEandDotNet.doc>
- [17] Sheil H., Monteiro M.: Rumble in the Jungle: J2EE versus .Net. How do J2EE and Microsoft's .Net compare in enterprise environments?. 2002; http://www.javaworld.com/javaworld/jw-06-2002/jw-0628-j2eevsnet_p.html
- [18] Sorli M., Romo J., Stach R., Bredehorst B.: REMOTE: GRD1-2000-25433. 2001, p. 40-42; www.remoteproject.org/downloads/REMOT E.D1.11.1--2001-07-09.pdf
- [19] Zamir S.: Handbook of Object Technology. CRC Press, 1999, 35ch.

Contact

Vita Speckauskiene

Biomedical Engineering Institute at
Kaunas University of Technology
Studentu str. 65 Kaunas, LT-51359
Lithuania
e-mail: vita.speckauskiene@med.kmu.lt

Wavelet Contour Based Detector for Arrhythmia Analysis

Dina Kičmerová, Ivo Provazník

Department of Biomedical Engineering, Brno University of Technology, Czech Republic

Summary: This paper deals with design of an automatic detector for classification of selected cardiac arrhythmias. The proposed algorithms employ the continuous wavelet transform (CWT) combined with an analysis of its contour envelopes. The CWT was used in a detector of R-waves, to distinguish between normal and abnormal beats, and for detection of atrial premature contractions (APCs) and premature ventricular contractions (PVCs). The algorithm was validated by extensive testing on the MIT/BIH database. Searching for a local maximum in wavelet contour envelopes efficiently detects R-peaks. The overall accuracy of its detection tested on 48 half-hour signals is 99.5%. Two types of classifications were tested: 1. classification based on the contour envelope and the detection of significant points with overall accuracy 94.6%, 96.1% for the sinus rhythm (SR), 30.4% APCs, 71.2% PVCs and 2. the localization of maximum of square modulus of CWT coefficients in the area of QRS complex for the determination of PVCs between SR, right bundle branch block (RBBB), APC and other narrow complex arrhythmias with the accuracy 96.8%.

Keywords: arrhythmias, ECG signal, wavelet transform, contours

Introduction

Automatic processing systems are frequently used in medical data analyses. Modern medicine generates huge amounts of data giving an opportunity to search for explicit understanding of physiology. New methods can help in dealing with this problem, they can simplify and usually speed up the processing of large volumes of data. The physician's work is a typical example. The physician has to decide very frequently upon a patient's diagnosis on the basis of a number of numerical values measured during an examination. Orientation in this amount of data is not always easy and

unambiguous. Therefore there are many consultation systems that help minimise human errors.

The modern analysis of the electrical activity of the heart uses simple and sophisticated algorithms of digital processing signals implemented in electrocardiographs [1,2,4-6,8-18]. We can sort these algorithms into three groups: working in time [8], frequency [5] and time-frequency domain [10-18]. First two classes belong to classical methods, which are successful in many clinical tested applications and analysis of ECG, detection of the QRS complex [4], its beginning and end, analysis of deviation of the ST segment, analysis of variability of heart rate, etc. New algorithms work in the time-frequency domain [10-18] and combine some advantageous characteristics known from classical methods – mainly they allow the frequency analysis with time information about analyzed features. Using the simple time analysis does not provide high quality results because of low sensitivity. It is caused by a small amplitude of the changes. The frequency analysis increases sensitivity, but it cannot tell in which phase of the heart cycle these changes originated.

At present, the wavelet transform (WT) is the most used method of the time-frequency analysis [10-18]. WT is popular, because it is implemented simply and its results are very well interpreted like the Fourier transform, using the frequency analysis. Many variants of WT provide large possibilities in selection, from a redundant detailed analysis to a fast analysis with minimum output data. Selection of a particular variant of the WT is based on the specific application (suppression of the noise, marking time and frequency components of useful signal, detection of important points, detection and observing important changes in waves in time, etc.).

In this paper we explore the use of the continuous wavelet transform in the analysis of beat morphologies and for ECG classification. The aim of this work was to distinguish different ECG arrhythmic patterns and enable an abnormality diagnosis. An arrhythmia classifier based on the continuous wavelet transform has been presented to identify normal beats, abnormal premature ventricular contraction (PVC) and atrial premature contraction in ECG. The occurrence of an arrhythmia is unpredictable. The purpose of this study is to develop a method using the CWT coefficients of ECG waveforms. Two kinds of cardiac arrhythmia, PVC (premature ventricular contractions) and APC (atrial premature contractions), which are the most common types of cardiac arrhythmias in ECG monitoring, will be discussed.

Methods

The wavelet transform (WT) is based on the use of a set of mathematical analyzing functions called "wavelets". Wavelets provide decomposition of the ECG signal into a set of wavelet coefficients. The analyzing functions are generated from a generally complex-valued function $\psi(t)$ by dilation and shift in time. Each analysing function $\psi_{a,b}(t) = \psi((t-b)/a)$ has its own parameters – time localization defined by shift b and a frequency band defined by a dilation factor a . Each resulting wavelet coefficient corresponds to the measurement in the signal in a given time instant and a given frequency band. Wavelet coefficients can be easily calculated using the following definition equation of the continuous wavelet transform (CWT)

$$WT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left(\frac{t-b}{a} \right) f(t) dt \quad (1)$$

where $f(t)$ is an ECG signal to be analysed and $*$ stands for complex conjugate.

Further background information concerning continuous wavelets and their properties can be found in [3, 6] and [16].

ECG components delineation

The proposed method is composed of two main parts: 1. continuous wavelet transform with discretized output, and 2. classical time-domain processing algorithms.

WT (Eq. (1)) results in a two-dimensional function which represents the signal in the time-frequency domain. For an easier analysis of the resulting function $WT(a,b)$, we may use its absolute value $abs(WT(a,b))$. Further, the function should be normalized for a later analysis. Thus, Eq. 1 can be rewritten as

$$WT_{abs}(a,b) = abs(WT(a,b)) / \max_{a,b} [abs(WT(a,b))] \quad (2).$$

An example of the ECG signal analysis using CWT with the Morlet wavelet is shown in Fig.1. Three main waves/peaks can be found in the signal (panel (a): small P-wave marked by (1), high QRS-complex marked by (2) and high T-wave marked by (3)). The time-frequency spectrum of the signal is shown as a shaded contour plot of the signal $WTabs(a,b)$ (minimum values in white, maximum values in black). One can easily find the three wave components mentioned above. However, $WT_{abs}(a,b)$ in Fig.1(b) is not suitable for simple and efficient detection of particular waves.

The image of $WT_{abs}(a,b)$ can be simplified by taking a z-axis slice for a chosen value $L \in \langle 0; 1 \rangle$. Thus, contour image C_L is

$$C_L(a,b) = \begin{cases} 1 & \text{if } WT_{abs}(a,b) \in \langle L - \varepsilon, L + \varepsilon \rangle \\ 0 & \text{if otherwise} \end{cases} \quad (3),$$

where ε is a small value. An example of a contour image is shown in Fig.2(b). Only that part of the contour, which is the closest to the highest frequency, is considered (4). Such a contour is called a contour envelope EC and is defined as

$$EC(b) = \min_{a \in A, C_L(a,b) \neq 0} [a] \quad (4)$$

for all b 's. The contour envelope EC is a 1-D function. An example of the contour

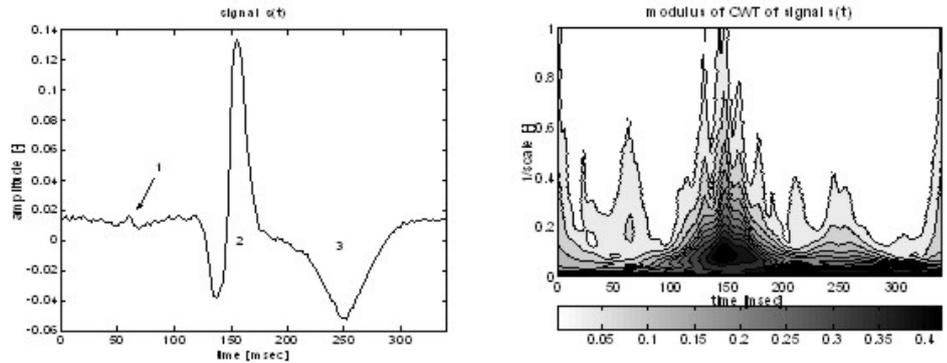


Figure 1. (a) ECG signal and its (b) modulus of CWT. P-wave (1), QRS-complex (2), and T-wave (3) are marked.

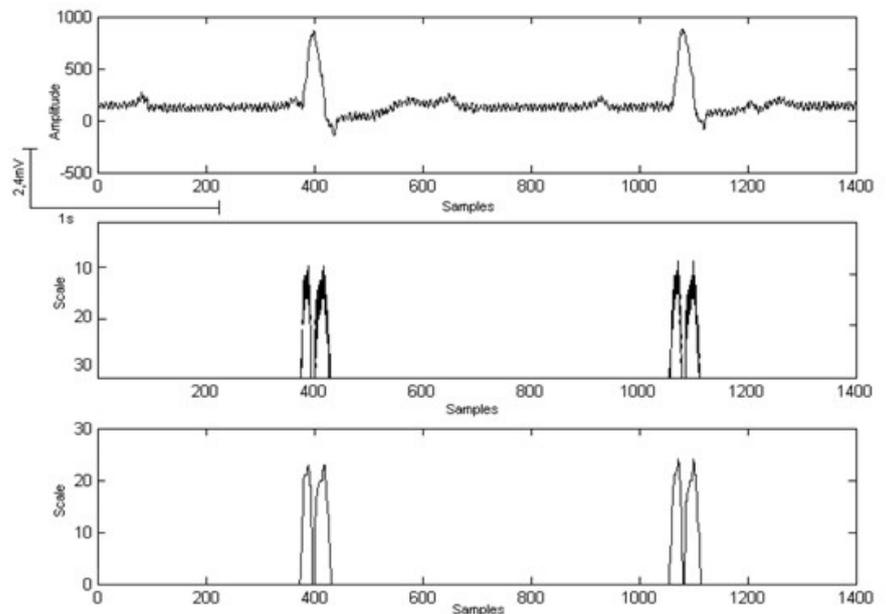


Figure 2. Contour envelope generation: (a) ECG signal, (b) slice through time-frequency spectrum of the signal, (c) contour envelope.

envelope is in Fig. 2(c). EC is further formed by classical signal processing algorithms. In the proposed method, EC is smoothed using a low-pass Lynn's filter defined as

$$H(z) = \frac{1 - z^{-7}}{1 - z} \quad (5).$$

The filtered EC is used for delineation of signal components (Fig.2 EC with a cutting level for QRS detection, Fig. 3 EC with a cutting level for other components).

Simply, the EC is thresholded to remove any noise and suppress the widening effect

due to the above described low-pass filtering (Fig. 3(c)).

While all small value areas are zeroed, all values above the threshold are set to 1. Then, ascending and descending edges represent on-sets and off-sets of particular signal components – P-wave, QRS-complex, and T-wave (Fig.3(d)). The high cutting level L is used only for QRS complex detection to remove other components and avoid false detection (Fig.2) and the low level for other waves to make them visible in the contour envelope (Fig.3). Inner edges are removed and on-sets and off-sets determined.

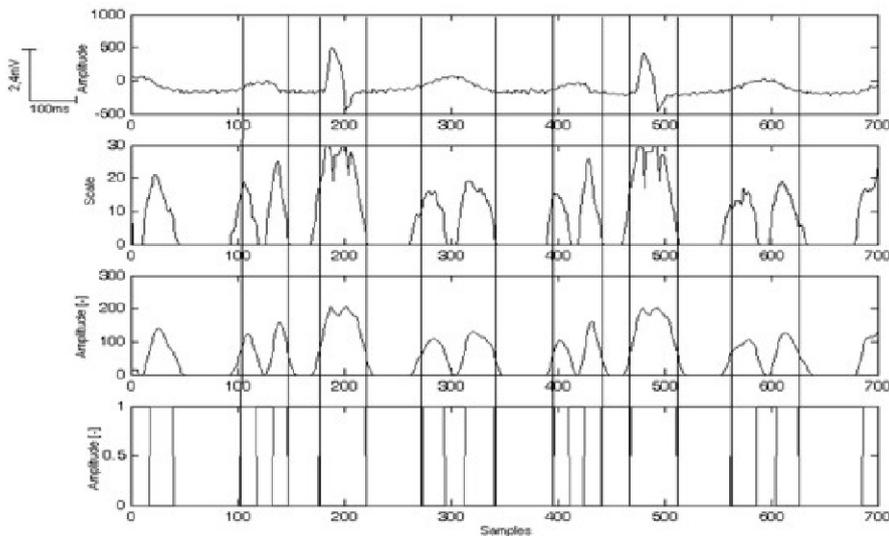


Figure 3. Detection of P-wave, QRS-complex and T-wave using Haar wavelet. (a) ECG signal, (b) contour envelope, (c) filtered conto

APC/PVC classifier

APC/PVC classifier distinguishes between APC (atrial premature contraction) and PVC (premature ventricular contraction).

We analyze the area before the detected QRS complex for the detection of APC. The premature P-wave can be detected in the wavelet contour envelope using the simple wavelets. Fig.4 shows the application of the Haar wave, scale 1-32, and the cutting level:

$$L = 0.125 \text{ mean}(WT_{abs}(a, b)) \quad (6).$$

We can see structures in the wavelet contour envelope before detected QRS, these structures detect the P-wave, which has an irregular P-P interval or an abnormal shape. This contour is also filtered for better performance. Detection edges are processed by a decision algorithm.

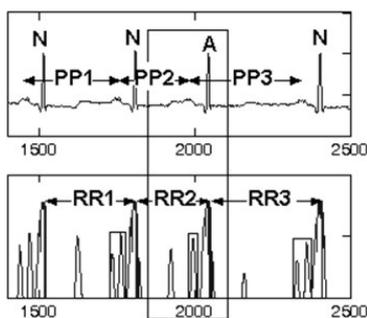


Figure 4. Atrial premature beat: (a) ECG signal, (b) contour envelope.

interval sliding window $[RR_1, RR_2, Rr_3]$ and an averaged RR interval is used to classify a middle RR_2 interval (Fig. 4). The R-R interval surrounding the premature ventricular beat is usually equal to double of the basic R-R interval (Fig. 6) and longer than double for APC (Fig. 4). Different criteria for the RR interval, P-P interval, P wave shape and QRS width are combined together.

Normal rhythm/ Abnormal rhythm classifier

This method shows that we can find differences between abnormal beats and normal beats in the matrix of wavelet coefficients. The Mexican hat wave and scale 1-64 were used. The maximum of square modulus of CWT coefficients of SR (sinus rhythm) in the area around QRS (QRS including) is usually for a between $a=1$ and $a=10$ and it does not change too much from one another. The square modulus of CWT coefficients of premature ventricular contractions has a bigger amplitude, the maximum of coefficients in the QRS area is in a higher scale than for normal beats and width of its shape is higher. It can be compared in Fig.6, where narrow QRS complexes of normal beats have a narrower response and a lower amplitude in the matrix of CWT coefficients than premature ventricular contractions. The position of maximum of CWT coefficients for PVC's in the area of QRS complex on a scale axis is more than 18.

The detected QRS complex has a normal width and it can distinguish APC between PVCs.

The width analysis of QRS in the contour envelope of square modulus of coefficients was done for the level $L=8 * \text{mean}(WT_{abs}(a, b))$, scale 1-32 and Mexican hat wave. The width analysis is documented in Fig. 5. Wide QRS complexes of PVCs has also wide response in contour envelope (Fig. 5(a) versus Fig. 5(b)). EC is filtered, thresholded and further processed by classical algorithms. The next analyzed feature is the R-R interval. A three RR

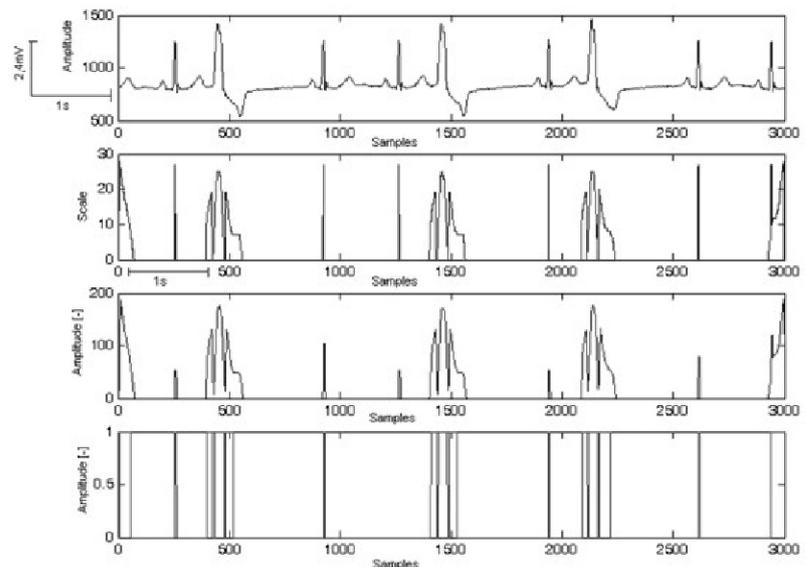


Figure 5. Width analysis: (a) signal No. 119, (b) contour envelope, (c) filtered contour envelope, (d) edges image.

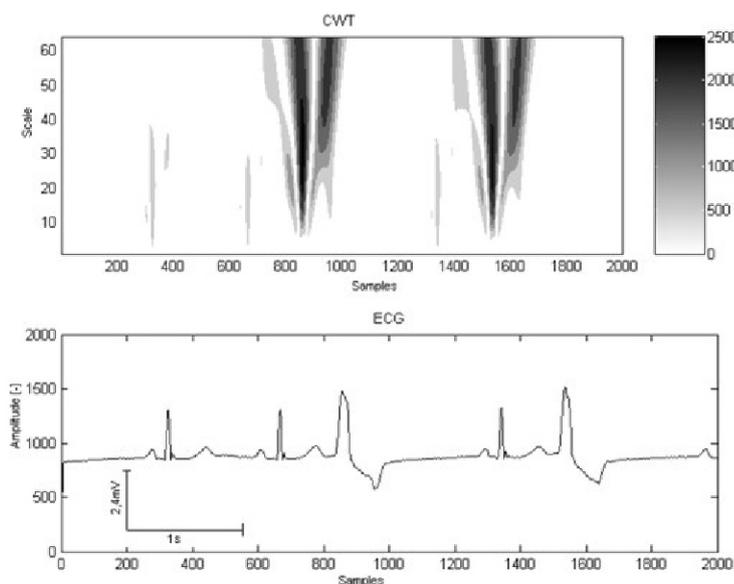


Figure 6. Discrimination of normal and ventricular premature contractions
(a) Square modulus of CWT coefficients – Mexican hat wave,
(b) ECG signal No. 119.

Each change in the size of maximum, position of maximum on a scale axis and width between beats can signalize some abnormality or noise.

RBBB (right bundle branch block) has maximum of coefficients in low scales, usually the same as SR, but it can be distinguished by virtue of QRS-width.

Results

Algorithms have been developed in the MATLAB environment with a Wavelet Toolbox. Signals were tested on selected recordings from the MIT-BIH arrhythmia database created by the Massachusetts Institute of Technology (MIT) and Boston's Beth Israel Hospital (BIH). All analysed recordings contained description of each heart beat used for evaluation of detection accuracy. A single lead MLII (modified limb lead II) was chosen for an analysis. The used sampling frequency was 360 Hz.

Detection of QRS-complexes

The algorithm was tested on 48 signals (no. 100-234) from the MIT-DB library in full length (each approximately of 30 minutes, 650 000 samples). In total, more than 24 hours of signals containing 99555 QRS-complexes were analyzed. The average detection rate was 99.45%, the lowest result was found for the signal no. 207 (88.63%). The signal No. 207 contained

large periods of ventricular fibrillation, where it was difficult to recognize particular QRS-complexes. The detector had the detection rate above 99.9% in 23 signals, and 99% in 43 of signals.

Fig.7 documents robustness of the algorithm. The upper panel shows an original signal with detected heart beats (N under the signal curve). The lower panel shows the same signal with the original description from the MIT-BIH database. In this example, no QRS complex was missed or false detected.

APC/PVC classifier

Classification of ECG was carried out on

34 of MIT-DB recordings. The following beat types were selected for the study: sinus rhythm (SR), premature ventricular contraction (PVC), and premature atrial contraction (APC). Classification was not trivial due to the wide variations of the morphology of the SR, PVC, and APC beats as it can be shown on the recordings. The average accuracy of this method in classification of SR, PVC and APC beats was 94.6%. The accuracy is represented by a ratio of true positive detections to a number of all beats. Numbers of tested SR, PVC, and APC beats and accuracy of detection is summarized in the table below. The accuracy of detection of SR, PVC, or APC beats in the table is represented by a ratio of true positive detections to a number of corresponding beats.

Table 1. Classification accuracy for APC/PVC classifier.

Rhythm	Database	Algorithm	Accuracy (%)
SR	69264	66821	96.1
PVC	5383	3807	71.2
APC	917	271	30.4

We tested 34 signals from the MIT-DB library where PVC and APC appear more frequently. The system described in this paper is patient-independent: an operator is not required to hand-segment samples of each new type of a beat before and during the analysis. The results obtained have shown that continuous wavelet transform enables an arrhythmia analysis and helped identify characteristics for arrhythmia classification.

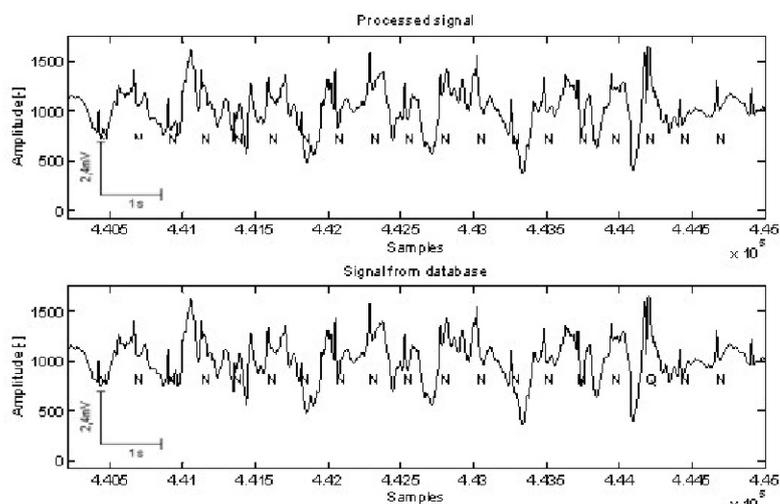


Figure 7. Detection of QRS-complexes – signal No. 105: (a) results, (b) database.

Normal rhythm/abnormal rhythm classifier

The algorithm was tested only on 40 signals from the MIT-DB library with PVC's, APC, RBBB, LBBB (left bundle branch block), SR. We can easily distinguish PVCs between RBBB, SR, APC and other narrow complex arrhythmias by searching for location of local maxima in matrix of CWT coefficients on a scale axis. This method has average accuracy 96.81% for detection of PVCs between APC, SR, RBBB. Signals with LBBBs were also tested, but there were not obtained reliable results. This algorithm is also independent, there is not necessary manual segmentation and manual classification of data.

Discussion and conclusion

These methods are found suitable for a continuing analysis. The cooperation with a cardiologist is necessary for further improvement. QRS complex detection accuracy is comparable to the recently published results. Classification algorithms can recognize clear examples of abnormality and typical progresses, but beats on the border of categories and atypical changes are not classified very successfully. APC/PVC classifier promises good results with some limitations: classification based on the detection of waves and measuring important intervals in contour envelopes can decide about abnormal intervals and values but it cannot decide precisely about the type of arrhythmias. This information cannot be compared to the database. This decision is not possible without practical medical experience. Finding of hidden differences between diseases and WT coefficients are promising techniques. The Normal rhythm/Abnormal rhythm classifier shows that differences can be discovered. We can see in the matrix of coefficients that RBBB has usually maximum of coefficients in the low part of the scale range (1:64), the same as SR (but we can separate them due to their width), LBBB in the middle part and PVC in the higher part. We can determine with high accuracy the PVCs in the matrix of coefficients. They have a maximum in the higher part of the scale range and a broad response in the wavelet image. Only this criterion gives good results for recognizing and distinguishing PVCs between SR, RBBB, APC, NPB (nodal premature beat), JEB (junctional escape

beat) and other narrow complex arrhythmias. It can be used as feature for the automatic computing and training methods. The CWT transform offers a detailed analysis, although this comes at a significant expense in computational load and memory usage. The creation of a general algorithm successful on all ECG cycle variants is a difficult problem. Automatic computing and training methods based on artificial intelligence approaches, hidden Markov models or others could solve the described difficulties. We can detect successfully deviations from the sinus rhythm; broad, wide and deep PVCs, etc., but there are still variations, where human input is necessary.

Acknowledgement

The research was supported by the grants GACR No. 102/07/1473 and the Research Programme of Brno University of Technology MSM0021630513.

References

- [1] Nabney I. T., Evans D. J., Tenner J., Gamlyn L.: Benchmarking Beat Classification Algorithms. *Computers in Cardiology*, 2001, 529-532.
- [2] Sahambi J. S., Prasad G. K.: Classification Of ECG Arrhythmias Using Multi-Resolution Analysis And Neural Networks. In *Proceedings of Convergent Technologies for the Asia-Pacific Region Conference*, Bangalore, India, IEEE 2003, Vol. 1.
- [3] Strang G., Nguyen T.: *Wavelets and filter banks*. Wellesley-Cambridge Press, USA, 1996.
- [4] Kohler B., Hennig C., Orglmeister R.: The principles of software QRS detection. *IEEE Eng. Med. Biol. Mag.*, 2002, pp. 42-54.
- [5] Small M., Yu D. J., Grubb N., Simonotto J., Fox K. A. A., Harrison R. G.: Automatic identification and recording of cardiac arrhythmia. *Comp. Cardiol.*, 2000, vol. 27, pp. 355-358.
- [6] Unser M.: A Practical Guide to the Implementation of the Wavelet Transform. In: Aldroubi, A, Unser M. *Wavelets in Medicine and Biology*. CRC Press, New York, 1996.
- [7] Braunwald E., Zipes D. P., Libby P.: *Heart disease*. W. B. Saunders Company, 2001.
- [8] Chen S. W., Clarkson P. M., Fan Q.: A robust sequential detection algorithm for cardiac arrhythmia classification. *IEEE Trans. Biomed. Eng.*, 1996, vol. 43, pp. 1120-1125.
- [9] Guvenir H. A., Acar B., Demiroz G., Cekin A.: A supervised learning algorithm for arrhythmia analysis. *Comp. Cardiol.*, 1997, vol. 24, pp. 433-436.
- [10] Dickhaus H., Heinrich H.: Classifying

biosignals with wavelet networks. *IEEE Eng. Med. Biol. Mag.*, 1996, pp. 103-104.

- [11] Provazník I., Bardonova J.: *Delineation of ECG Waves Using Complex-Valued Continuous Wavelet Transform*. Springer Berlin Heidelberg, 2008, vol. 18, pp. 158-161, ISSN 1680-0737.
- [12] Chesnokov Yu. V., Chizhikov V. I.: *Continuous Wavelet Transformation in Processing Electrocardiograms in Ventricular Arrhythmia*. *Journal Measurement Techniques*. Springer New York, 2004, no. 4, vol. 47, pp. 417-421, ISSN 0543-1972.
- [13] Shyu L. Y., Wu Y. H., Hu W.: Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. *IEEE Trans. Biomed. Eng.*, 2004, vol. 51, pp. 1269-73.
- [14] Romero Legaretta I., Addison P. S., Reed M. J., Grubb N. R., Clegg G. R., Robertson C. E., Watson J. N.: Continuous wavelet transform modulus maxima analysis of the electrocardiogram: Beat-to-beat characterisation and beat-to-beat measurement, *Int. J. Wavelets, Multiresolution and Information Processing*, 2005, vol.3(1), pp. 19-42.
- [15] Addison P. S.: Wavelet transforms and the ECG: a review. *Physiol. Meas.*, 2005, vol. 26.
- [16] Addison P. S.: *The Illustrated Wavelet Transform Handbook: Applications in Science, Engineering, Medicine and Finance*. Institute of Physics Publishing, Bristol, 2002.
- [17] Mahmoodabadi S. Z., Ahmadian A., Abolhasani M. D.: ECG Feature Extraction Using Daubechies Wavelets. *Proceedings of the Fifth IASTED International Conference, Visualization, Imaging, and Image Processing*, Benidorm, Spain. September 7-9, 2005.
- [18] Daqrouq K., Abu-Isbeih I. N.: Arrhythmia Detection using Wavelet Transform. *EUROCON*, 2007. The International Conference on Computer as a Tool, pp.122-126, 9-12 Sept. 2007.

Contact

Prof. Ing. Ivo Provazník, Ph.D.
 Department of Biomedical Engineering
 Faculty of Electrical Engineering and Communication
 Brno University of Technology
 Kolejní 4, Brno 61200
 Czech Republic
 e-mail: provazni@feec.vutbr.cz

Coronary Heart Disease Mortality in Czech Men, 1980-2004

Jindra Reissigová, Marie Tomečková

Centre of Biomedical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Background: The Czech Republic belongs to countries with high coronary heart disease (CHD) mortality. The aim was to analyze age, period and cohort specific CHD mortality in men from the Czech Republic.

Design and Methods: National data on mortality of men (30-74 yrs.) in the Czech Republic in 1980-2004 were explored. The Poisson regression model was applied to evaluate age, period and cohort effects on mortality.

Results: An adequate fit of CHD men's mortality provided the age-period-cohort model ($p=0.121$). The ratio of mortality of the age group 45-49 to 40-44 was by 17 % lower than that of the age group 40-44 to 35-39. The ratio of the relative risk of period 1995-1999 to 1990-1994 was by 19 % lower than that of the period 1990-1994 to 1985-1989. The ratios of the relative risks between adjoining birth cohort-groups were close to 1.

Conclusions: The fatal CHD risk in Czech men was significantly depended on age (the risk increasing with age), calendar period (the risk reduced in the recent years), and birth cohort (the risk decreasing with birth cohort). The crucial positive change in the CHD mortality trend was observed after the collapse of communism in 1989, and in men past age 40 years.

Keywords: coronary heart disease, cardiovascular, mortality, risk function

Introduction

At the turn of the 21st century in the European Union countries cardiovascular diseases (CVD) cause between 26 % of deaths (France) and 62 % (Bulgaria) in men and between 31 % of deaths (France) and 71 % (Bulgaria) in women [1]. CVD remain the main cause of death in men and in women in all European countries, except men in France, the Netherlands, and Spain. Over a third of deaths from CVD are from coronary heart disease (CHD), which

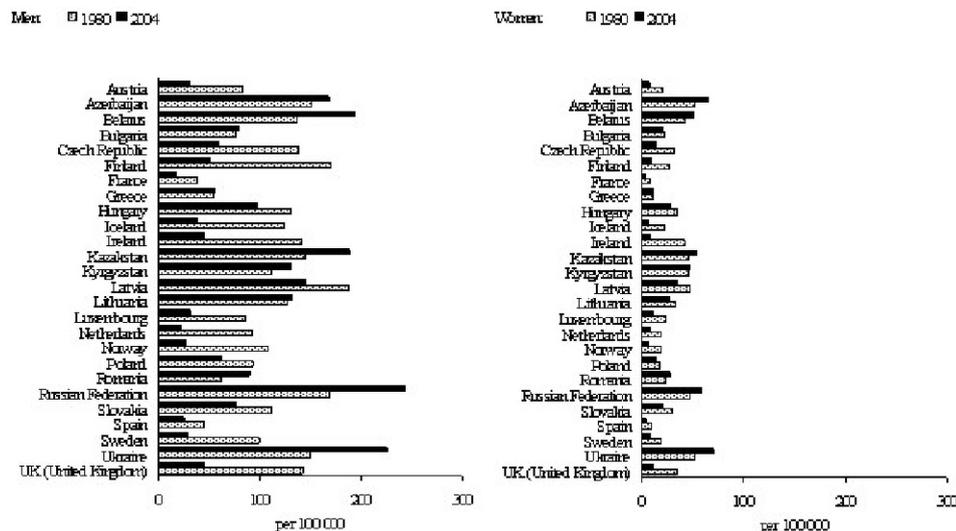


Figure 1. European age-standardized coronary heart disease mortality in some European countries, by sex, 0-64 yrs., 1980¹, 2004.

¹In Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Lithuania and Ukraine mortality in 1981 is presented. Source [1].

by itself is the most common cause of death in Europe. Figure 1 shows age-standardized mortality from CHD in some European countries.

In 1970-2000 mortality from CVD declined almost linearly in western European countries [2]. In eastern European countries, CVD mortality reached a maximum in 1990-1994, followed by a decline in Poland, in Hungary, and in the Baltic states (Estonia, Latvia, Lithuania). An appreciable reversal of the trend was also observed in the Czech Republic (CR) [3]. A main role in the decline in CVD mortality in CR plays most likely a decreasing trend in the occurrence of most major CVD risk factors observed in the 1985-2000/01 period [4].

Nevertheless, in 2006 diseases of the circulatory system were the most frequent cause of death in CR (the 2nd most frequent cause of death were malignant neoplasms, the 3rd injuries) [5]. They were responsible for 50.3 % of all deaths. The leading cause of death from the cardiovascular diseases was CHD. It

caused 47.1 % of all death from CVD in men, and 40.8 % in women. The aim of this article is to analyze mortality from CHD in Czech men in 1980-2004.

Material and Methods

Source data

The numbers of CHD (codes D410-D414 in the 8th Revision of International Classification of Diseases (ICD-8), codes 410-414 in ICD-9, codes I20-I25 in ICD-10) and mid-year population by the 5-year age groups (0-4, 5-9,...80-84, 85+) in 1980-2004 are from the publications "Population changes in the Czech Republic" yearly issued by the Czech statistical office. Mortality from CHD was directly age-standardized to the European population [6].

Statistical methods

An age-period-cohort (APC) modelling was applied to analyse mortality from CHD in men aged 30-74 years from CR in 1980-2004 (note: mortality in the younger age group was low and the reliability of death cause in the elderly can be arguable, these age groups were excluded from the

analysis). The number of CHD deaths was modelled by a log-linear Poisson regression as a function of the age at death, period of death, and birth cohort. The APC modelling was done hierarchically as summarized by Arbyn et al. [7], and finally presented in Table 1 in Results. A goodness-of-fit of a given model was measured by the residual deviance with a chi-square test. The contribution of an additional term to the model was evaluated by comparing the difference in the deviances between the given model and the model without the term against chi-square distribution with the degree of freedom (df) equalled to the difference in df between these two models. The adequacy of the model was further verified by the Akaike's information criterion (AIC) for the fit, and an examination of the Pearson residuals. To test for extra-Poisson variation, the Cameron-Trivedi (C-T) test was applied.

It is well known that a linear dependency exists among the variables of age (a_i , $i=1, \dots, I$), period (p_j , $j=1, \dots, J$) and cohort (c_k , $k=1, \dots, K$): $k=i+j$. In the model involving the variables of age, period and cohort simultaneously, the linear dependency results in non-unique estimations of these parameters (note: the multiple estimations, however, yield the same fitted number of deaths). To obtain unique estimations, a constraint on the estimated parameters is required. However, the different constraint leads to estimations of age, period and cohort trends. In our case, the constraints of fixing two period parameters (1985-1989, 1990-1994), and one cohort parameter (1935-1944) to zero and the constraints so-called second differences by Clayton and Schifflers [8] were also used to estimate the parameters of the age, period and cohort. Other differences are the contrasts expressed as the ratio (difference on a logarithmic scale) of two adjacent relative values (more in Results). The statistical analysis was performed using the R software version 2.3.1.

Results

Figure 2 shows the development of age-standardized CHD mortality in men from CR in 1980-2004, in all age groups together, in the age group of 0-64 years, and in the age group of 30-74 years. For

the ability of comparison with mortality in other European countries (Figure 1), this mortality was adjusted (recalculated) on the age-structure of the European standard population.

Figure 3 shows the age, period, and birth cohort specific (yearly) mortality in the age group of 30-74 years. Mortality is plotted on a logarithmic scale. Mortality was increasing with age (Figure 3a), decreasing in the recent years (Figure 3b), and decreasing with birth cohort (Figure 3c). In other words, Figure 3 depicts mortality data subsequently stratified by age, period and birth cohort, which we have at our disposal for the statistical analysis testing simultaneously to what extent age, period, and birth cohort influence mortality.

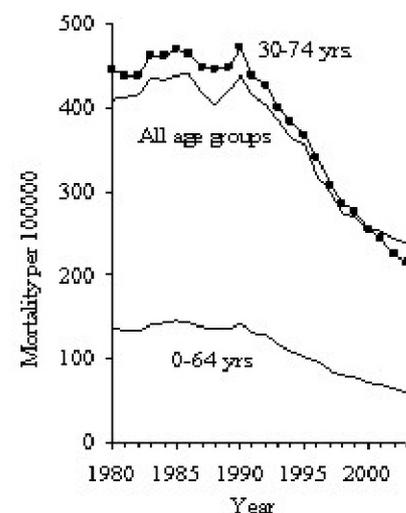


Figure 2. European age-standardized coronary heart mortality in the Czech Republic, men, in all and specific age groups, 1980-2004.

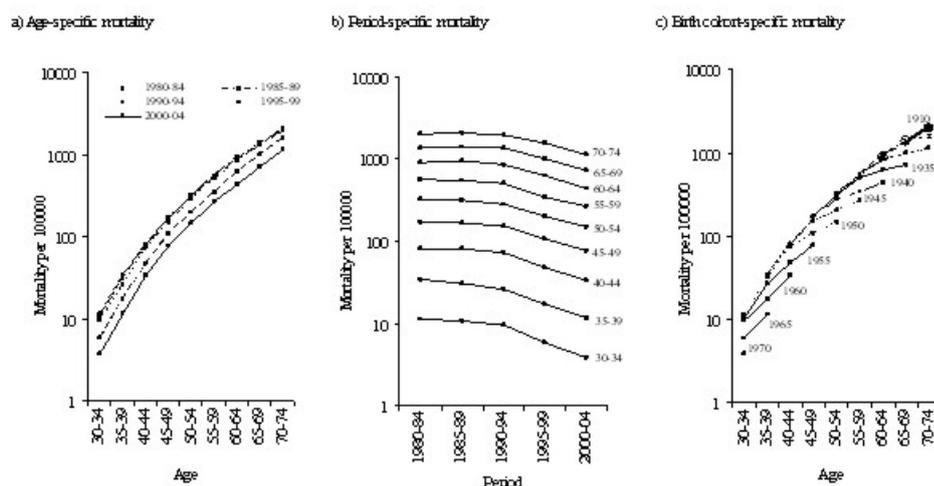


Figure 3. Age, period, birth cohort-specific coronary heart mortality in the Czech Republic, men, 30-74 yrs., 1980-2004¹.

¹In c), birth cohort notation represents the centre of the 10-year interval, e.g. birth cohort of 1940 represents the cohort of men born from 01/01/1935 to 31/12/1944.

The statistical analysis is presented in Table 1, which summaries the results of APC modelling. The effect of age was significant on mortality (the model 1). The fits of the models 2 and 3a-3c were significantly better compared with the model 1, however, still not sufficient. A more adequate fit provided the APC model (the model 4). It means that the risk of death was significantly depended on all three effects: age, calendar period, and birth cohort. In the APC model, no extra-Poisson variation (in other words, no significance difference between mean and variance) was detected ($p=0.500$), and the

Pearson residuals ranged from -1.864 to 1.784.

The self effects of age, period, and cohort on mortality are demonstrated in Figure 4. The age effect was expressed as age-specific mortality per 100 000, the period and cohort effects as relative risk of death from CHD. Increasing age was elevating CHD mortality (Figure 4a). The risk of death from CHD was reducing in the most recent periods (Figure 4b, compared with 1985-1995, i.e. the reference period), and with birth cohort (Figure 4c).

The cohort risks were related to the risk of the cohort 1940 (the reference cohort). The cohort 1940 included men born around (1935-1944) the beginning of the Second World War in 1939. In fact, it was a central cohort and compared with this cohort, the younger and older cohorts are based on fewer data, Figure 3c (it resulted in the varying width of confidence intervals, Figure 4c).

The numbers in brackets in Figure 4 indicate the ratio of two adjacent values. Remind that these ratios are the same for various age, period and cohort parameter estimates, and thus they solve the non-identifiability problem in the APC model as mentioned in Material and Methods. It is important to emphasize that the ratios indicate neither increasing nor decreasing trend. In fact, the values under 1 indicate "deceleration" of the trend, and the values above 1 indicate "acceleration" of the trend. Let's describe the most essential sudden changes in the trends:

- Following Figure 4a): The ratio of mortality of the age group 45-49 to the age group 40-44 was lower by 17 % than that of the age group 40-44 to the age group 35-39, because the mentioned mortality rate was $\left(\frac{165.2}{87.7}\right) / \left(\frac{87.7}{38.6}\right) = 83\%$.

The sequent mortality rates were 89 % for the age group of 45-49 years, and 93 % for the age group of 50 years and over. In summary, the largest deceleration of the mortality trend was observed in men in the age group 40-49 years.

- Following Figure 4b): The ratio of the relative death risk of the period 1995-1999 to the period 1990-1994 was lower by 19 % than that of the period 1990-1994 to the period 1985-1989, because the mentioned risk ratio is $\left(\frac{0.81}{1.0}\right) / \left(\frac{1.0}{1.0}\right) = 81\%$.

Other more than 10% changes in the trend were not observed. So that, the largest decrease in the mortality trend was achieved in 1995-1999 compared to the previous development of the trend.

- Following Figure 4c): No larger sudden changes in the relative death risk were observed by birth cohorts; the ratio of the relative risks between adjoining cohort-groups ranged from 0.92 to 1.06.

Table 1. Age, period, cohort modelling of coronary heart mortality, men, 30-74 yrs., Czech Republic, 1980-2004.

No.1	Model	Residual deviance (D)	Degree of freedom (df)	p-value	Compared with model	$\Delta(D)^2$	$\Delta(df)_3$	p-value	AIC ⁴
0	Interception	355388.0	44	<0.001					
1	Age	15148.0	36	<0.001	0	340240.0	8	<0.001	15600.6
2	Age-Drift	3255.5	35	<0.001	1	11892.5	1	<0.001	3710.1
3a	Age-Age*Drift	2922.5	27	<0.001	2	333.0	8	<0.001	3393.1
3b	Age-Period	388.2	32	<0.001	2	2867.3	3	<0.001	848.8
3c	Age-Cohort	1872.6	24	<0.001	2	1382.9	11	<0.001	2349.2
4	Age-Period-Cohort	28.7	21	0.121	3b	359.5	11	<0.001	511.3

¹Follow Figure 2. If the model 1 holds, the age-specific mortality curves by the periods are identical. If the model 2 applies, the age-specific mortality curves are parallel at constant distances (so-called drift). In the case of model 3a, the age-specific mortality curves are not parallel; the age-specific mortality in successive periods differ by the constant specific for each age group. In the case of the model 3b, the age-specific (or equivalently period-specific) mortality curves, are parallel, but at different distances. If the model 3c holds, the birth cohort-specific mortality curves are parallel at different distances. The model 4 allows for the non-parallel age-specific mortality curves as a function of the period and cohort variables.

²Difference in residual deviances.

³Difference in degree of freedoms.

⁴Akaike's information criterion.

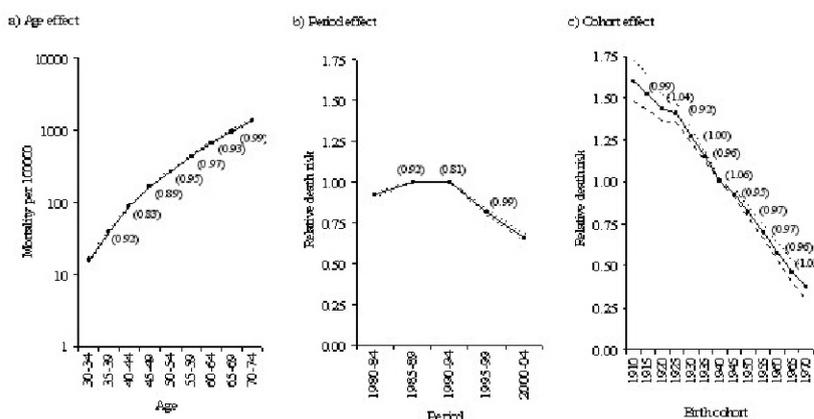


Figure 4. Estimated effects of age, period¹ and birth cohort² (with 95% CI) on coronary heart mortality in the Czech Republic, men, 30-74 yrs., 1980-2004 (the numbers in brackets indicate the ratio of two adjacent relative risks³).

¹Periods of 1985-1989 and 1990-1994 was chosen as reference.

²Birth cohort of 1940 was chosen as reference; birth cohort notation represents the centre of the 10-year interval, e.g. birth cohort of 1940 represents the cohort of men born from 01/01/1935 to 31/12/1944.

³For instance, the ratio of mortality of age group 40-44 (87.7) versus age group 35-39 (38.6) to mortality of age group 35-39 (38.6) versus age group 30-34 (15.6) is $(87.7/38.6) / (38.6/15.6) = 0.92$.

Discussion

APC modelling is used as the method to analyse incidence and mortality from various causes ranged from cancer to cardiovascular diseases. It is a descriptive tool, mainly used for nationally registered data. It helps not only to model but also to predict occurrence of diseases. In this paper, we used APC modelling to explore mortality from CHD in men from CR.

CVD causes nearly half of all death in Europe (48 %), and CHD by itself is the

single most common cause of death in Europe [1]. Primary prevention programmes in many countries attempt to reduce occurrence of CHD through risk factor modification. However, in studies so far conducted, such interventions have limited the effect on mortality, as summarized in the literature [9,10]. On the other hand, as stated there, a small but potentially important benefit (about a 10% reduction in CHD mortality) may have been missed in those studies.

Major cardiovascular risk factors are smoking, elevated blood pressure, elevated cholesterol level, and diabetes mellitus. CR belongs to the European countries with high CHD mortality even if mortality from CHD has been decreasing in CR from the beginning of 1990s (Figure 1). In this study, using APC modelling, we found out that mortality in Czech men was significantly dependent on the age of death, calendar period of death and birth cohort.

Mortality from CHD was increasing with age (Figure 4a). The effect of age on mortality was likely to reflect the dependency of mortality on atherosclerosis ("hardening of blood vessels" – the main underlying cause of CVD), which as has been known may begin in childhood, and getting worse as a person gets older. Likely that decisive age for men is the age around 40/50 years, when the largest deceleration of increasing mortality was observed (the mortality ratio of the age group 45-49 to 40-44 was by 17 % lower than that of the age group 40-44 to 35-39; the mortality ratio of the age group 50-54 to 45-49 was by 11 % lower than that of the age group 45-49 to 40-44).

The risk of death from CHD was reduced in the recent period (Figure 4b). The period effect on mortality reflects the risk factors and the level of health care that act in the period of death. The period risks of 1985-1989, 1990-1994 were fixed to 1 (the reference periods) due to been assumed that the risks were approximately same in these years around the collapse of communism in 1989. Decrease of the risk from fatal CHD in the recent periods (compared to 1985-1994) is explained by the collapse of communism, followed by lifestyle changes and modern treatment. The ratio of the relative death risks of the period 1995-1999 to the period 1990-1994 was by 19 % lower than that of period 1990-1994 to period 1985-1989.

The relative risk of dying from CHD was linearly decreasing with birth cohort, and ranged from 0.38 to 1.60 (Figure 4c). No sudden larger trend fluctuations were observed; the ratios of the relative risks between adjoining cohort-groups were close to 1.0. The effect of birth cohort on mortality can be viewed as risk factors and

environmental exposures typical for a given generation. Etiology is often related to birth cohort.

One limitation of our APC model is that we did not have data on background characteristics of individuals. Therefore we could not relate the age, period, and cohort trends to trends in cardiovascular risk factors (smoking, blood pressure etc.) and precisely to investigate time from risk factor modification to decline in mortality. Our study is not the analytical study, but the descriptive study, which could help medicine workers understand and speculate about the development of mortality from CHD in CR. For instance, our data suggests that a sharper decline in mortality was demonstrated in 5-10 years (1995-1999) after the risk factor modification, which is supposed to start around 1990. Further, even if the aim of our study was not to predict the future CHD mortality trend, mortality among men in CR can be expected to decrease, assuming the trend in the risk factors remain unaltered. The decline is mainly expected due to the fact that the risk of CHD death is decreasing with the year of birth, and in the recent periods. APC modelling of CHD was recently also applied in Finland [11] and Japan [12], in which decline in future CHD mortality is also expected. The results of both studies emphasize the significance of maintaining the recent decline of CHD mortality among middle-aged adults. Few studies on APC mortality were found that were conducted in the past century and predicted development mortality from the past to the beginning of the 21st century. For instance, age-period-cohort effects on CHD mortality in Sweden from 1969 to 1993 were analysed, with forecasts up to 2003 [13]. It would be worth to compare predicted and observed mortality in these studies.

In conclusion, the risk of CHD death in Czech men was significantly dependent on the age of death, calendar period, and birth cohort. The crucial positive change in the mortality trend was observed after the collapse of communism in 1989, and in men at the age of 40-49 years. Assuming age, period and cohort trends remain unchanged, the further decrease in mortality from CHD can be expected in Czech men.

References

- [1] European cardiovascular disease statistics 2008 edition. London: British heart foundation, 2008.
- [2] Kesteloot H., Sans S., Kromhout D.: Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *Eur Heart J* 2006; 27(1):107-13.
- [3] Levi F., Lucchini F., Negri E., La Vecchia C.: Trends in mortality from cardiovascular and cerebrovascular diseases in Europe and other areas of the world. *Heart* 2002; 88(2):119-24.
- [4] Cifkova R., Skodova Z.: Longitudinal trends in major cardiovascular disease risk factors in the Czech population. *Cas Lek Cesk* 2004; 143(4):219-226 (in Czech).
- [5] Deaths 2006. Institute of Health Information and Statistics of the Czech Republic, Prague 2007. [6] Doll R., Cook P.: Summarizing indices for comparison of cancer incidence data. *Int J Cancer*, 1967; 2(3):269-279.
- [7] Arbyn M., Van Oyen H., Sartor F., Tibaldi F., Molenberghs G.: Description of the influence of age, period and cohort effects on cervical cancer mortality by loglinear Poisson models (Belgium, 1955-94), *Arch Public Health* 2002; 60:73-100.
- [8] Clayton D., Schifflers E.: Models for temporal variation in cancer rates. II: Age-period-cohort models. *Stat Med* 1987; 6(4):469-481.
- [9] Ebrahim S.: Systematic review of randomised controlled trials of multiple risk factor intervention for preventing coronary heart disease. *BMJ* 1997; 314(7095):1666-1674.
- [10] Ebrahim S., Beswick A., Burke M., Davey Smith G.: Multiple risk factor interventions for primary prevention of coronary heart disease. *Cochrane Database Syst Rev* 2006; (4):CD001561
- [11] Huovinen E., Härkänen T., Martelin T., Koskinen S., Aromaa A.: Predicting coronary heart disease mortality-assessing uncertainties in population forecasts and death probabilities by using Bayesian inference. *Int J Epidemiol* 2006; 35(5):1246-52.
- [12] Ma E., Iso H., Takahashi H., Yamagishi K., Tanigawa T.: Age-Period-Cohort analysis of mortality due to ischemic heart disease in Japan, 1955 to 2000. *Circ J* 2008; 72(6):966-72.

- [13] Peltonen M., Asplund K.: Age-period-cohort effects on ischaemic heart disease mortality in Sweden from 1969 to 1993, and forecasts up to 2003. *Eur Heart J* 1997; 18(8): 1307-1312.

Acknowledgments

The paper was supported by the project 1M06014 of the Ministry of Education, Youth and Sports CR.

Contact

RNDr. Jindra Reissigová, Ph.D.

CBI, Institute of Computer Science AS CR
Pod Vodárenskou věží 2, Prague 8 182 07
Czech Republic
e-mail: reissigova@euromise.cz

Guideline Knowledge Representation Model (GLIKREM)

David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský

Centre of Biomedical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Summary: The guideline knowledge representation system (GLIKREM) is based on a GLIF model which was published in a GLIF3.5 specification. GLIKREM contains some changes and extensions of the definition and implementation of the original GLIF model. The aim of this article is to give a description of GLIKREM, its construction, its implementation in XML, a realization of the data interface and use of the result model.

Keywords: knowledge representation, GLIF model, guidelines

Introduction

Medical guidelines are used for clinician decision support. They are intended to improve the quality of patient care and reduce costs. Unfortunately, finding information contained in conventional (free text form) guidelines may be difficult. A prerequisite for developing decision support systems that use guidelines is creating computer interpretable representations of the medical knowledge contained in medical guidelines. A number of groups are actively developing computer interpretable guideline representation languages for this purpose. The Arden Syntax [1] is perhaps the best known language for representing medical knowledge in decision support systems. It is a rule based formalism developed for encoding individual clinical rules as Medical Logic Modules. Other approaches share a hierarchical decomposition of guidelines into networks of component tasks that unfold over time [2]. All of the following modelling methods can combine guideline steps in directed cyclical graphs. Asbru is being collaboratively developed at Ben Gurion University and the Vienna University of Technology [3]. It is a time oriented, intention based, skeletal plan specification language that is used to represent clinical protocols.

EON was developed at Stanford University and provides a suite of models and software components for creating guideline-based applications [4]. EON uses a task based approach to define decision support services that can be implemented using alternative techniques [5]. They use the Protégé-2000 environment to build a patient-data information model, a medical-specialty model, and a guideline model that formalizes the knowledge needed to generate recommendations regarding clinical decisions and actions.

GUIDE is part of a guideline modelling and execution framework developed at the University of Pavia [6]. It supports integrating modelled guidelines into organizational workflows, using decision analytical models such as decision trees and influence diagrams, and simulating guideline implementation in terms of Petri nets.

PRODIGY was developed at the University of Newcastle upon Tyne [7]. The PRODIGY project's aim is to produce the simplest, most readily comprehensible model necessary to represent the class of guidelines. Teams of clinicians have used Protégé's knowledge engineering environment [8] to encode three complex chronic disease management guidelines. PROforma was developed at the Advanced Computation Laboratory of Cancer Research [9] in the United Kingdom. It combines logic programming and object oriented modelling and is formally grounded in the R2L Language. PROforma supports four tasks: actions, compound plans, decisions, and enquiries. All tasks share attributes describing goals, control flow, preconditions, and post conditions.

The Guideline Interchange Format version 3 (GLIF) has been collaboratively developed by groups at Columbia, Stanford and Harvard Universities (the

InterMed Collaboratory). The most recent GLIF specification is the GLIF3.5 specification [10,11,12]. Its expression language was originally based on the Arden Syntax. GLIF used the Guideline Expression Language (GEL) to specify criteria and expressions. GLIF now uses GELLO, an extensible object oriented expression language that supports a superset of the functions supported by GEL [13].

GLIF was selected as the model for representing the Computerized Clinical Guidelines in SAPHIRE too [14]. The SAPHIRE system continuously monitors the patients through dedicated agents and supports the healthcare professionals through an intelligent decision support system.

The aim of this article is to describe the design of a guideline knowledge representation model (GLIKREM). GLIKREM is based on a GLIF model. GLIKREM contains some changes in and extensions of the definition and implementation of the original GLIF model. This article describes GLIKREM, its construction, its implementation in XML (eXtensible Markup Language), a realization of the data interface and the use of the model. GLIKREM is used in the Medical Knowledge Representation System (MEKRES) which is developed now [15].

Guideline knowledge representation system

The whole process of the construction of a knowledge model (GLIKREM) from the free text of guidelines, its representation in XML and the use of the model is illustrated in figure 1 [16].

In the stage of model construction from textual guidelines it is important to find the logical and process structure of guidelines (the decision algorithm), all fundamental

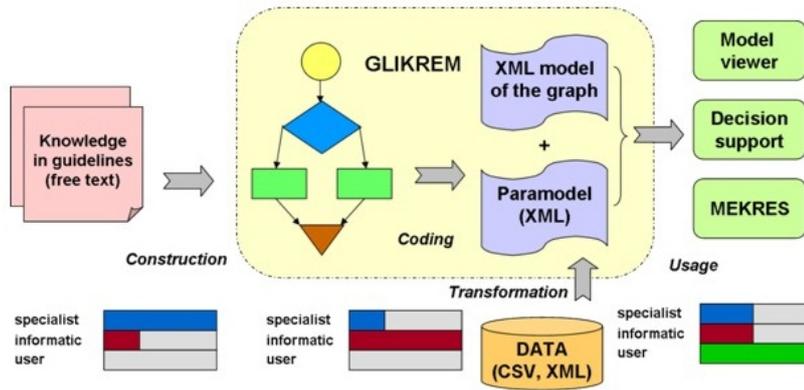


Figure 1. Process of construction, coding and use of GLIKREM. Source [1].

parameters and their interrelationships [17].

In the stage of the model coding the graphic model of guidelines is coded into XML. Besides this a list of basic and derived parameters of the model (paramodel) is created.

The paramodel is used as an interface between the model and real data [18]. In the stage of model construction from textual guidelines it is important to find the logical and process structure of guidelines (the decision algorithm), all fundamental parameters and their interrelationships [17].

In the stage of the model coding the graphic model of guidelines is coded into XML. Besides this a list of basic and derived parameters of the model (paramodel) is created. The paramodel is used as an interface between the model and real data [18].

Graphic model

Main parts (steps) of graph

The model of textual guidelines created in the construction stage is an oriented graph (see fig. 2) which is composed of five main parts (steps):

- **Action step** – specify clinical actions that are to be performed. It can be an application of some therapy, carrying out some examination or measurement etc. Action steps also may refer to sub-guidelines (subgraph), which provide details of the action.
- **Decision steps** are used for conditional branching. These steps are used when

branching is determined by evaluation of logical criteria based on the value of data items. If the decision cannot be made automatically the user can select himself one of the alternative next steps.

- **Branch and synchronization steps** enable concurrence in the model. Guideline steps that follow a branch step can be performed concurrently. Branches have a root in the branch step and eventually converge in a synchronization step. In this step all branches are synchronized after evaluation of a synchronizing condition.

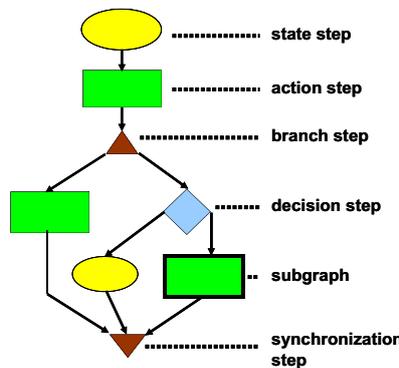


Figure 2. Parts (steps) of the graphic model.

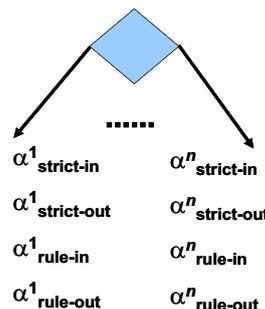


Figure 3. Decision criteria in a decision step.

- **State step** characterizes an object state that has to be observed after execution of the previous step or a state that should exist at the beginning of the model.

Decision criteria

Each decision step specifies four criteria for each decision option $\alpha^1 \dots \alpha^n$ (see fig. 3). The subsequent flow through the model is automatically or manually chosen on the basis of the evaluation of these criteria.

- **Strict-in** – if a *strict-in* is true then the control flows to the guideline step that is specified by that decision option's destination.
- **Strict-out** – if a *strict-out* is true then the decision option's destination is forbidden. The *strict-out* can be but do not have to be opposite of *strict-in*.
- **Rule-in** – if a *rule-in* is true then it is only recommended to flow to the guideline step that is specified by that decision option's destination. The user should select himself one of the next steps with positive *rule-in*.
- **Rule-out** – if a *rule-out* is true the decision option's destination is not recommended but it is not forbidden. The user should not select one of the next steps with positive *rule-out*.

The *strict-out* criterion is evaluated at first. If the *strict-out* criterion (of some option) is evaluated as true the rest of the criteria (of this option) is not evaluated. This option is forbidden. In the opposite case the *strict-in* criterion is evaluated. If the both of *strict-in* and *strict-out* criteria are false, the *rule-in* and *rule-out* criteria are evaluated. The ranking of *rule-ins* and *rule-outs* (of all option's criteria) is left to the users who may use their clinical judgement or may develop their own ranking schemes.

Criteria evaluation

When evaluating the criteria (*strict-in*, *strict-out*, *rule-in*, *rule-out*), it often happens that input parameter values are not known. Therefore the criteria are evaluated in three-value (or multi-value) logic. The logic formulas contain variables (parameters of the paramodel) and logic or relational operands and can take the values:

- **true** – the result of the formula evaluation is true – it is represented by the value 1.

- **false** – the result of the formula evaluation is false – it is represented by the value 0.
- **unknown** – when the formula cannot be evaluated – it is represented by the value $\frac{1}{2}$.

Each logic formula should be expressed in the disjunctive or conjunctive normal form that contains only the logical operations of negation, conjunction (logical multiply) or disjunction (logical sum). This form is equivalent to the original formula.

The definition of the operations of negation, conjunction and disjunction in three-value logic are as follows:

- **negation:** $\neg p = 1 - p$
- **conjunction:** $(p_1 \wedge p_2 \wedge \dots \wedge p_n) = \min(p_1, p_2, \dots, p_n)$
- **disjunction:** $(p_1 \vee p_2 \vee \dots \vee p_n) = \max(p_1, p_2, \dots, p_n)$

A well-designed model has to fulfil several rules for each decision step:

- The automatic flow to the decision option's destination is possible if its *strict-in* criterion is true and its *strict-out* criterion is false.
- If the first rule does not happen the user must chose some destination manually. He (she) can use only the destination whose *strict-out* criterion is false.
- The recommended and not recommended options are offered to the user after the evaluation of *rule-in* and *rule-out* criteria.
- The *strict-in* and *strict-out* criteria must have definite values (true or false). If some *strict-in* or *strict-out* criterion is evaluated as unknown, the user will have to add the missing data.
- Each decision step must be consistent i.e. at the most one of the *strict-in* criteria (of all options) can be true and both of *strict-in* and *strict-out* criteria are not true in the same option.
- Criteria can be dependent on several different parameters. The evaluation of the decision step must be correct (at least one of the *strict-in* or *rule-in* is true in some option and *strict-out* is false in this option) for all of possible values of each parameter.

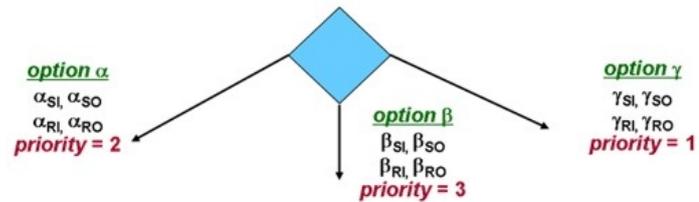


Figure 4. Option priority in the decision step.

Priority in the decision

If these rules are thoroughly applied the user can insert missing data when it is necessary. If a strict-in criterion of some option (destination) is true the evaluation of criteria of other options is not needed. The amount of necessary essential data is dependent on order of single option evaluations. Therefore it is necessary to set an order of evaluation i.e. to set the priority of decision options. The priority of each option is chosen by a specialist (see fig.4).

In the example (see fig. 4) the criteria of the option γ is evaluated first then the criteria of the option α and finally the criteria of the option β . If strict-in criterion of the option γ is true the options α and β will be not evaluated.

Synchronization conditions

In modelling parallel branches (with branch and synchronization steps) it is necessary to specify which branch is necessary and which one is optional. That is why the synchronization conditions are set for each synchronization step (see fig.5).

The synchronization condition has to be expressed in the disjunctive or conjunctive normal form. The onfly parameters of the paramodel are used as variables in the synchronization condition. Values of onfly parameters are:

- value 1 – the user goes through the branch which is represented by this onfly parameter,
- value 0 – the user do not go through this branch.

A value of the onfly parameter is set by a "shadow" operation (see XML representation of model) in the last step of the branch.

In the example (see fig. 5) the synchronization condition $(\alpha_1 \vee \alpha_2) \wedge \alpha_3$ will be true if

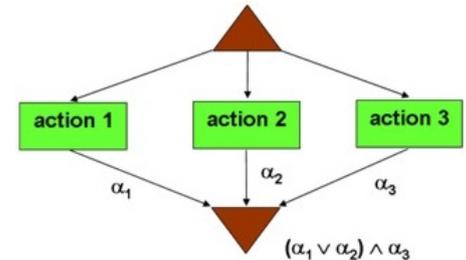


Figure 5. Synchronization condition.

the user goes through branch α_3 and goes through one of the branches α_1 or α_2 .

Representation of the graphic model in XML

We specified our own XML scheme for representation of graphic model and developed a graphical editor for construction of the graphical model and its translation into XML. The whole model consists of a sequence of XML elements which represent steps of the graphic model. The XML syntax of the model contains elements for description of attributes of model steps, elements for graphical symbols of steps and elements for the decision support (see fig. 6).

The description of elements is as follows:

- **glikrem** – the root element of the whole model (GLIKREM),
- **head** – the definition of head – see the paragraph Model extension for using in MEKRES,
- **graph** – the root element of the graphic model contains elements step,
- **step** – the attributes of one step, it contains elements:
 - **name** – the unique identification of the step,
 - **type** – the type of the step, the type can be: *action, case, branch, synchronization, state, subgraph,*
 - **caption** – the caption of the graphical symbol of the step,

- text** – the description of the step,
x – the x-coordinate of the graphical symbol,
y – the y-coordinate of the graphical symbol,
width – the width of the graphical symbol (in pixels),
height – the height of the graphical symbol (in pixels),
focus – the highlighting of the step for visualization of a path in the graph, possible values are:
not = not highlighted,
auto = automatically highlighted,
user = highlighted by the user simulation,
status – the location of the step in the model, possible values are:
start = the first step (the graphical model starts with this step),
end = the last step (usually the final patient's state),
in = the internal step,
- operation** – the list of "shadow" operations of the step (see fig. 7), each operation (element ops) contains elements:
otype – the type of operation, possible values are:
insert = the insertion of the parameter value,
get = the getting of the parameter value,
put = the putting of the parameter value,
open = the opening of some file (e.g. subgraph),
oparam – the name of the parameter or the file,
ovalue – the value onfly for *put* operation,
- next** – the list of follow branches (options), each option (element option) contains elements (see fig. 8):
nname – the identification of the option,
nstep – the identification of the target step,
ncaption – the caption of the option,
npriority – the priority of the option, the default value is 1,
nstrictin – *strict-in* criterion,
nstrictout – *strict-out* criterion,
nrulein – *rule-in* criterion,
nruleout – *rule-out* criterion,
nnote – the note of the option,
nline – the description of line to the target step in a form of the sequence of points (elements point):

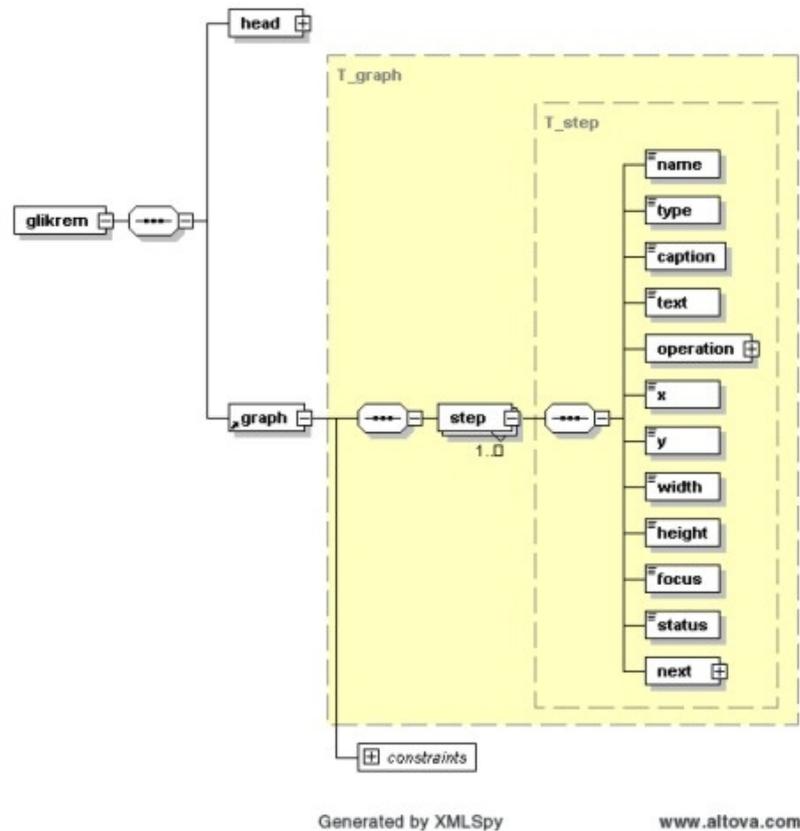


Figure 6. XML scheme of the graphic model representation.

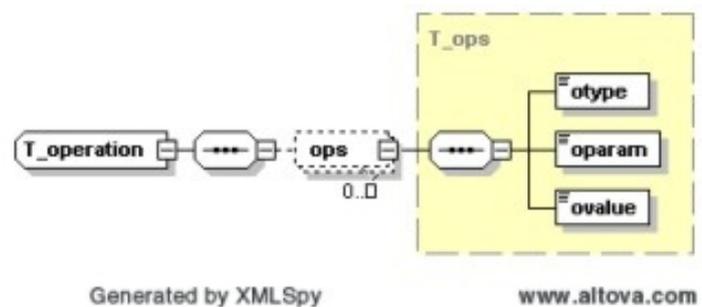


Figure 7. XML scheme of shadow operations.

- px** – the x-coordinate of the point,
- py** – the y-coordinate of the point.

pvalue and
 params/param[pid="FINDINGS"]/pvalue.

Every type of decision condition is usually a logic formula with basic and derived parameters of the paramodel as variables in these formulas. The notation of the decision conditions (*strict-in*, *strict-out*, *rule-in*, *rule-out*) adheres the **XPath** syntax.

The notation of the synchronization conditions is analogical (in XPath) but we need onfly parameters as variables for these conditions.

For example the definition of condition with parameters SYMPTOMS and FINDINGS:
 /params/param[pid="SYMPTOMS"]/

For example the definition of synchronization condition with onfly parameters (branches) DG_S3 and DG_S4:
 /params/param[pid="DG_S3"]/pvalue or /params/param[pid="DG_S4"]/pvalue.

Model of parameters (paramodel)

The essential parameters for GLIKREM are often saved (in databases or clinical information systems) in a form which is not efficient for direct use in the knowledge model. The model of parameters (paramodel) is a better solution. The paramodel is constructed along with GLIKREM. The paramodel serves as a data interface between real data and GLIKREM. The values of parameters of the paramodel are derived from real data by data transformation (see the paragraph Data transformation). Only the parameter names of the paramodel are used in the decision conditions.

Representation of paramodel in XML

The paramodel is constructed by user in the graphic model editor and represented in the XML the same way as a graphic model. The whole paramodel contains elements which represent the respective parameter (see fig. 9). The parameters of the GLIKREM can be as follows:

- **basic** – the basic parameter is directly measurable or has a value which is found by another method,
- **derived** – the parameter is derived from basic parameters by the application of a logical or arithmetical operation,
- **onfly** – the parameter for synchronization conditions which contains a logical value. The value is set by a "shadow" operation in the last step of the branch which is antecedent to the synchronization step.

The description of the paramodel elements is as follows:

- **paramodel** – the root element of the paramodel,
- **head** – the definition of the head,
- **params** – the root element of the parameters which contains elements param,
- **param** – the description of one parameter,
 - pid** – the unique identification of the parameter,
 - ptype** – the type of the parameter, possible values are: basic, derived and onfly,
 - pname** – the full name of the parameter,
 - pdatatype** – the data type of the parameter (XML data type),
 - pvalue** – the value of the parameter,

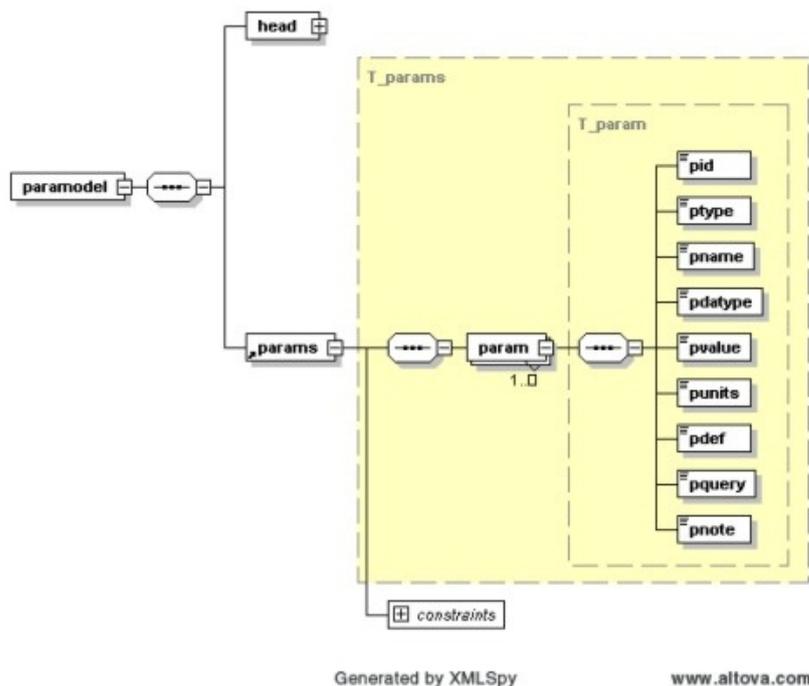


Figure 9. XML scheme of the paramodel.

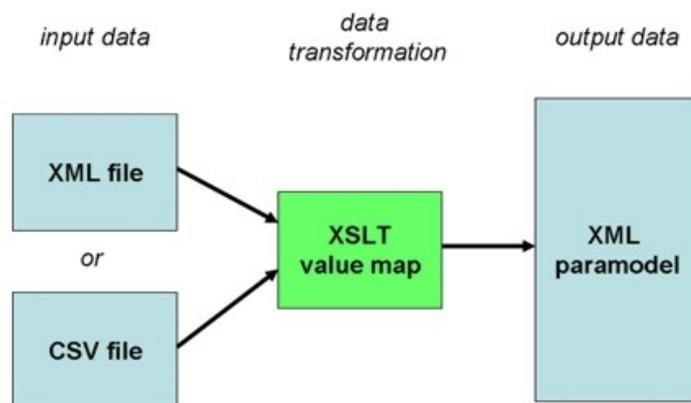


Figure 10. The process of the data transformation into the paramodel.

- punits** – the unit of the parameter,
- pdef** – the definition of the derived parameter in Xpath,
- pquery** – the definition of the database query (e.g. in SQL),
- pnote** – the note of the parameter.

Transformation of input data

The real data of patients are stored in clinical information systems and databases of different types. It is usually possible to export data from these systems in the form of XML or CSV (Comma-Separated Values) files. From these files it is necessary to transform data into the paramodel (in XML). A simplified approach of the transformation you can see in figure

10. It is necessary to define the value-map that contains an input and a result. The input is the parameter and its value as exported by an XML-file or a CSV file and the result is a mapping to the parameter name and value as used in the paramodel. The definition of the value-map is in the XSLT (eXtensible Stylesheet Language Transformations) file form. For example the determination of the systolic blood pressure value as an average from measured values which are stored in the XML file. The input data are the values of the different systolic blood pressure measuring and the result is the average value (see fig. 11).

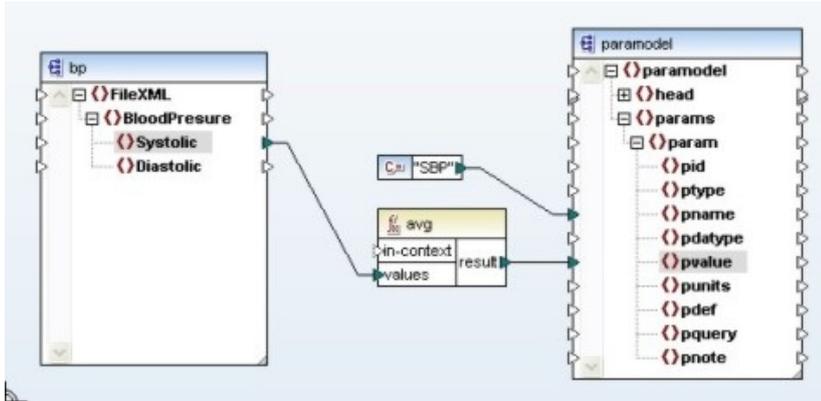


Figure 11. Transformation of systolic blood pressure value

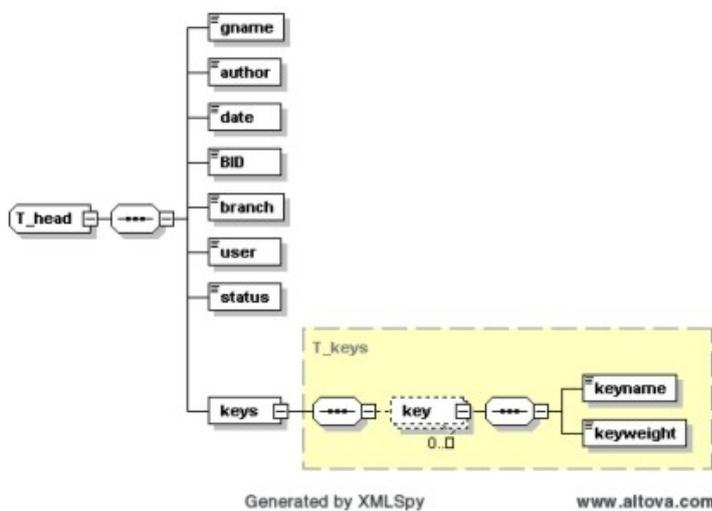


Figure 12. XML scheme of key attributes.

Model extension for using in MEKRES

The resulting GLIKREM is extended by key attributes for use in MEKRES (Medical Knowledge REpresentation System). These key attributes are used in the selection algorithm of a relevant model and they are coded in XML and stored in the element head along with the graphic model (see fig. 12).

The description of the element head is as follows:

- **gname** – the name of knowledge model (GLIKREM),
- **author** – the author(s) of the model,
- **date** – the last update date,
- **BID** – the branch identification – e.g. International Classification of Diseases (ICD),
- **branch** – the branch described by the model,

user – the user of system to whom the model is primary determined, possible values are: *patient, GP, specialist, operator (of emergency rescue), everybody*,

- **status** – the model validity, possible values are: *valid, draft, expired*,
- **keys** – the list of keys described by the model contains elements **key**, **keyname** – the name of the **key**, **keyweight** – the weight of the key – the model description rate of the key.

Selection algorithm of relevant model

The principle of the whole medical knowledge representation system (MEKRES) and the algorithm to select therelevant GLIKREM model is illustrated in figure 13. The select algorithm is described as follow:

- A set of all branches and keys (key attributes) which corresponds to participant state (his/her attributes) is determined for every specific participant (user, patient, physician, operator, ...) and his attributes.
 - e.g. branches "diabetes" and "hypertension" which corresponds to patient attributes "SBP", "DBP", "glycaemia".
- For each branch and key the set of knowledge models (GLIKREM) are chosen from a set of all GLIKREMs where each model contains attribute branch and key weight (element keyweight) is non zero.
 - e.g. models G_1 (hypertension) and G_2 (diabetes).
- For each chosen model a general aggregate operator (e.g. $R_g = \sum keyweight(k)$) is defined. The operator determines the relevant score of the model.

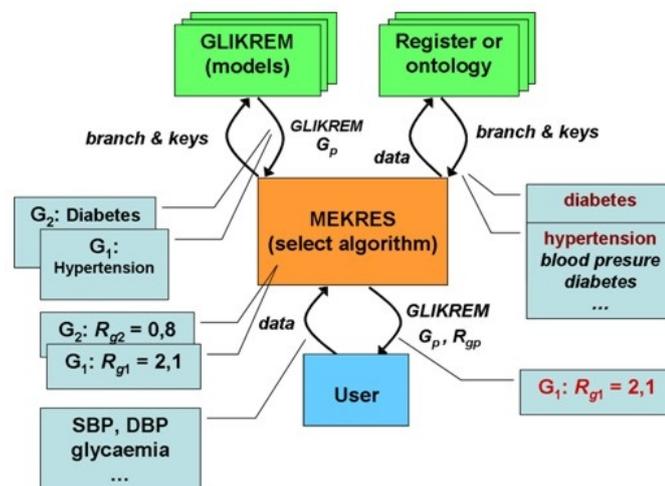


Figure 13. Medical knowledge representation system.

- e.g. the model G_1 has operator $R_{g_1} = 2,1$ and the model G_2 has $R_{g_2} = 0,8$.
- The model with the highest relevance value or a list of models ordered by the value of relevance is offered to participant.
 - e.g. the model G_1 (the highest relevance value) is offered to a patient.

Conclusion

The designed knowledge representation model (GLIKREM) based on the GLIF3.5 specification is the universal tool for the formalization of knowledge stored in a free text form (e.g. medical guidelines). The XML representation of the graphic model and the creation of the data interface in a paramodel form makes it possible to use it in different types of applications and to connect to real data from different sources (e.g. XML or CSV files). The definition in XSLT file form is used for the transformation of input data and the graphic model too.

The main extensions and benefits of the designed model are as follows:

- Our own formal representation of the graphic model in XML. The result of this representation is a XSD (XML Schema Definition).
- A more accurate definition of decision criteria and their evaluation in a three value logic.
- A definition of priority attribute for a reduction of a number of needed option evaluations in decision steps.
- A design of a paramodel (in XML) which is used as an interface between GLIKREM and real patient data in a clinical information system.
- An extension of GLIKREM by key attributes. This extension is used in a Medical Knowledge Representation System.

Acknowledgment

The paper is supported by the grant project No. AV0Z10300504 and No. 1M06014 of Ministry of Education, Youth and Sports CR.

References

- [1] Peleg M., Ogunyemi O., Tu S., et al.: Using features of Arden Syntax with object-oriented medical data models for guideline modeling. Proc AMIA Symp. 2001:523-7.
- [2] Peleg M., Tu S.W., et al.: Comparing Computer-interpretable Guideline Models: A Case-study Approach. The Journal of the American Medical Informatics Association, 2003 Jan-Feb; 10(1): 52-68.
- [3] Shahar Y., Miksch S., Johnson P.: The Asgaard Project: a task-specific Framework for the application and critiquing of time-oriented clinical guidelines. Artif Intell Med 1998;14:29-51.
- [4] Tu S.W., Musen M.A.: A flexible approach to guideline modeling. Proc AMIA Symp 1999:420-424.
- [5] Tu S.W., Musen M.A.: From guideline Modeling to guideline execution: Defining guideline-based decision-support Services. Proc AMIA Annu Symp 2000:863-867.
- [6] Quaglini S., Stefanelli M., Lanzola G., Caporusso V., Panzarasa S.: Flexible guideline-based patient careflow systems. Artif Intell Med 2001;22:65-80.
- [7] Johnson P.D., Tu S.W., Booth N., Sugden B., Purves I.N.: Using scenarios in chronic disease management guidelines for primary care. Proc AMIA Annu Fall Symp 2000:389-393.
- [8] Gennari J.H., Musen M.A., Fergerson R.W., Grosso W.E., Crubezy M., Eriksson H., et al.: The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. Int J Hum Comput Stud. 2003;58(1):89-123.
- [9] Bury J., Fox J., Sutton D.: The PROforma guideline specification language: progress and prospects. Proceedings of the First European Workshop, Computer-based Support for Clinical Guidelines and Protocols (EWGLP 2000), 2000.
- [10] Ohno-Machado L., Gennari J.H., Murphy S.N., Jain N.L., Tu S.W., Oliver D., et al.: The GuideLine Interchange Format: A model for representing guidelines, Journal of the American Medical Informatics Association. 1998. ISSN 1067-5027.
- [11] Zeng Q.: Guideline Interchange Format 3.5 technical specification. Available at: <http://www.glif.org> - Peleg, Boxwala, et al.
- [12] Boxwala A.A., Peleg M., Tu S.W., Ogunyemi O., Zeng Q., Wang D., et al.: GLIF3: A Representation Format for Sharable Computer-Interpretable Clinical Practice Guidelines. J Biomed Inform. 2004;37(3):147-61.
- [13] Sordo M., Boxwala A., Ogunyemi O., Greenes R.: Description and Status Update on GELLO: a Proposed Standardized Object-oriented Expression Language for Clinical Decision Support. Medinfo; 2004; 2004. p. 164-8.
- [14] Laleci G.B., Dogac A., et al.: SAPHIRE: A Multi-Agent System for Remote Healthcare Monitoring through Computerized Clinical Guidelines. online at <http://www.srdc.metu.edu.tr/webpage/projects/saphire/>
- [15] Buchtela D., Peleska J., Vesely A., Zvarova J.: Medical Knowledge Representation System, 21st International Congress MIE 2008 in Göteborg, 2008, Sweden.
- [16] Buchtela D., Peleska J., Vesely A., Zvarova J.: Method of GLIF model Construction and Implementation, The XIX International Congress of the European Federation for Medical Informatics. 2005.
- [17] Peleška J., Anger Z., Buchtela D., Šebesta K., Tomečková M., Veselý A., Zvára K., Zvárová J.: Formalization of Medical Guidelines. European Journal for Biomedical Informatics. 2005. Prague.
- [18] Veselý A., Zvarova J., Peleška J., Buchtela D., Anger Z.: Medical Guidelines Presentation and Comparing with Electronic Health Record. International Journal of Medical Informatics 75 (3-4). 2006. 240-245

Contact

Ing. David Buchtela

Centre of Biomedical Informatics
Institute of Computer Science AS CR
Pod Vodárenskou věží 2
Prague 8 182 07
Czech Republic
e-mail: buchtela@euromise.cz



EJBI 2008

ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 4 (2008)

Czech version

Cooperating journal – Lékař a technika

Editor-in-Chief: Karel Roubík (Czech Republic)

Publisher: Společnost biomedicínského inženýrství a lékařské informatiky ČLS JEP
a Fakulta biomedicínského inženýrství ČVUT v Praze

www.ejbi.eu



Content

Czech version

- cs 1 Předmluva
- cs 2 - 6 [Konturový vlnkový detektor pro analýzu arytmií](#)
Dina Kičmerová, Ivo Provazník
- cs 7 - 10 [Ischemická choroba srdeční u mužů v České republice, 1980-2004](#)
Jindra Reissigová, Marie Tomečková
- cs 11 -17 [Model reprezentace znalostí v doporučeních](#)
David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský

Předmluva

Jana Zvárová
Vedoucí redaktorka

Biomedicínská informatika je rychle se rozvíjející obor, s důležitými aplikacemi a dopady na celý biomedicínský svět a poskytování zdravotnické péče. European Journal for Biomedical Informatics (EJBI) reaguje na velkou evropskou potřebu sdílet informace v mnohojazyčné a multikulturní Evropě.

EJBI otevírá nový model elektronického publikování pro obor biomedicínské informatiky. EJBI vydává přijaté recenzované příspěvky v angličtině a současně v dalších evropských jazycích. Tím otevírá nové možnosti pro rychlejší přenos vědecko-výzkumných poznatků z mnoha evropských zemí širokému mezinárodnímu společenství výzkumných

pracovníků v biomedicínských oborech, lékařů, ostatních zdravotnických pracovníků a občanů. Kromě toho časopis umožňuje, aby se výsledky vědecko-výzkumné práce a praktické zkušenosti zahraničních odborníků staly přístupné a srozumitelné široké veřejnosti v každé zemi, kde je používán některý z evropských jazyků.

Cílem redakční rady je dosáhnout co nejvyšší vědecké úrovně časopisu a současně ukázat osvědčené postupy aplikací metod biomedicínské informatiky široké veřejnosti. Redakční rada je složena z vynikajících evropských odborníků v oboru biomedicínské informatiky. Redakční rada usiluje o to, aby

prezentace vědeckých prací v angličtině s jejich odbornými překlady do dalších evropských jazyků významně přispěla ke sjednocování odborné terminologie v evropských jazycích. Jsem přesvědčena, že tato činnost pro EJBI přispěje k propagaci dobrého kreditu časopisu.

V roce 2008 EJBI zahájilo spolupráci s nakladatelstvím Schattauer ve zpřístupňování kvalitních informací z biomedicínské informatiky v dalších evropských jazycích. Více informací získáte v Předmluvě časopisu *Methods of Information in Medicine*, Issue 2 2008. První společnou publikaci EJBI a *Methods* naleznete zde.

Konturový vlnkový detektor pro analýzu arytmií

Dina Kičmerová, Ivo Provazník

Ústav biomedicínského inženýrství, Fakulta elektrotechniky a komunikačních technologií VUT v Brně, Česká republika

Souhrn: Článek se věnuje návrhu automatického detektoru vybraných arytmií. Navržený algoritmus využívá spojitou vlnkovou transformaci (CWT) v kombinaci s analýzou její konturové obálky. Vlnková transformace byla využita v detektoru R vln, k rozlišení normálních a abnormálních stahů a pro detekci předčasných atriálních kontrakcí (APC) a předčasných komorových kontrakcí (PVC). Algoritmus byl ověřen rozsáhlým testováním na MIT/BIH databázi. Hledáním lokálních maxim ve vlnkové konturové obálce jsou účinně detekovány R vlny. Celková úspěšnost detekce testovaná na 48 půlhodinových signálech je 99,5 %. Byly otestovány dva typy klasifikace: 1. klasifikace založená na konturové obálce a detekci význačných bodů s celkovou úspěšností 94,6 %, 96,1 % pro sinusový rytmus (SR), 30,4 % pro APC, 71,2 % pro PVC a 2. lokalizace maxima umocněných koeficientů spojitě vlnkové transformace v oblasti QRS komplexu k určení PVC mezi SR, blokádu pravého Tawarova raménka (RBBB), APC a dalšími arytmiemi s úzkým QRS komplexem s přesností 96,8 %.

Klíčová slova: arytmie, EKG signál, vlnková transformace, kontury

Úvod

Pro analýzu medicínských dat jsou často využívány automatické systémy. Moderní zdravotnictví produkuje obrovské množství dat, která dávají příležitost k hledání explicitního porozumění fyziologii. Nové metody mohou pomoci zjednodušit a urychlit zpracování velkého objemu dat. Lékaři musí často určovat pacientovu diagnózu na základě mnoha numerických hodnot měřených během vyšetření. Orientace v takovém množství dat není vždy jednoduchá a jednoznačná. Proto existuje řada konzultačních systémů, které se snaží minimalizovat lidské chyby. Moderní analýza elektrické aktivity srdce používá jednoduché

i sofistikované algoritmy digitálního zpracování signálu implementované v elektrokardiografech [1,2,4-6,8-18]. Algoritmy mohou být rozděleny do tří skupin: pracující v časové [8], frekvenční [5] a časově-frekvenční oblasti [10-18]. První dvě třídy patří mezi klasické metody, které jsou úspěšné v mnoha klinicky testovaných aplikacích, v analýze EKG, detekci QRS komplexu [4], jeho začátku a konce, v analýze změn ST segmentu, analýze variability srdečního rytmu, atd. Nové algoritmy pracují v časově-frekvenční oblasti [10-18] a kombinují některé výhodné vlastnosti známé z klasických metod – zejména umožňují frekvenční analýzu se zachováním časové informace o analyzovaných jevech. Použití prosté časové analýzy neposkytuje dostatečně kvalitní výsledky díky nízké citlivosti, což je způsobeno malou amplitudou změn. Frekvenční analýza zvyšuje senzitivitu, ale nedokáže detekovat, v které fázi srdečního cyklu změna vznikla.

V současnosti je pro časově-frekvenční analýzu nejvíce využívaná vlnková transformace (WT) [10-18]. WT je populární pro snadnou implementaci a její výsledky lze snadno interpretovat stejně jako Fourierovu transformaci pro frekvenční analýzu. Mnoho variant WT nabízí mnoho možností při výběru od redundantní detailní analýzy až po rychlou analýzu s minimem výstupních dat. Výběr patřičné varianty WT záleží na typu aplikace (potlačování šumu, značení časově-frekvenčních komponent užitečného signálu, detekce významných bodů, detekce a pozorování významných změn v čase, atd.).

V tomto článku se zabýváme použitím vlnkové transformace k analýze morfologie srdečních stahů a k EKG klasifikaci. Cílem této práce je rozlišit různé typy arytmií a detekce abnormalit. Prezентujeme klasifikátor arytmií založený

na spojitě vlnkové transformaci (CWT) k identifikaci stahů normálních, abnormálních předčasných komorových kontrakcí (PVC) a předčasných atriálních kontrakcí v EKG signálu. Výskyt arytmií je nepředvídatelný. Cílem této studie je vyvinout metodu používající CWT koeficienty EKG vln. Budou diskutovány dva typy srdečních arytmií: PVC (předčasné komorové kontrakce) a APC (předčasné atriální kontrakce), které jsou nejběžnějšími typy arytmií při monitorování EKG.

Metoda

Vlnková transformace (WT) je založena na použití sady matematických analyzujících funkcí nazývaných „vlnky“. Vlnky provádějí rozklad EKG signálu do sady vlnkových koeficientů a jsou odvozeny od obecně komplexní funkce $\psi(t)$ dilatací a posunem v čase. Každá analyzující funkce $\psi_{a,b}(t) = \psi((t-b)/a)$ má vlastní parametry – časovou lokalizaci definovanou posunem b a frekvenční pásmo definované dilatačním faktorem (měřítkem) a . Každý výsledný vlnkový koeficient koresponduje s měřením signálu v daném čase a v daném frekvenčním pásmu. Vlnkové koeficienty mohou být jednoduše vyjádřeny následující definicí vlnkové transformace (CWT)

$$WT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left(\frac{t-b}{a} \right) f(t) dt \quad (1),$$

kde $f(t)$ je analyzovaný EKG signál a $*$ označuje komplexní sdruženost. Další informace ohledně vlnkové transformace a jejích vlastností mohou být nalezeny v [3, 6] a [16].

Rozměrování EKG komponent

Navržená metoda se skládá ze dvou hlavních částí: 1. spojitě vlnkové transformace s diskretizovaným výstupem a 2. klasických algoritmů pracujících v časové oblasti.

WT (Rov. (1)) je dvouřozměrnou funkcí, která reprezentuje signál v časově-frekvenční oblasti. Pro snazší analýzu výsledné funkce $WT(a,b)$ použijeme její absolutní hodnotu $abs(WT(a,b))$ a dále normalizaci. Rovnice 1 může být přepsána do tvaru

$$WT_{abs}(a,b) = \frac{abs(WT(a,b))}{\max_{a,b}[abs(WT(a,b))]} \quad (2).$$

Příklad EKG analýzy s použitím CWT a Morletovy vlnky je dokumentován na obrázku Obr. 1. V signálu vidíme tři hlavní vlny (vrcholy) (panel (a): malá P-vlna označená (1), vysoký QRS-komplex označený (2) a vysoká T-vlna označená (3)). Časově-frekvenční spektrum signálu je zobrazeno jako stínované kontury $WTabs(a,b)$ (kde minimální hodnota odpovídá bílé a maximální hodnota černé). V obrázku mohou být jasně identifikovány komponenty zmiňované dříve. Nicméně $WT_{abs}(a,b)$ na Obr. 1(b) není vhodná pro jednoduchou a efektivní detekci patřičných vln.

Obrázek $WTabs(a,b)$ může být zjednodušen řezem v ose z pro zvolenou hodnotu $L \in \langle 0; 1 \rangle$. Tím vznikne konturový obraz C_L

$$C_L(a,b) = \begin{cases} 1 & \text{if } WT_{abs}(a,b) \in \langle L - \varepsilon; L + \varepsilon \rangle \\ 0 & \text{if otherwise} \end{cases} \quad (3),$$

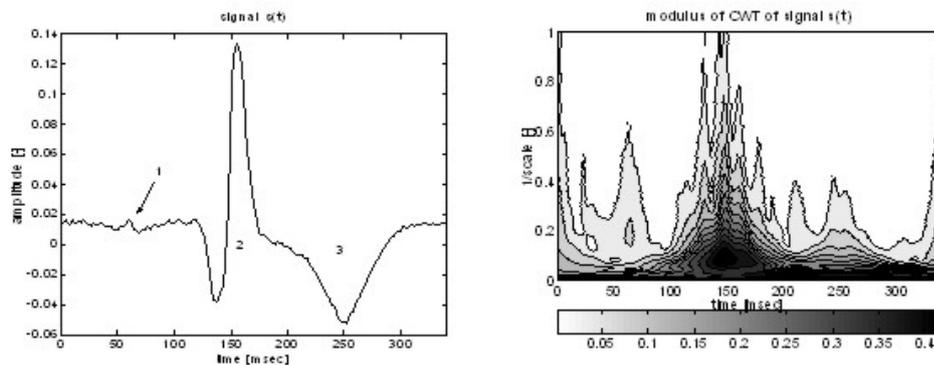
kde ε je malé číslo. Příklad konturového obrázku je vyobrazen na Obr. 2(b). Bereme v úvahu pouze tu část kontury, která je nejbližší vysokým frekvencím (4). Takovou konturu nazýváme konturovou obálkou EC a je definována jako

$$EC(b) = \min_{a \in A, C_L(a,b) > 0} [a] \quad (4)$$

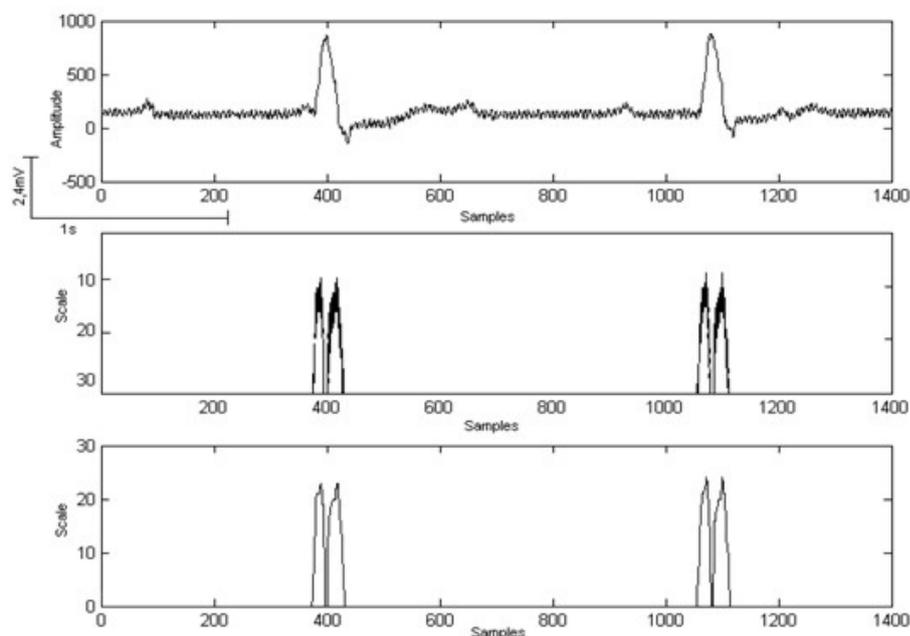
pro všechna b . Konturová obálka EC je jednorozměrná funkce. Příklad konturové obálky je na Obr. 2(c).

EC je dále zpracovávána klasickými algoritmy. V předložené metodě je EC vyhlazena Lyyňovou dolní propustí definovanou vztahem

$$H(z) = \frac{1 - z^{-7}}{1 - z} \quad (5).$$



Obr. 1. (a) EKG signál a jeho (b) absolutní hodnota CWT. P-vlna (1), QRS-komplex (2) a T-vlna (3) jsou vyznačeny.



Obr. 2. Vytvoření konturové obálky: (a) EKG signál, (b) řez časově-frekvenčním spektrem signálu, (c) konturová obálka.

Filtrovaná obálka EC je použita pro rozměřování signálových komponent (Obr. 2 EC s úrovní řezu pro QRS detekci, Obr. 3 EC s úrovní řezu pro další komponenty).

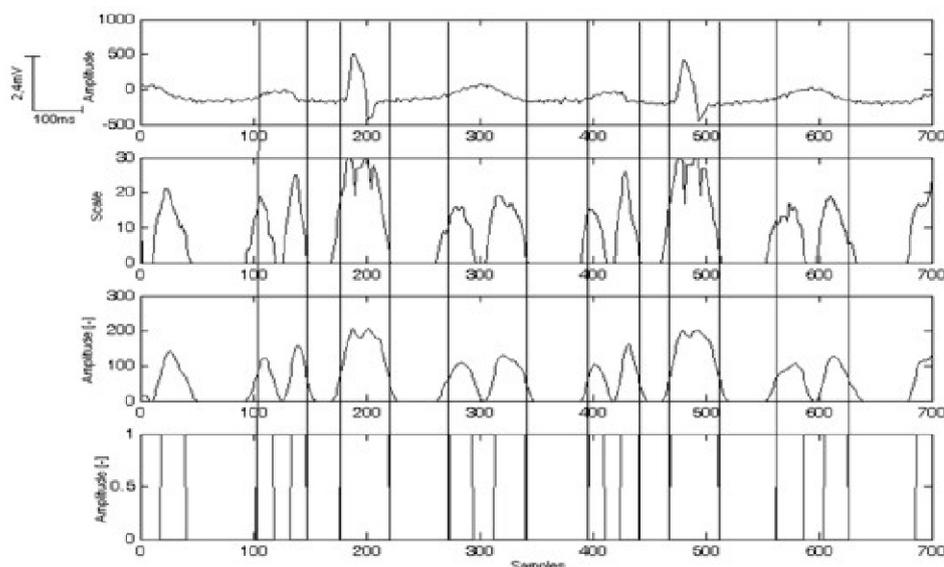
Konturová obálka je prahována za účelem odstranění šumu a potlačení rozšiřujícího efektu vlivem filtrace dolní propustí (Obr. 3(c)). Malé hodnoty jsou nulovány a nadprahovým hodnotám je přiřazena hodnota jedna. Poté vzestupné a sestupné hrany reprezentují začátky a konce jednotlivých vln – P-vlny, QRS-komplexu a T-vlny (Obr. 3(d)). Pro detekci QRS komplexu je použita vysoká úroveň řezu k

falešných detekcí (Obr. 2) a nízká úroveň řezu pro zviditelnění dalších EKG vln v konturové obálce (Obr. 3). Vnitřní hrany jsou odstraněny a stanoveny začátky a konce vln.

APC/PVC klasifikátor

APC/PVC klasifikátor rozlišuje APC (předčasné atriální kontrakce) a PVC (předčasné komorové kontrakce).

Pro detekci APC analyzujeme oblast před detekovaným QRS komplexem.



Obr. 3. Detekce P-vln, QRS-komplexu a T-vln s použitím Haarovy vlnky: (a) EKG signál, (b) konturová obálka, (c) filtrovaná konturová obálka, (d) detekční hrany.

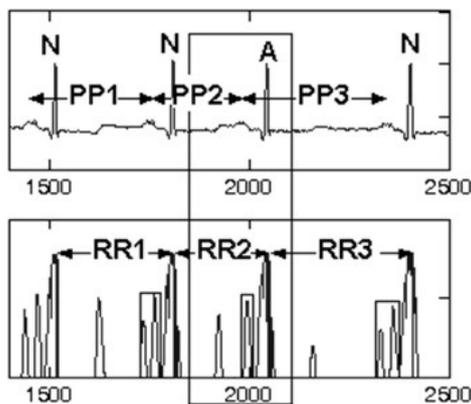
Předčasná P vlna může být detekována v konturové obálce s použitím jednoduchých vlnek. Obr. 4 dokumentuje použití Haarovy vlnky, měřítka 1-32 a úroveň řezu:

$$L = 0.125 \text{ mean}(WT_{abs}(a, b)) \quad (6).$$

Před QRS komplexem jsou v konturové obálce zřetelné struktury, které představují P vlnu, která má nepravidelný P-P interval nebo abnormální tvar. Tato kontura je rovněž filtrována k dosažení lepších výsledků.

Detekční hrany jsou dále zpracovány rozhodovacím algoritmem.

Detekovaný QRS komplex má normální šířku a tak můžeme jednoduše rozlišit APC od PVC.



Obr. 4. Předčasná atriální kontrakce: (a) EKG signál, (b) konturová obálka.

Analýza šířky QRS komplexu v konturové obálce byla provedena s následujícími parametry: $L=8 \cdot \text{mean}(WT_{abs}(a, b))$ měřítko 1-32 a vlnkou typu mexický klobouk. Analýza šířky je dokumentována na Obr. 5.

Široké QRS komplexy předčasné komorové kontrakce mají také širší odezvu v konturové obálce.

(Obr. 5(a) versus Obr. 5(b)). EC je filtrovaná, prahována a dále zpracována klasickými algoritmy.

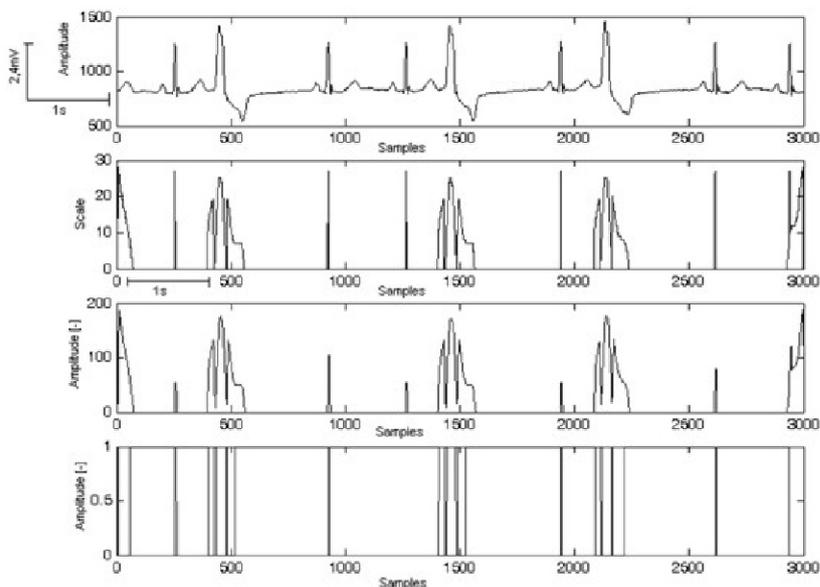
Dalším analyzovaným parametrem je R-R interval.

Kluzavé okno tří RR intervalů $[RR_1, RR_2, RR_3]$ a průměrný RR interval jsou použity ke klasifikaci prostředního RR_2 intervalu (Obr. 4). R-R interval obklopující předčasnou komorovou kontrakci je obvykle dvojnásobný než základní interval (Obr. 6) a delší než dvakrát pro APC (Obr. 4). Kombinují se různá kritéria pro RR interval, P-P interval, tvar P vlny a šířka QRS komplexu.

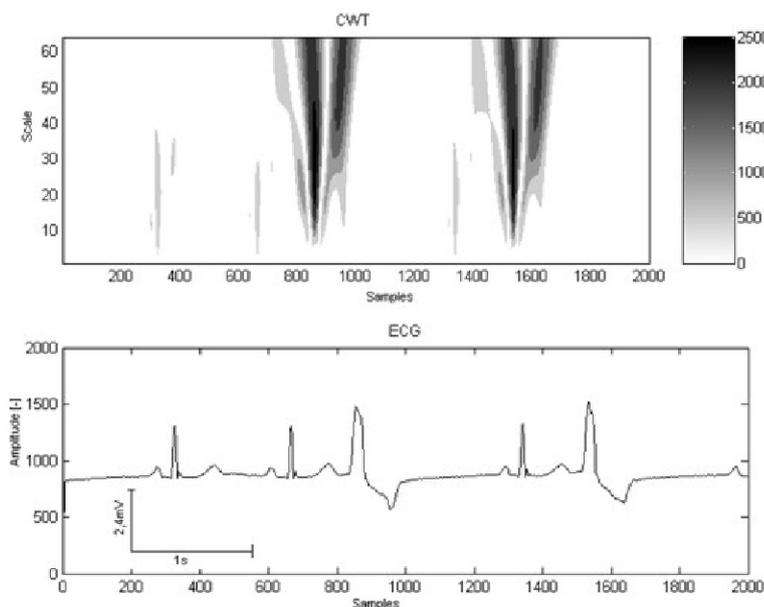
Klasifikátor normálního a abnormálního rytmu

Tato metoda ukazuje, že můžeme nalézt rozdíly mezi abnormálními a normálními stahy v matici vlnkových koeficientů. Byla použita vlnka typu mexický klobouk a měřítka 1-64. Maximum umocněných koeficientů spojitě vlnkové transformace v oblasti QRS komplexu u sinusového rytmu je obvykle v rozsahu $a = 1$ a $a = 10$ a příliš se mezi sebou neliší. Umocněné koeficienty CWT předčasných komorových kontrakcí mají větší amplitudu a maximum koeficientů v QRS oblasti je ve vyšším měřítku než pro normální stahy a šířka útvaru je větší. Rozdíly mohou být porovnány na Obr. 6, kde úzké komplexy normálních stahů mají užší odezvu a menší amplitudu v matici CWT koeficientů než předčasné komorové kontrakce.

Pozice maxima CWT koeficientů v oblasti QRS komplexu pro PVC je na měřítkové ose větší než 18.



Obr. 5. Analýza šířky: (a) signál č. 119, (b) konturová obálka, (c) filtrovaná konturová obálka, (d) detekční hrany.



Obr. 6. Rozlišování normálních a předčasných komorových kontrakcí: (a) umocněná absolutní hodnota CWT koeficientů – vlnka mexický klobouk, (b) EKG signál č. 119.

Každá změna ve velikosti maxima, pozici maxima na měřítkové ose a šířce mezi stahy může signalizovat abnormalitu nebo rušení.

RBBB (blokady pravého Tawarova raménka) mají maximum koeficientů pro nízká měřítka, obvykle stejně jako SR, ale mohou být snadno odlišeny šířkou QRS-komplexu.

Výsledky

Algoritmy byly vyvinuty v prostředí MATLAB s využitím Wavelet Toolboxu. Signály byly testovány na vybraných záznamech z MIT-BIH arytmiické databáze vytvořené Massachusetts Institute of Technology (MIT) a Boston's Beth Israel Hospital (BIH). Všechny analyzované signály obsahují popis každého srdečního stahu využívaného pro hodnocení detekční účinnosti. Analýza byla provedena na svodu MLII (modifikovaný končetinový svod II). Signály jsou vzorkovány frekvencí 360 Hz.

Detekce QRS-komplexu

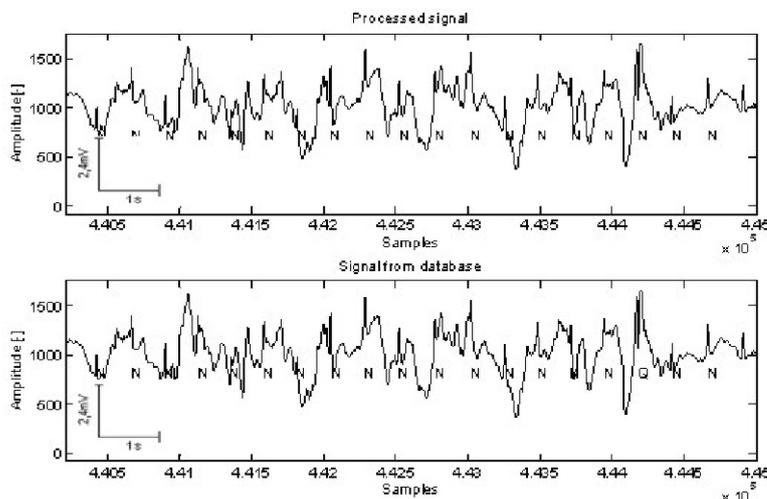
Algoritmus byl testován na 48 signálech (č. 100-234) z knihovny MIT-DB v plné délce (každý přibližně 30 minut, 650 000 vzorků). Celkem více než 24 hodin signálů obsahujících 99555 QRS-komplexů.

Průměrná detekční účinnost byla 99,45 %, nejhorší výsledek byl pro signál č. 207 (88,63 %), který obsahuje dlouhé úseky komorových fibrilací, kde je obtížné rozpoznat dílčí QRS-komplexy. Detektor měl detekční účinnost nad 99,9 % u 23 signálů, a nad 99 % u 43 signálů.

Obr. 7 dokumentuje robustnost algoritmu. Horní panel zobrazuje původní signál s detekovanými stahy (N pod křivkou signálu).

Dolní panel zobrazuje stejný signál s původním popisem z MIT-BIH databáze.

V tomto případě nebyl žádný QRS komplex vynechán nebo falešně detekován.



Obr. 7. Detekce QRS-komplexu – signál č. 105: (a) výsledky, (b) databáze.

APC/PVC klasifikátor

Klasifikace byla provedena na 34 záznamech z MIT-DB databáze. Pro studium byly vybrány následující kategorie: sinusový rytmus (SR), předčasné komorové kontrakce (PVC), a předčasné atriální kontrakce (APC). Klasifikace nebyla jednoduchá díky rozmanitosti morfologie SR, PVC, a APC stahů, což můžeme vidět na záznamech. Průměrná úspěšnost této metody v klasifikaci SR, PVC a APC stahů byla 94,6 %. Úspěšnost je prezentována jako poměr pravdivě pozitivních detekcí k celkovému počtu stahů. Počet testovaných SR, PVC, a APC stahů a úspěšnost detekce je shrnutá v následující tabulce. Přesnost detekce SR, PVC, nebo APC stahů v tabulce je reprezentována poměrem pravdivě pozitivních detekcí k počtu korespondujících stahů.

Tab. 1. Klasifikační účinnost pro APC/PVC klasifikátor.

Rytmus	Databáze	Algoritmus	Úspěšnost (%)
SR	69264	66821	96,1
PVC	5383	3807	71,2
APC	917	271	30,4

Bylo otestováno celkem 34 signálů z MIT-DB knihovny, kde se PVC a APC objevují nejčastěji.

Systém popsáný v tomto článku je nezávislý na pacientovi: není potřeba ruční segmentace vzorků každého nového typu stahu před a během analýzy.

Získané výsledky ukazují, že spojitá vlnková transformace umožňuje analýzu arytmií a pomáhá nacházet vlastnosti pro klasifikaci arytmií.

Klasifikátor normálního a abnormálního rytmu

Algoritmus byl testován pouze na 40 signálech z MIT-DB knihovny s PVC, APC, RBBB, LBBB (blokáda levého Tawarova raménka) a SR.

Můžeme jednoduše rozlišit PVC mezi RBBB, SR, APC a dalšími arytmiemi s úzkým QRS komplexem hledáním polohy maxima vlnkových koeficientů na měřítkové ose. Tato metoda má průměrnou úspěšnost 96,81 % pro detekci PVC mezi APC, SR, RBBB. Byly testovány také signály s LBBB, ale zde nebyly dosaženy hodnověrné výsledky. Tento algoritmus je také nezávislý bez potřeby manuální segmentace nebo manuální klasifikace dat.

Diskuze a závěr

Tyto metody byly shledány vhodnými pro pokračující analýzu. Spolupráce s kardiologií je nezbytná pro další zdokonalování. Úspěšnost detekce QRS komplexu je srovnatelná se současně publikovanými výsledky. Klasifikační algoritmus umí rozpoznat jasné případy abnormality a typické průběhy, ale stahy na hranici kategorií a atypické změny nejsou klasifikovány příliš úspěšně.

APC/PVC klasifikátor slibuje dobré výsledky s některými omezeními: klasifikace založená na detekci vln a měření důležitých intervalů v konturové obálce může rozhodovat o abnormálních intervalech a hodnotách, ale nemůže přesně rozhodnout o typu arytmiie. Tato informace nemůže být porovnána s databází. A toto rozhodování není možné bez praktické lékařské zkušenosti.

Hledání skrytých rozdílů mezi nemocemi a vlnkovými koeficienty je slibná technika. Klasifikátor normálního a abnormálního rytmu ukazuje, že rozdíly mohou být nalezeny. V matici vlnkových koeficientů můžeme vidět, že RBBB má obvykle maximum koeficientů v dolním rozsahu měřítek (1:64), stejně jako SR (ale můžeme je separovat na základě šířky),

LBBB v prostřední části rozsahu a PVC v horní části. Můžeme s velkou přesností určit PVC v matici koeficientů. Mají maximum ve vyšších měřících a širokou odezvu ve vlnkovém obraze. Toto kritérium dává dobré výsledky při rozpoznávání PVC mezi SR, RBBB, APC, NPB (nodální předčasný stah), JEB (junctional escape beat) a dalšími arytmiemi s úzkými komplexy. Toho může být využito jako příznaku pro automatické výpočty a trénovací metody. CWT transformace nabízí detailní analýzu za cenu zvýšení výpočetní náročnosti a využití paměti. Vytvoření obecného algoritmu úspěšného na všech typech EKG stahů je obtížný problém. Automatické výpočetní a trénovací metody založené na umělé inteligenci, skrytých Markovových modelech a dalších, by mohly vyřešit popsané problémy. Můžeme jednoduše detekovat odchylku od sinusového rytmu, široké a hluboké PVC, ale stále jsou zde variace, kde je vstup člověka nezbytný.

Poděkování

Výzkum vznikl za podpory grantu GACR č. 102/07/1473 a výzkumného záměru Vysokého učení technického v Brně MSM 0021630513.

Literatura

- [1] Nabney I. T., Evans D. J., Tenner J., Gamlyn L.: Benchmarking Beat Classification Algorithms. *Computers in Cardiology*, 2001, 529-532.
- [2] Sahambi J. S., Prasad G. K.: Classification Of ECG Arrhythmias Using Multi-Resolution Analysis And Neural Networks. In *Proceedings of Convergent Technologies for the Asia-Pacific Region Conference*, Bangalore, India, IEEE 2003, Vol. 1.
- [3] Strang G., Nguyen T.: *Wavelets and filter banks*. Wellesley-Cambridge Press, USA, 1996.
- [4] Kohler B., Hennig C., Orglmeister R.: The principles of software QRS detection. *IEEE Eng. Med. Biol. Mag.*, 2002, pp. 42-54.
- [5] Small M., Yu D. J., Grubb N., Simonotto J., Fox K. A. A., Harrison R. G.: Automatic identification and recording of cardiac arrhythmia. *Comp. Cardiol.*, 2000, vol. 27, pp. 355-358.
- [6] Unser M.: A Practical Guide to the Implementation of the Wavelet Transform. In: Aldroubi, A., Unser M. *Wavelets in Medicine and Biology*. CRC Press, New York, 1996.
- [7] Braunwald E., Zipes D. P., Libby P.: *Heart disease*. W. B. Saunders Company, 2001.
- [8] Chen S. W., Clarkson P. M., Fan Q.: A robust sequential detection algorithm for cardiac

- arrhythmia classification. *IEEE Trans. Biomed. Eng.*, 1996, vol. 43, pp. 1120-1125.
- [9] Guvenir H. A., Acar B., Demiroz G., Cekin A.: A supervised learning algorithm for arrhythmia analysis. *Comp. Cardiol.*, 1997, vol. 24, pp. 433-436.
- [10] Dickhaus H., Heinrich H.: Classifying biosignals with wavelet networks. *IEEE Eng. Med. Biol. Mag.*, 1996, pp. 103-104.
- [11] Provazník I., Bardanova J.: *Delineation of ECG Waves Using Complex-Valued Continuous Wavelet Transform*. Springer Berlin Heidelberg, 2008, vol. 18, pp. 158-161, ISSN 1680-0737.
- [12] Chesnokov Yu. V., Chizhikov V. I.: Continuous Wavelet Transformation in Processing Electrocardiograms in Ventricular Arrhythmia. *Journal Measurement Techniques*. Springer New York, 2004, no. 4, vol. 47, pp. 417-421, ISSN 0543-1972.
- [13] Shyu L. Y., Wu Y. H., Hu W.: Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. *IEEE Trans. Biomed. Eng.*, 2004, vol. 51, pp. 1269-73.
- [14] Romero Legaretta I., Addison P. S., Reed M. J., Grubb N. R., Clegg G. R., Robertson C. E., Watson J. N.: Continuous wavelet transform modulus maxima analysis of the electrocardiogram: Beat-to-beat characterisation and beat-to-beat measurement. *Int. J. Wavelets, Multiresolution and Information Processing*, 2005, vol.3(1), pp. 19-42.
- [15] Addison P. S.: Wavelet transforms and the ECG: a review. *Physiol. Meas.*, 2005, vol. 26.
- [16] Addison P. S.: *The Illustrated Wavelet Transform Handbook: Applications in Science, Engineering, Medicine and Finance*. Institute of Physics Publishing, Bristol, 2002.
- [17] Mahmoodabadi S. Z., Ahmadian A., Abolhasani M. D.: ECG Feature Extraction Using Daubechies Wavelets. *Proceedings of the Fifth IASTED International Conference, Visualization, Imaging, and Image Processing*, Benidorm, Spain. September 7-9, 2005.
- [18] Daqrouq K., Abu-Isbeih I. N.: Arrhythmia Detection using Wavelet Transform. *EUROCON, 2007. The International Conference on Computer as a Tool*, pp.122-126, 9-12 Sept. 2007.

Kontakt

Prof. Ing. Ivo Provazník, Ph.D.

Ústav biomedicínského inženýrství,
Fakulta elektrotechniky a komunikačních
technologií

Vysoké učení technické v Brně
Kolejní 4, Brno 61200

Česká republika

e-mail: provaznik@feec.vutbr.cz

Ischemická choroba srdeční u mužů v České republice, 1980-2004

Jindra Reissigová, Marie Tomečková

Centrum Biomedicínské Informatiky, Ústav Informatiky AV ČR, v.v.i., Praha, Česká republika

Pozadí: Česká republika patří mezi země s vysokou úmrtností na ischemickou chorobu srdeční (ICHS). Cílem bylo analyzovat specifickou úmrtnost na ICHS podle věku, období a kohorty narození u mužů v České republice.

Design a metody: Analyzovala se národní data úmrtnosti mužů (30-74 let) v České republice v letech 1980-2004. K hodnocení vlivu věku, období a kohorty na úmrtnost se použil Poissonův regresní model.

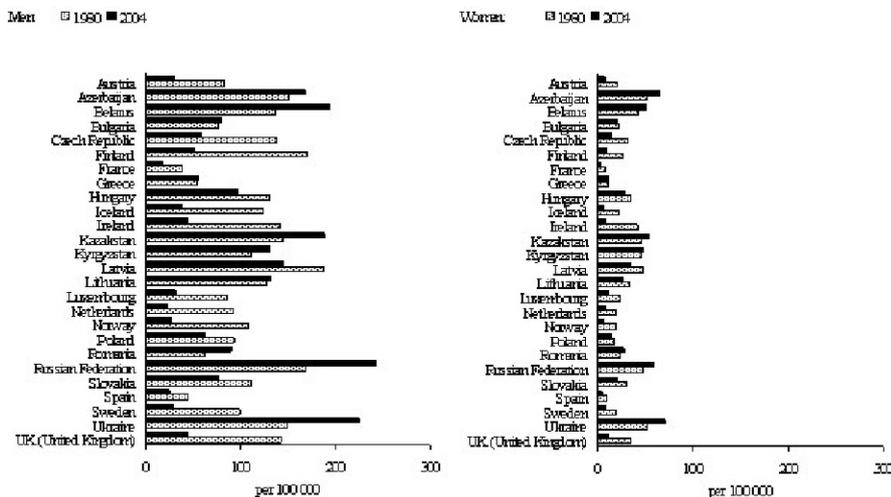
Výsledky: Úmrtnost na ICHS u mužů adekvátně popisoval model zahrnující věk, období a kohortu ($p=0,121$). Poměr úmrtností věkových skupin 45-49 k 40-44 byl o 17 % nižší než odpovídající poměr věkových skupin 40-44 k 35-39. Poměr relativních rizik v obdobích 1995-1999 k 1990-1994 byl o 19 % nižší než odpovídající poměr v obdobích 1990-1994 k 1985-1989. Poměry relativních rizik mezi sousedícími kohortami kolísaly kolem hodnoty 1.

Závěry: Riziko úmrtnosti na ICHS bylo u mužů v České republice významně závislé na věku (riziko se zvyšovalo s věkem), kalendářním období (riziko se snížilo v posledních letech) a kohortě narození (riziko se snižovalo s kohortou narození). Zásadní pozitivní změna ve vývoji úmrtnosti na ICHS nastala po pádu komunismu v roce 1989 a u mužů po 40. roce věku.

Klíčová slova: ischemická choroba srdeční, kardiovaskulární, úmrtnost, riziko

Úvod

Na přelomu 21. století způsobují kardiovaskulární onemocnění (KVO) v zemích Evropské unie mezi 26 % (Francie) až 62 % (Bulharsko) úmrtí u mužů a mezi 31 % (Francie) až 71 % (Bulharsko) úmrtí u žen [1]. KVO zůstávají hlavní příčinou úmrtí u mužů a žen ve všech evropských zemích, vyjma mužů ve Francii, Holandsku a ve Španělsku. Více než jedna třetina z úmrtí na KVO jsou úmrtí na ischemickou chorobu srdeční (ICHS), jež je sama nejčastější příčinou úmrtí



Obr. 1. Na evropskou populaci věkově standardizovaná úmrtnost na ischemickou chorobu srdeční v některých evropských zemích, podle pohlaví, 0-64 let, 1980¹, 2004.

¹ V Ázerbájdžánu, Bělorusku, Kazachstánu, Kyrgyzstánu, Litvě a Ukrajině je znázorněna úmrtnost v roce 1981.

V Evropě. Obr. 1 zobrazuje věkově standardizovanou úmrtnost na ICHS v některých zemích Evropy.

V průběhu let 1970-2000 úmrtnost na KVO v zemích západní Evropy klesá téměř lineárně [2]. V zemích východní Evropy úmrtnost na KVO dosáhla svého maxima v letech 1990-1994, s následováním poklesem v Polsku, Maďarsku a v baltických státech (Estonsko, Lotyšsko, Litva). Pozoruhodný opačný trend ve vývoji úmrtnosti byl také pozorován v České republice (ČR) [3]. Hlavní vliv na pokles úmrtnosti na KVO v ČR měl nejspíše klesající trend výskytu hlavních rizikových faktorů KVO pozorovaný v letech 1985-2000/01 [4].

Nicméně, v roce 2006 byla onemocnění oběhového soustavy nejčastější příčinou úmrtí v ČR (druhou nejčastější příčinou úmrtí byly zhoubné novotvary, třetí příčinou byly vnější příčiny) [5]. Jsou zodpovědná za 50,3 % úmrtí ze všech úmrtí. Hlavní příčinou úmrtí z kardiovaskulárních onemocnění je ICHS. Zapříčinila 47,1 % ze všech úmrtí na KVO u mužů a 40,8 % u žen. Cílem tohoto

článku je analyzovat úmrtnost na ICHS u mužů v České republice v letech 1980-2004.

Materiál and metody

Zdrojová data

Počet onemocnění ICHS (kódy D410-D414 podle 8. revize Mezinárodní klasifikace nemocí (MKN-8), kódy 410-414 podle MKN-9, kódy I20-I25 podle MKN-10) a střední stav populace v 5letých věkových skupinách (0-4, 5-9, ..., 80-84, 85+) v letech 1980-2004 jsou převzaty z publikací „Pohyb obyvatelstva v České republice“ ročně vydávaných Českým statistickým úřadem. Úmrtnost byla věkově standardizovaná metodou přímé standardizace na evropskou populaci [6].

Statistické metody

Na úmrtnost na ICHS u mužů ve věku 30-74 let v ČR v letech 1980-2004 se aplikoval model „věku-období-kohorty“ (APC, age-period-cohort) (pozn. úmrtnost v mladších věkových skupinách byla nízká a spolehlivost udávaných příčin úmrtí u starších osob může být diskutabilní, proto byly tyto věkové skupiny z analýzy vyloučeny).

Počet úmrtí na ICHS se modeloval pomocí log-lineární Poissonovy regrese jako funkce věku při úmrtí, období úmrtí a kohorty narození. APC modelování se provádělo hierarchicky, jak bylo sumarizováno Arbynem a kol. [7] a jeho konečné výsledky jsou prezentovány v Tab. 1 ve Výsledcích. Vhodnost modelu (goodness of fit) byla vyhodnocována reziduální deviancí a chí-kvadrát testem. K vyhodnocení vlivu přidávaného efektu do modelu se porovnal rozdíl deviancí mezi daným model a modelem bez tohoto efektu s chí-kvadrát rozdělením se stupni volnosti (df) rovnými rozdílu df těchto dvou modelů. Adekvátnost modelu se dále vyhodnotila na základě Akaikeho informačního kritéria (AIC, Akaike's information criterion) a analýzou Pearsonových reziduí. K testování extra-Poissonova kolísání (extra-Poisson variation) se použil Cameronův-Trivediův (C-T) test.

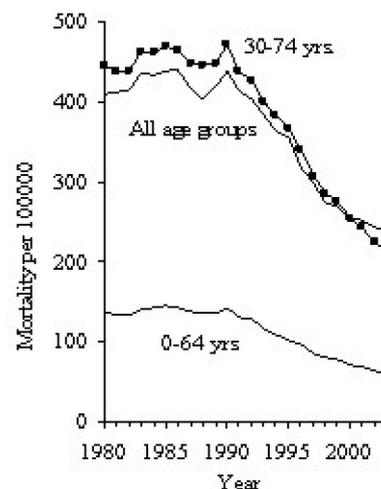
Je dobře známo, že mezi proměnnými věk ($a_i, i=1, \dots, I$), období ($p_j, j=1, \dots, J$) a kohorta ($c_k, k=1, \dots, K$) existuje lineární závislost: $k=i+j$. V modelu zahrnujícím současně věk, období a kohortu se lineární závislost projeví tak, že neexistuje jednoznačné řešení odhadů parametrů (pozn. odhadované počty úmrtí modelem jsou však stejné pro různé odhady koeficientů). K získání jednoznačných odhadů je nutné klást na odhadované parametry různá omezení. Různá omezení však vedou k různým odhadům trendů věku, období a kohorty. V našem případě se při odhadování parametrů věku, období a kohorty fixovaly dva parametry období (1985-1989, 1990-1994) a jeden parametr kohorty (1935-1944) k hodnotě 0 a také se odhadly takzvané druhé rozdíly parametrů podle Claytona a Schiffersa [8]. Druhé rozdíly se vyjadřují poměrem (rozdílů na logaritickém měřítku) dvou sousedních relativních hodnot (více ve Výsledcích). K provedení statistické analýzy byl použit R software verze 2.3.1.

Výsledky

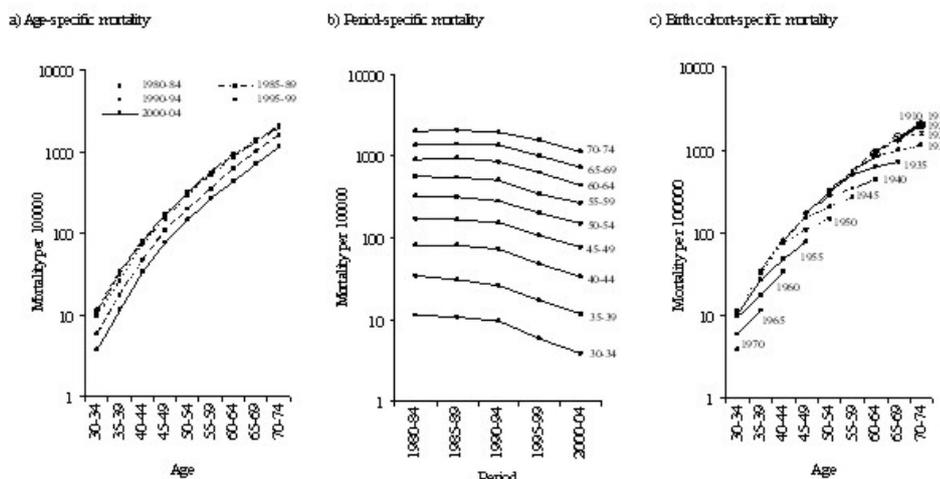
Obr. 2 ukazuje vývoj věkově standardizované úmrtnosti na ICHS u mužů v ČR v letech 1980-2004 ve všech věkových skupinách, ve věkové skupině 0-64 let a ve věkové skupině 30-74 let. Pro možnost srovnání s úmrtností v jiných evropských

státech (Obr. 1) je tato úmrtnost adjustována (přepočtena) na věkovou strukturu evropské standardní populace.

Obr. 3 znázorňuje (roční) specifickou úmrtnost podle věku, období a kohorty narození ve věkové skupině 30-74 let. Úmrtnost je znázorněna na logaritickém měřítku. Úmrtnost se zvyšovala s věkem (Obr. 3a), snižovala se v posledních letech (Obr. 3b) a snižovala se s kohortou narození (Obr. 3c). Jinými slovy, Obr. 3 znázorňuje úmrtnost postupně stratifikovanou podle věku, období a kohorty narození, kterou jsme měli k dispozici pro statistickou analýzu testující simultánně do jaké míry věk, období, kohorta narození ovlivňují úmrtnost.



Obr. 2. Na evropskou populaci věkově standardizovaná úmrtnost na ischemickou chorobu srdeční v České Republice, muži, ve všech a ve specifických věkových skupinách, 1980-2004.



Obr. 3. Specifická úmrtnost na ischemickou chorobu srdeční v České republice podle věku, období a kohorty narození, muži, 30-74 let, 1980-2004¹.

¹V grafu c) kohorta narození označuje střed 10letého intervalu, např. kohorta narození 1940 reprezentuje kohortu mužů narozených od 01/01/1935 do 31/12/1944.

Statistická analýza je prezentována v Tab. 1, která sumarizuje výsledky APC modelování. Vliv věku na úmrtnost byl významný (model 1). Modely 2 a 3a-3b byly významně lepší ve srovnání s modelem 1, avšak stále nepostačující. Nejvhodnější byl APC model (model 4). To znamená, že riziko úmrtí bylo významně závislé na všech třech efektech: věku, kalendářním období a kohortě narození. V APC modelu nebylo detekováno žádné extra-Poissonovo kolísání (jinými slovy rozdíl mezi průměrem a rozptylem byl nevýznamný) ($p=0,500$) a Pearsonova rezidua se pohybovala v rozpětí od -1,864 do 1,784.

Vlastní trendy vlivu věku, období a kohorty na úmrtnost jsou znázorněny na Obr. 4. Vliv věku se vyjádřil jako věkově specifická úmrtnost na 100 000, vliv periody a kohorty jako relativní riziko úmrtí na ICHS. Se zvyšujícím se věkem se zvyšovala úmrtnost na ICHS (Obr. 4a). Riziko úmrtí na ICHS se snižovalo v posledních letech (Obr. 4b, porovnáváno s lety 1985-1994, tj. referenční období) a s kohortou narození (Obr. 4c). Kohortová rizika úmrtí byla porovnávána s rizikem kohorty 1940 (referenční kohorta). Kohorta 1940 zahrnovala muže narozené kolem (1935-1944) počátku druhé světové války v roce 1939.

Vlastně jde o středovou kohortu a ve srovnání s ní starší a mladší kohorty vycházejí z méně dat, Obr. 3c (to se odráží v kolísavé šířce intervalů spolehlivosti, Obr. 4c).

Čísla v závorkách na Obr. 4 uvádějí poměr dvou sousedních hodnot. Připomeňme, že tyto poměry jsou stejné pro různé odhady vlivu věku, období a kohorty a řeší tedy problém nejednoznačného odhadu parametrů v APC modelu, jak bylo zmíněno v sekci Materiál a metody. Je důležité zdůraznit, že poměry neindikují rostoucí ani klesající trend. Hodnoty nižší než 1 vlastně vyjadřují, že došlo ke „zpomalení“ trendu a hodnoty vyšší než 1, že došlo k „zrychlení trendu“. Popišme nejdůležitější náhlé změny v trendech:

- Viz Obr. 4a): Poměr úmrtností ve věkové skupině 45-49 k věkové skupině 40-44 byl o 17 % nižší než poměr úmrtností věkové skupiny 40-44 k věkové skupině 35-39, a to protože zmíněný poměr úmrtností byl $\left(\frac{165.2}{87.7}\right) / \left(\frac{87.7}{38.6}\right) = 83\%$.

Následující poměry úmrtností byly 89 % pro věkovou skupinu 45-49 a 93 %-99 % pro věkovou skupinu 50 a více let. Souhrnně řečeno k největšímu zpomalení trendu úmrtností došlo u mužů ve věkové skupině 40-49 let.

- Viz Obr. 4b): Poměr relativních rizik období 1995-1999 k období 1990-1994 byl o 19 % nižší než poměr relativních rizik období 1990-1994 k období 1985-1989, protože zmíněný poměr rizik je

$$\left(\frac{0.81}{1.0}\right) / \left(\frac{1.0}{1.0}\right) = 81\%.$$

Jiné více než 10% změny trendu nebyly pozorovány. Takže k nevyššímu poklesu úmrtností došlo v letech 1995-1999 ve srovnání s předcházejícím vývojem.

- Viz Obr. 4c): U jednotlivých kohort nebyly pozorovány žádné náhlé změny relativního rizika; poměry relativních rizik mezi sousedními kohortami kolísaly od 0,92 do 1,06.

Diskuse

APC modelování je užívána metoda k analýze incidence a úmrtnosti z různých příčin od rakovin až po kardiovaskulární onemocnění. Je to deskriptivní nástroj především aplikovaný na národně registrovaná data. Pomáhá nejen

Tab. 1. Modelování úmrtnosti na ischemickou chorobu srdeční v závislosti na věku, období a kohortě, muži, 30-74 let, 1980-2004.

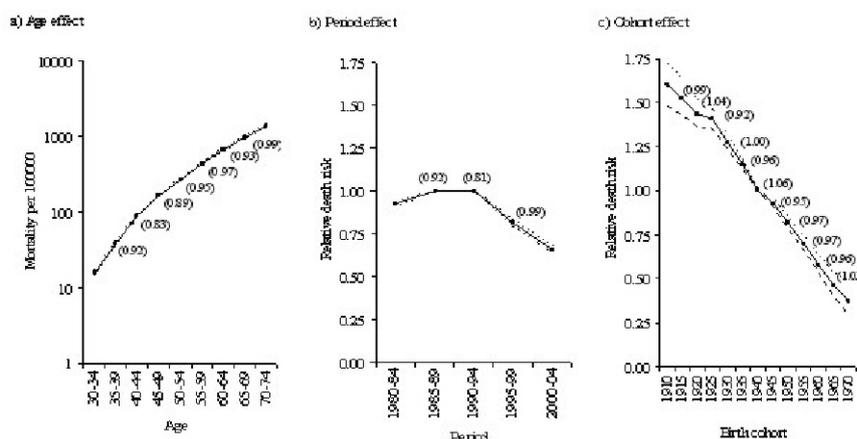
No. ¹	Model	Residual deviance (D)	Degree of freedom (df)	p-value	Compared with model	$\Delta(D)^2$	$\Delta(df)^3$	p-value	AIC ⁴
0	Interception	355388.0	44	<0.001					
1	Age	15148.0	36	<0.001	0	340240.0	8	<0.001	15600.6
2	Age-Drift	3255.5	35	<0.001	1	11892.5	1	<0.001	3710.1
3a	Age-Age*Drift	2922.5	27	<0.001	2	333.0	8	<0.001	3393.1
3b	Age-Period	388.2	32	<0.001	2	2867.3	3	<0.001	848.8
3c	Age-Cohort	1872.6	24	<0.001	2	1382.9	11	<0.001	2349.2
4	Age-Period-Cohort	28.7	21	0.121	3b	359.5	11	<0.001	511.3

¹Viz Obr. 2. Jestliže platí model 1, věkově specifické křivky úmrtnosti podle období jsou identické. Jestliže platí model 2, věkově specifické křivky úmrtnosti jsou rovnoběžné v konstantní vzdálenosti (tzv. drift). V případě modelu 3a, věkově specifické křivky úmrtnosti nejsou rovnoběžné; věkově specifická úmrtnost v sousedních obdobích se liší o konstantu specifickou pro každou věkovou skupinu. V případě modelu 3b, věkově specifické křivky úmrtnosti (nebo ekvivalentně křivky specifické úmrtnosti dle období) jsou rovnoběžné, ale v různých vzdálenostech. Jestliže platí model 3c, specifické křivky úmrtnosti podle kohort jsou rovnoběžné v různých vzdálenostech. Model 4 vyjadřuje, že věkově specifické křivky úmrtnosti nejsou rovnoběžné a jsou funkcí období a kohorty.

²Rozdíl reziduálních deviací.

³Rozdíl stupňů volnosti.

⁴Akaikeho informační kritérium.



Obr. 4. Odhadované vlivy věku, období¹ a kohorty narození² (s 95% intervalem spolehlivosti) na úmrtnost na ischemickou chorobu srdeční v České republice, muži, 30-74 let, 1980-2004 (čísla v závorkách označují poměr dvou sousedních relativních hodnot³).

¹Období 1985-1989 a 1990-1994 byla vybrána jako referenční období.

²Kohorta narození 1940 byla vybrána jako referenční; kohorta narození označuje střed 10letého intervalu, např. kohorta narození 1940 reprezentuje kohortu mužů narozených od 01/01/1935 do 31/12/1944.

³Například poměr úmrtností věkové skupiny 40-44 (87,7) k věkové skupině 35-39 (38,6) a úmrtností věkové skupiny 35-39 (38,6) k věkové skupině 30-34 (15,6) je $(87,7/38,6) / (38,6/15,6) = 0,92$.

modelovat ale i předpovídat výskyt onemocnění. V tomto článku jsme použili APC modelování k analýze úmrtnosti na ICHS u mužů v České republice.

KVO jsou příčinou skoro poloviny všech úmrtí v Evropě (48 %) a sama ICHS je nejčastější příčinou úmrtí v Evropě [1]. Primárně preventivní programy v mnoha zemích se pokoušejí snížit výskyt ICHS modifikací rizikových faktorů. V doposud uskutečněných studiích však takové intervence měly omezený vliv na úmrtnost,

jak je sumarizováno v literatuře [9,10]. Na druhé straně, jak je tam uvedeno, malý ale potenciálně důležitý prospěch (asi 10% snížení úmrtnosti na ICHS) nemusel být v těchto studiích odhalen.

Hlavními kardiovaskulárními rizikovými faktory jsou kouření, zvýšený krevní tlak, zvýšená hladina cholesterolu a diabetes mellitus. ČR patří k evropským zemím s vysokou úmrtností na ICHS, přestože úmrtnost na ICHS v ČR klesá od počátku 90. let minulého století (Obr. 1).

V této studii jsme použitím APC modelování zjistili, že úmrtnost u českých mužů byla významně závislá na věku úmrtí, na kalendářním období úmrtí a na kohortě narození.

Úmrtnost na ICHS se zvyšovala s věkem (Obr. 4a). Vliv věku na úmrtnost pravděpodobně odrážel závislost této úmrtnosti na ateroskleróze („tvrdnutí cév“ – hlavní základní příčina KVO), o které je známo, že může začít již v dětství a s věkem se zhoršuje. Pravděpodobně rozhodujícím věkem pro muže je věk kolem 40/50 let, kdy bylo pozorováno největší zpomalení nárůstu úmrtnosti (poměr úmrtností věkové skupiny 45-49 k 40-44 byl o 17 % nižší než odpovídající poměr věkové skupiny 40-44 k 35-39; poměr úmrtností věkové skupiny 50-54 k 45-49 byl o 11 % nižší než odpovídající poměr věkové skupiny 45-49 k 40-44). Relativní riziko úmrtí na ICHS se v poslední době snížilo (Obr. 4b). Vliv období na úmrtnost zrcadlí rizikové faktory a úroveň zdravotní péče, které působily v době úmrtí. Rizika v obdobích 1985-1989 a 1990-1994 se fixovala k hodnotě 1 (referenční období), neboť se dá očekávat, že riziko bylo v těchto letech kolem pádu komunismu v roce 1989 přibližně stejné. Pokles rizika na fatální ICHS v poslední době (ve srovnání s lety 1985-1994) se vysvětluje pádem komunismu, jenž měl za následek změnu životního stylu a moderní léčbu. Poměr rizik v letech 1995-1999 k 1990-1994 byl o 19 % nižší než odpovídající poměr rizik období 1990-1994 k 1985-1989.

Relativní riziko úmrtí na ICHS lineárně klesalo s kohortou narození v rozmezí od 0,38 do 1,60 (Obr. 4c). Nebyly pozorovány žádné větší změny trendu; poměry relativních rizik mezi sousedními kohortami byly blízko 1. Vliv kohorty narození na úmrtnost může být chápán jako vliv rizikových faktorů a environmentální expozice, které jsou typické pro danou generaci. Ke kohortě narození se často vztahuje etiologie.

Jedním z nedostatků našeho APC modelu je okolnost, že jsme neměli k dispozici data týkající se charakteristik jednotlivců. Nemohli jsme tak vztáhnout trendy věku, období a kohorty k trendům

kardiovaskulárních rizikových faktorů (kouření, krevní tlak atd.) a přesně prozkoumat čas od modifikace rizikových faktorů k poklesu úmrtnosti. Naše studie není studie analytická, ale studie deskriptivní a mohla by pomoci zdravotnickým pracovníkům k lepšímu porozumění vývoje úmrtnosti na ICHS v ČR a spekulacím o něm. Například naše data naznačují, že strmější pokles úmrtnosti se projevil v 5-10 letech (1995-1999) po modifikaci rizikových faktorů, o které se předpokládá, že začala kolem roku 1990. Dále, i když cílem naší studie nebylo předpovídat budoucí vývoj úmrtnosti na ICHS, dá se očekávat, že úmrtnost u mužů v ČR bude dále klesat za předpokladu, že trendy rizikových faktorů se nezmění. Pokles je především očekáván proto, že riziko úmrtí se snižuje s rokem narození a v posledních letech. APC modelování ICHS bylo v nedávné době aplikováno také ve Finsku [11] a Japonsku [12], u nichž je rovněž očekáván pokles úmrtnosti na ICHS. Výsledky obou studií zdůrazňují důležitost nedávného poklesu úmrtnosti na ICHS u středně starých osob. Našli jsme několik studií, které byly uskutečněny v minulém století a předpovídaly vývoj úmrtnosti do začátku 21. století. Například se analyzovaly vlivy věku, období a kohorty na úmrtnost na ICHS ve Švédsku v letech 1969-1993, s predikcí do roku 2003 [13]. Stálo by za to porovnat předpovídanou a pozorovanou úmrtnost v těchto studiích.

Závěrem shrňme, že riziko úmrtí na ICHS u českých mužů bylo významně závislé na věku úmrtí, kalendářním období a kohortě narození. Podstatné pozitivní změny v trendu úmrtnosti byly pozorovány po pádu komunismu v roce 1989 a u mužů ve věku 40-49 let. Za předpokladu, že trendy věku, období a kohorty se nezmění, dá se u českých mužů očekávat další pokles úmrtnosti na ICHS.

Literatura

- [1] European cardiovascular disease statistics 2008 edition. London: British heart foundation, 2008.
- [2] Kesteloot H., Sans S., Kromhout D.: Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *Eur Heart J* 2006; 27(1):107-13.
- [3] Levi F., Lucchini F., Negri E., La Vecchia C.: Trends in mortality from cardiovascular and

cerebrovascular diseases in Europe and other areas of the world. *Heart* 2002; 88(2):119-24.

- [4] Cifkova R., Skodova Z.: Longitudinal trends in major cardiovascular disease risk factors in the Czech population. *Cas Lek Cesk* 2004; 143(4):219-226 (in Czech).
- [5] Deaths 2006. Institute of Health Information and Statistics of the Czech Republic, Prague 2007. [6] Doll R., Cook P.: Summarizing indices for comparison of cancer incidence data. *Int J Cancer*, 1967; 2(3):269-279.
- [7] Arbyn M., Van Oyen H., Sartor F., Tibaldi F., Molenberghs G.: Description of the influence of age, period and cohort effects on cervical cancer mortality by loglinear Poisson models (Belgium, 1955-94). *Arch Public Health* 2002; 60:73-100.
- [8] Clayton D., Schifflers E.: Models for temporal variation in cancer rates. II: Age-period-cohort models. *Stat Med* 1987; 6(4):469-481.
- [9] Ebrahim S.: Systematic review of randomised controlled trials of multiple risk factor intervention for preventing coronary heart disease. *BMJ* 1997; 314(7095):1666-1674.
- [10] Ebrahim S., Beswick A., Burke M., Davey Smith G.: Multiple risk factor interventions for primary prevention of coronary heart disease. *Cochrane Database Syst Rev* 2006; (4):CD001561
- [11] Huovinen E., Härkänen T., Martelin T., Koskinen S., Aromaa A.: Predicting coronary heart disease mortality-assessing uncertainties in population forecasts and death probabilities by using Bayesian inference. *Int J Epidemiol* 2006; 35(5):1246-52.
- [12] Ma E., Iso H., Takahashi H., Yamagishi K., Tanigawa T.: Age-Period-Cohort analysis of mortality due to ischemic heart disease in Japan, 1955 to 2000. *Circ J* 2008; 72(6):966-72.
- [13] Peltonen M., Asplund K.: Age-period-cohort effects on ischaemic heart disease mortality in Sweden from 1969 to 1993, and forecasts up to 2003. *Eur Heart J* 1997; 18(8): 1307-1312.

Poděkování

Práce byla podporována projektem 1M06014 MŠMT ČR.

Kontakt

RNDr. Jindra Reissigová, Ph.D.
Centrum biomedicínské informatiky
Ústav informatiky AV ČR, v.v.i.
Pod Vodárenskou věží 2
Praha 8 182 07
Česká republika
e-mail: reissigova@euromise.cz

Model reprezentace znalostí v doporučeních

David Buchtela, Jan Peleška, Arnošt Veselý, Jana Zvárová, Miroslav Zvolský

Centrum biomedicínské informatiky, Ústav informatiky Akademie věd ČR, v.v.i., Praha, Česká republika

Souhrn: Model reprezentace znalostí obsažených v lékařských doporučeních GLIKREM (GuideLines Knowledge REpresentation Model) vychází z GLIF modelu, který byl publikován ve specifikaci GLIF3.5. GLIKREM obsahuje některé změny a rozšíření definice a implementace původního GLIF modelu. Cílem tohoto příspěvku je popis znalostního modelu GLIKREM, jeho konstrukce, implementace v XML, realizace datového rozhraní a použití výsledného modelu.

Klíčová slova: reprezentace znalostí, GLIF model, doporučení

Úvod

Lékařská doporučení jsou potřebná pro podporu rozhodování v klinické praxi. Hlavním záměrem je zvýšení kvality péče o pacienty a snížení nákladů. Přístup k informacím obsaženým v konvenční podobě doporučení (ve formě volného textu) může být bohužel obtížný. Předpokladem pro vývoj systémů podpory rozhodování je tedy vytvoření počítačově zpracovatelné reprezentace znalostí obsažených v lékařských doporučeních. Proto se vývojem takové reprezentace doporučení zabývá řada výzkumných skupin.

Pravděpodobně nejznámějším jazykem pro reprezentaci lékařských doporučení v systémech podpory rozhodování je Ardenská syntaxe [1]. Jedná se o formalismus založený na pravidlech pro kódování individuálních klinických pravidel jako medicínské logické moduly. Řada přístupů sdílí hierarchickou dekompozici doporučení ve formě sítí dílčích úkolů, které se rozkládají v čase [2]. Všechny dále uvedené metody modelování mohou kombinovat jednotlivé kroky doporučení v orientovaných cyklických grafech.

Asbru je vyvíjen ve spolupráci Univerzity Bena Guriona a Vídeňské technologické univerzity [3]. Jedná se časově

orientovaný jazyk založený na specifikaci záměrů a základní kostry plánu, který je používán pro reprezentaci klinických protokolů.

EON byl vyvinut na Stanfordské univerzitě a poskytuje sadu modelů a programových komponent pro vytváření aplikací postavených na doporučeních [4]. EON používá úkolově orientovaný přístup k definici služeb podpory rozhodování, které mohou být implementovány různými technikami [5]. V EON architektuře se používá prostředí Protége2000 ke konstrukci informačního modelu pacientových dat, model medicínské specializace a model doporučení, který formalizuje znalosti potřebné k realizaci doporučení s ohledem na klinická rozhodnutí a aktivity.

GUIDE je část modelovacího a výkonného rámce, který je vyvíjen na Univerzitě v Pavii [6]. Podporuje integraci modelů doporučení do pracovních procesů v organizaci použitím analyticko rozhodovacích modelů jako jsou rozhodovací stromy a diagramy vlivu, a simulaci implementace doporučení v podmínkách Petriho sítí.

PRODIGY byl vyvinut v Univerzitě v Newcastlu nad Tynem [7]. Cílem projektu PRODIGY je vytvoření nejjednoduššího rychleji pochopitelného modelu reprezentace třídy doporučení. Týmy klinických lékařů zde používají vývojové prostředí Protégé [8] pro kódování tří komplexních doporučení řízení chronických nemocí.

PROforma byla vyvinuta v Moderní výpočetní laboratoři výzkumu rakoviny v Britském království [10]. Kombinuje logické programování a objektově orientované modelování, formálně zakotveném v jazyce R2L. PROforma podporuje čtyři typy činností: akce, složené plány, rozhodování a dotazy.

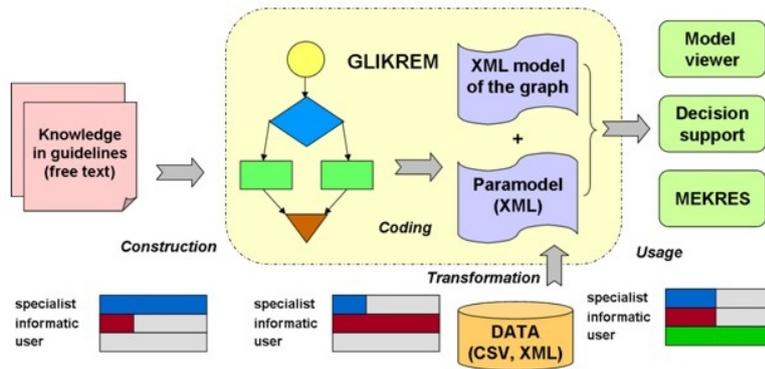
Všechny činnosti sdílejí atributy popisující cíle, kontrolní toky, předpoklady a následné podmínky.

GLIF (Guideline Interchange Format) verze 3 byl vyvinut ve spolupráci Univerzity Columbia, Harvardské, McGillovy a Stanfordské univerzity (InterMed). GLIF byl publikován ve specifikaci GLIF3.5 [10, 11, 12]. Jeho výrazový jazyk byl původně založen na Ardenské syntaxi a pro popis kritérií a výrazů byl používán jazyk GEL (Guideline Expression Language). V současnosti je používán jazyk GELLO, tj. rozšířený objektově orientovaný jazyk, který také podporuje množinu funkcí z jazyka GEL [13].

GLIF model byl také použit jako model reprezentace počítačových klinických doporučení v projektu SAPHIRE [14]. Systém SAPHIRE průběžně monitoruje pacienty prostřednictvím jednoúčelových agentů a pomocí inteligentního systému podpory rozhodování pomáhá profesionálům v zajištění zdravotní péče. Cílem tohoto příspěvku je návrh znalostního modelu doporučení GLIKREM (Guidelines Knowledge Representation Model). GLIKREM je založený na GLIF modelu, oproti původnímu modelu obsahuje navíc rozšíření v definici a implementaci. Příspěvek popisuje vlastní znalostní model doporučení, jeho konstrukci, implementaci v XML (eXtensible Markup Language), realizaci datového rozhraní a použití výsledného modelu. Znalostní model GLIKREM je dále použit v systému reprezentace lékařských znalostí (MEKRES), který v současnosti vyvíjen [15].

Model reprezentace znalostí GLIKREM

Celý proces konstrukce modelu znalostí (GLIKREM) z volného textu lékařských doporučení, jeho následnou reprezentaci v XML a použití výsledného modelu znázorňuje obrázek 1 [16].

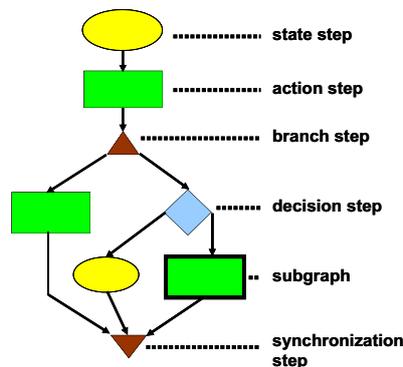


Obr. 1. Proces konstrukce, kódování a použití GLIKREM.

Ve fázi konstrukce modelu z textových doporučení je důležité najít procesní strukturu doporučení (rozhodovací algoritmus) a všechny podstatné parametry modelu a jejich vzájemné vztahy [17].

Ve fázi kódování je grafický model doporučení zakódován v XML.

Mimoto je vytvořen i seznam základních a odvozených parametrů modelu (paramodel), který slouží jako rozhraní mezi modelem a reálnými daty [18].



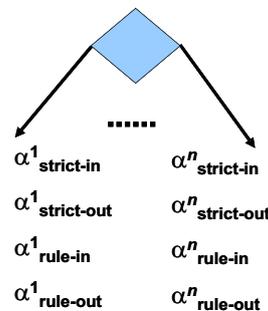
Obr. 2. Prvky grafického modelu.

Grafický model

Hlavní části (kroky) grafu

Znalostním modelem textových doporučení, vytvořeným ve fázi konstrukce, je orientovaný graf (viz obr. 2) skládající se z pěti hlavních částí (kroků):

- **Akce** – představuje specifickou činnost nebo událost. Akcí může být i podgraf, který dále zjemňuje danou činnost.
- **Rozhodování** – představuje větvení (výběr) na základě automatického splnění logického kritéria, kdy další postup grafem je dán výsledkem aritmetického nebo logického výrazu nad konkrétními daty. V případě, že nelze rozhodnout automaticky, má uživatel v tomto místě možnost výběru, kterou částí grafu bude dále pokračovat.
- **Větvení a synchronizace** – větvení se používá při modelování nezávislých kroků, které mohou probíhat paralelně a synchronizace slouží pro tyto kroky jako slučovací bod.
- **Stav** – značí stav, ve kterém se zkoumaný objekt nachází při vstupu do modelu nebo po provedení některého předchozího kroku.



Obr. 3. Rozhodovací kritéria v rozhodovacím kroku.

Rozhodovací kritéria

V každém rozhodovacím kroku jsou pro každou následnou volbu (větev) $\alpha^1 \dots \alpha^n$ definovány čtyři rozhodovací kritéria (viz obr. 3). Na základě jejich vyhodnocení je vybrán, automaticky nebo ručně, následný postup modelem.

- **Strict-in** – jestliže je splněna podmínka (např. logický výraz) nezávisle na uživateli, určitě se bude pokračovat příslušnou větví.
- **Strict-out** – jestliže je splněna podmínka, příslušnou větví se určitě nebude pokračovat – tato větev je zakázána.

- **Rule-in** – při splnění této podmínky je pouze doporučeno pokračovat následnou větví – je vyžadován zásah uživatele, uživatel by si měl vybrat pouze z větví s pravdivou podmínkou *rule-in*.
- **Rule-out** – při splnění této podmínky není doporučeno následnou větví – opět vyžadován zásah uživatele.

Při výběru následné větve se postupuje tak, že se nejprve vyhodnotí podmínka typu *strict-out* a při jejím splnění se další typy již nevyhodnocují, z pohledu možného průchodu grafem je příslušná větev zakázána. V opačném případě se vyhodnotí podmínka typu *strict-in*. Je-li splněna pokračuje se touto větví. Jestliže není splněna ani jedna podmínka typu *strict-in*, je vyžadován zásah uživatele s tím, že na základě vyhodnocení podmínek *rule-in* a *rule-out* je doporučen nebo nedoporučen výběr příslušné větve.

Vyhodnocování kritérií

Při vyhodnocování logických kritérií (*strict-in*, *strict-out*, *rule-in*, *rule-out*) se velice často setkáme se situací, kdy nejsou známy jednoznačné hodnoty všech proměnných. Neznámá hodnota proměnné nemusí být přítom chybou a je třeba takový stav brát v úvahu. Logické formule, složené z proměnných (parametrů modelu) a logických či relačních operandů, mohou nabývat následující stavy:

- **pravda** – formule je splněná – označíme hodnotou 1,
- **nepravda** – formule není splněna – označíme hodnotou 0,
- **neznámo** – stav formule není jednoznačně určený – označíme hodnotou $\frac{1}{2}$.

Z matematické logiky plyne, že jakoukoliv logickou formuli lze převést do disjunktivního (resp. konjunktivního) normálního tvaru, složeného pouze z logických operací negace, konjunkce (logický součin) a disjunkce (logický součet). Výsledný tvar formule je ekvivalentní s původní formulí.

Definici operací negace, konjunkce a disjunkce v tříhodnotové logice uvádějí následující vztahy:

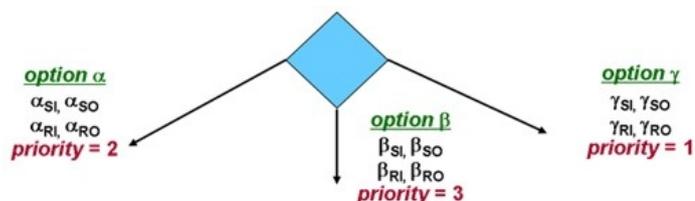
- $\neg p = 1 - p$
- $(p_1 \wedge p_2 \wedge \dots \wedge p_n) = \min(p_1, p_2, \dots, p_n)$
- $(p_1 \vee p_2 \vee \dots \vee p_n) = \max(p_1, p_2, \dots, p_n)$

Při vyhodnocování rozhodovacích kritérií a výběru následné větve (volby) ze všech možných větví (voleb) vedoucích z rozhodovacího kroku definujeme následující pravidla:

- Automatický postup větví vycházející z rozhodovacího kroku je možný pokud její *strict-in* podmínka je vyhodnocena jako pravda (1) a pokud na této větvi neleží současně *strict-out* podmínka vyhodnocená jako pravda.
- Pokud nenastanou první bod lze postupovat dále jenom tak, že uživatel některou větev ručně vybere. K výběru jsou uživateli nabídnuty pouze větve, na nichž nejsou *strict-out* podmínky vyhodnocené jako pravda (nejsou zakázané).
- Doporučení či nedoporučení větve je uživateli nabídnuto po vyhodnocení podmínek *rule-in* a *rule-out*.
- Jestliže některé z kritérií *strict-in* nebo *strict-out* je vyhodnoceno jako neznámo ($\frac{1}{2}$), je třeba doplnit hodnoty parametrů, které se v těchto podmínkách vyskytují. Po doplnění parametrů se provede opakované vyhodnocení. Kritéria *strict-in* a *strict-out* musí být vždy vyhodnoceny jednoznačně (pravda nebo nepravda).
- V každém rozhodovacím uzlu musí být zajištěna konzistence, tj. při všech možných kombinacích hodnot parametrů je splněna (hodnota pravda) maximálně jedno *strict-in* kritérium a nejsou na stejné větvi (volbě) splněny (pravdivé) obě kritéria *strict-in* a *strict-out*.
- Rozhodovací kritéria mohou být závislá na několika různých parametrech. Vyhodnocení rozhodovacího kroku musí být korektní (minimálně jedno z kritérií *strict-in* nebo *rule-in* některé volby musí být pravdivé a *strict-out* v této volbě nepravdivé) pro všechny možné kombinace hodnot každého parametru.

Priorita v rozhodování

Při důsledném uplatňování předchozích pravidel může nastat situace, ve které bude po uživateli vyžadováno upřesnění hodnot vstupních parametrů i v případě, kdy to pro další průchod modelem není



Obr. 4. Priorita větví v rozhodovacím kroku.

nezbytně nutné. V případě, že u některé z větví nabývá kritérium *strict-in* hodnoty pravda, bude se pokračovat určitě touto větvi a vyhodnocování kritérií ostatních větví je zbytečné.

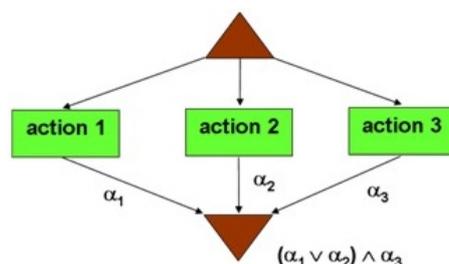
Počet potřebných upřesnění je závislé na pořadí vyhodnocování jednotlivých větví. Z uvedených důvodů je výhodné stanovit pořadí jejich vyhodnocování, tj. určit prioritu.

Priorita je přiřazena jednotlivým větvím rozhodování při návrhu modelu po expertní analýze modelovaného problému (viz obr. 4).

V uvedeném příkladu se nejprve budou vyhodnocovat rozhodovací kritéria větve γ , poté větve α a nakonec větve β . Pokud bude splněno kritérium *strict-in* větve γ , větve α a β se již vyhodnocovat nebudou.

Synchronizační podmínky

Při modelování paralelně probíhajících větví pomocí prvků větvení a synchronizace je nutné upřesnit, které z větví musí uživatel určitě projít a které jsou pouze volitelné.



Obr. 5. Synchronizační podmínka.

Uvedené případy lze realizovat pomocí *synchronizační podmínky* v synchronizačním kroku (viz obr. 5). Podmínku je možné opět vyjádřit v disjunktivním (konjunktivním) normálním tvaru, kde vstupními proměnnými jsou plovoucí (onfly) parametry reprezentující

průchod jednotlivými větvemi. Hodnota onfly parametru je nastavena pomocí skryté operace (viz reprezentace modelu v XML) v posledním kroku příslušné větve. Ohodnocení parametrů je pak následující:

- **hodnota 1** – touto větvi uživatel prošel,
- **hodnota 0** – touto větvi uživatel neprošel.

V ukázkové situaci se bude pokračovat v průchodu grafem, bude-li splněna synchronizační podmínka $(\alpha_1 \vee \alpha_2) \wedge \alpha_3$, tj. jestliže uživatel určitě projde větvi α_3 a zároveň alespoň jednou z větví α_1 nebo α_2 .

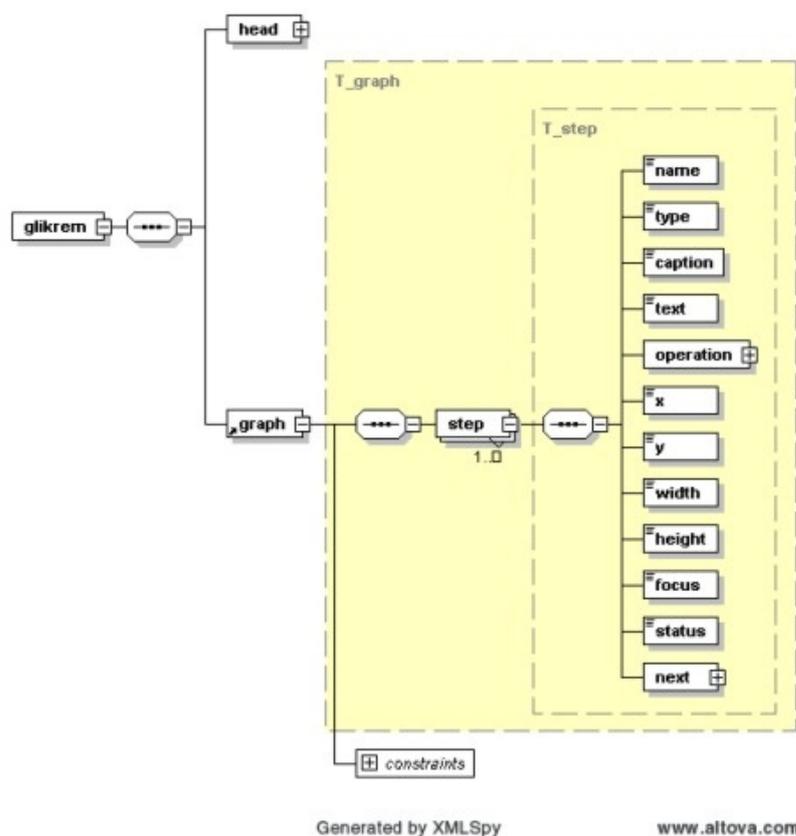
Reprezentace grafického modelu v XML

Pro reprezentaci grafického modelu jsme navrhli vlastní XML schéma a vyvinuli grafický editor pro konstrukci grafického modelu a jeho překlad do XML. Celý model se skládá z jednotlivých kroků reprezentujících příslušné prvky grafického modelu. Kromě atributů vyplývajících přímo z modelu obsahuje navržená syntaxe XML i atributy vhodné pro další zpracování, například pro grafické zobrazení nebo pro podporu rozhodování (viz obr. 6).

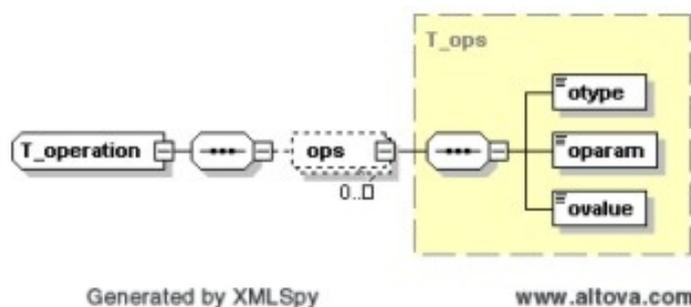
Význam jednotlivých elementů je následující:

- **glikrem** – kořenový element celého modelu (GLIKREM),
- **head** – definice hlavičky modelu - viz kapitola rozšíření modelu pro MEKRES,
- **graph** – kořenový element vlastního grafického modelu, skládá se z kroků (step),
- **step** – parametry jednoho kroku, obsahuje elementy:
 - name** – jednoznačná identifikace (název) kroku,
 - type** – typ kroku, je možný pouze výběr z typů: action, case, branch, synchronization, state, subgraph,
 - caption** – text v záhlaví grafického symbolu kroku,

- text** – vlastní text popisující krok,
 - x** – x-ová (horizontální) souřadnice grafického symbolu,
 - y** – y-ová (vertikální) souřadnice grafického symbolu,
 - width** – šířka (v pixelech) grafického symbolu,
 - height** – výška (px) grafického symbolu,
 - focus** – zvýraznění kroku pro zobrazení cesty v grafu, možné hodnoty jsou:
 - not* = nezvýrazněn, *auto* = zvýrazněn (automaticky),
 - user* = zvýrazněn (uživatelská simulace),
 - status** – postavení kroku v rámci modelu, možné hodnoty jsou:
 - start* = počáteční krok,
 - end* = koncový krok,
 - in* = vnitřní krok (všechny ostatní),
- operation** – seznam skrytých operací probíhajících „na pozadí“ kroku (viz obr. 7), každá operace je ohraničena elementy **ops** a obsahuje elementy:
 - otype** – typ operace, je možný výběr z:
 - insert* = vložení (zadání) hodnot parametrů,
 - get* = načtení hodnot parametrů (zjištění hodnoty),
 - put* = nastavení (uložení) hodnoty parametru,
 - open* = otevření souboru (např. podgrafu),
 - oparam** – název parametru nebo název souboru,
 - ovalue** – pouze pro operaci *put* obsahuje hodnotu, na kterou má být parametr nastaven,
- next** – seznam následných větví (voleb) kroku (viz obr. 8), každá volba je ohraničena elementy **option** a obsahuje elementy:
 - nname** – název, identifikace následné větve,
 - nstep** – identifikátor (název) kroku, kterým se má pokračovat,
 - ncaption** – text zobrazený u příslušné volby,
 - npriority** – priorita vyhodnocení větve udaná číslem, defaultně hodnota 1,
 - nstrictin** – *strict-in* rozhodovací kritérium,
 - nstrictout** – *strict-out* rozhodovací kritérium,
 - nrulein** – *rule-in* rozhodovací kritérium,



Obr. 6. XML schéma reprezentace grafického modelu.



Obr. 7. XML schéma operací na pozadí kroku.

- nruleout** – *rule-out* rozhodovací kritérium,
 - nnote** – poznámka popisující příslušnou větev,
 - nline** – popis čáry vedoucí z tohoto kroku do následného, ve tvaru sekvence bodů (elementů **point**):
 - px** – x-ová souřadnice bodu,
 - py** – y-ová souřadnice bodu.

Všechny typy rozhodovacích podmínek mají obvykle tvar logické formule, kde proměnnými jsou základní a odvozené parametry paramodelu (modelu parametrů). Zápis rozhodovacích

podmínek (*strict-in*, *strict-out*, *rule-in*, *rule-out*) dodržuje syntaxi jazyka **Xpath**.

Příklad definice podmínky testující pravdivost obou parametrů PRIZNAKY a NALEZY:

```
/params/param[pid="PRIZNAKY"]/pvalue
and params/param[pid="NALEZY"]/pvalue
```

Zápis synchronizačních podmínek je obdobný (opět v jazyku **XPath**), předpokládá ale definici plovoucích (*onfly*) parametrů definujících průchod příslušnou větví.

Příklad definice synchronizační podmínky, kdy je vyžadován průchod větví S3 nebo větví S4:

```
/params/param[pid="DG_S3"]/pvalue or
/params/param[pid="DG_S4"]/pvalue
```

Reprezentace paramodelu v XML

Pro reprezentaci paramodelu je, stejně jako u grafického modelu, využito jazyku XML. Celý model se skládá z elementů reprezentující příslušný parametr (viz obrázek 9). Parametry GLIKREM mohou být trojího typu a to:

- **basic** – základní parametr, tj. přímo měřitelná nebo získatelná hodnota,
- **derived** – parametr odvozený ze základních provedením logické, aritmetické či logicko-aritmetické operace,
- **onfly** – pomocný parametr (obsahující logickou hodnotu) pro definici synchronizačních podmínek.

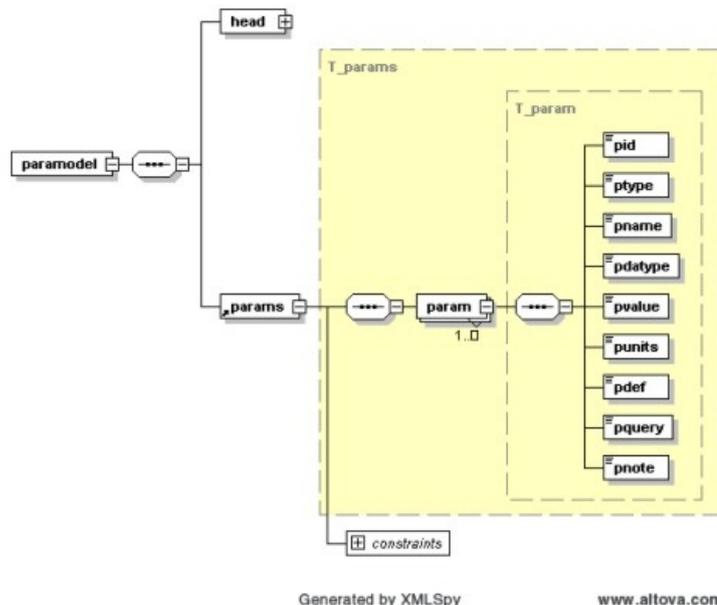
Význam jednotlivých XML elementů paramodelu je následující:

- **paramodel** – kořenový element celého paramodelu,
- **head** – hlavička paramodelu,
- **params** – kořenový element vlastních parametrů, obsahuje elementy **param**,
param – popis jednoho parametru,
pid – jedinečná identifikace (název) parametru
pptype – typ parametru, možné hodnoty jsou basic, derived a onfly,
pname – text úplného názvu parametru,
pdatatype – datový typ parametru odpovídající některému z XML datových typů,
pvalue – vlastní hodnota parametru,
punits – jednotky parametru,
pdef – definice operace nad základními parametry pro odvozené parametry,
pquery – definice dotazu do databáze na hodnotu parametru (např. SQL),
pnote – text poznámky k parametru.

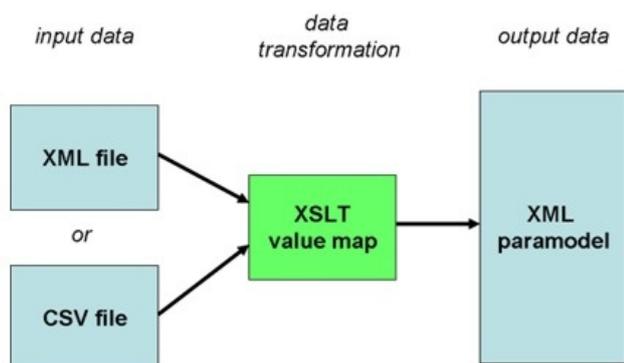
Definice odvozených parametrů (element pdef) využívá opět zápis v jazyce XPath.

Transformace vstupních dat

Reálná data o pacientech mohou být uložena v klinických informačních systémech a databázích různých typů. Není však velký problém data z těchto systémů exportovat do formátů XML nebo CSV (*Comma-Separated Values*). Před



Obr. 9. XML schéma paramodelu.



Obr. 10. Proces transformace dat do paramodelu.

použitím těchto dat v GLIKREM je třeba je vhodným způsobem transformovat do paramodelu (v XML). Zjednodušený proces transformace je znázorněn na obrázku 10.

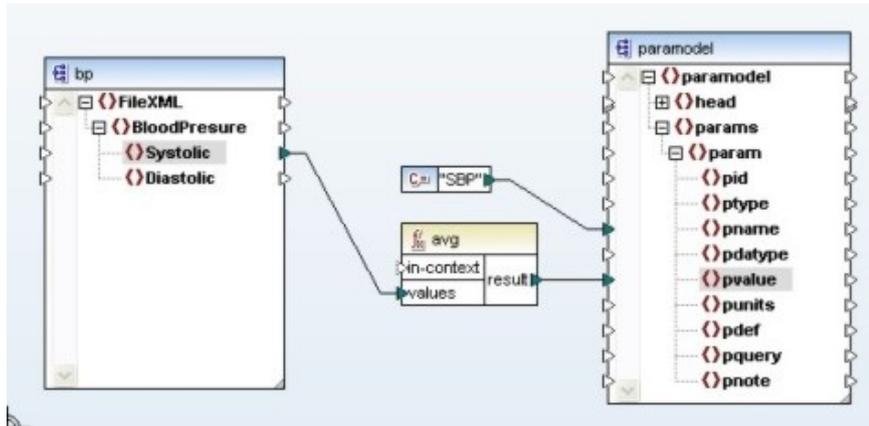
Pro každý parametr v paramodelu je třeba, s ohledem na uložení reálných vstupních hodnot v datovém souboru (XML nebo CSV), definovat specifickou hodnotovou mapu (value-map). Definice hodnotové mapy je ve formě XSLT (*eXtensible Stylesheet Language Transformations*) souboru. Například stanovení hodnoty systolického tlaku jako průměr z naměřených hodnot (viz. obr. 11) uložených v XML souboru.

Rozšíření modelu pro použití v systému MEKRES

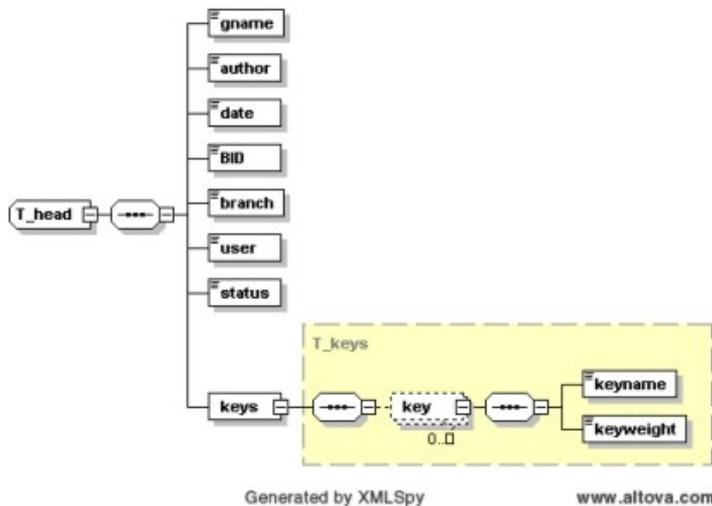
Pro použití GLIKREM v systému MEKRES (*Medical Knowledge REpresentation System*) je model rozšířen o klíčové atributy. Klíčové atributy se uplatní v algoritmu výběru relevantních formalizovaných znalostí účastníkům systému. Klíčové atributy jsou zakódovány opět v jazyce XML a uloženy v hlavičce (element head) společně s grafickým modelem (viz obr. 12).

Význam jednotlivých XML elementů hlavičky (elementu head) je následující:

- **gname** – název znalostního modelu,
- **author** – autor(ři) modelu,
- **date** – datum poslední aktualizace,



Obr. 11. Transformace systolického tlaku do paramodelu.



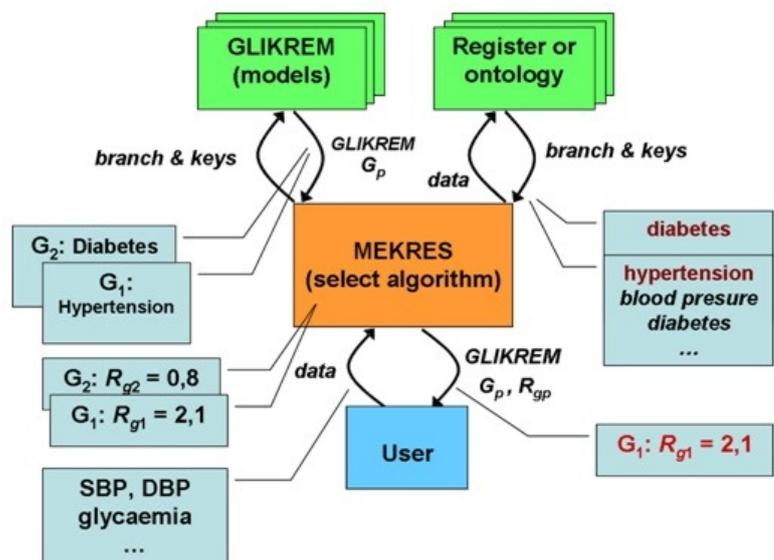
Obr. 12. XML schéma klíčových atributů.

- **BID** – identifikátor oblasti – například pro medicínu podle mezinárodního číselníku nemocí (MSN),
- **branch** – oblast (obor), které se model týká,
- **user** – označení účastníka systému, kterému je model primárně určen (možné hodnoty jsou *patient*, *GP*, *specialist*, *operator*, *everybody*),
- **status** – označení platnosti modelu (možné hodnoty jsou *valid*, *draft*, *expired*),
- **keys** – seznam klíčů (klíčových atributů) vyskytujících se v modelu obsahuje elementy **key**,
keyname – označení klíče,
keyweight – váha klíče udává, do jaké míry je klíč popsán modelem.

Algoritmus výběru relevantního modelu

Celý systém reprezentace znalostí MEKRES a algoritmus výběru relevantního znalostního modelu je znázorněn na obrázku 13. Algoritmus výběru relevantního modelu lze popsat následovně:

- Pro uživatele systému a jeho atributy (data o uživateli) je vybrána množina oborů (branch) a klíčů, které odpovídají stavu uživatele (datům).
 - např. atributům pacienta „STK“, „DTK“ a „glycaemia“ odpovídají obory „diabetes“ a „hypertenze“.
- Pro každý obor (branch) a klíč je vybrána množina znalostních modelů (GLIKREM), které obsahují tento obor v hlavičce (element head) a váha klíče (element keyweight) uloženého v modelu je nenulová.
 - např. modely G_1 (hypertenze) a G_2 (diabetes).
- Pro každý vybraný model je stanovena hodnota obecného agregovaného operátoru (např. $R_g = \sum keyweight(k)$) udávající relevantnost daného modelu.
 - např. model G_1 má operátor $R_{g_1} = 2,1$ a model G_2 má $R_{g_2} = 0,8$.
- Uživateli je zpětně zobrazen model s největší hodnotou relevance nebo seznam modelů uspořádaných podle relevance.
 - např. uživateli je nabídnut model G_1 s největší hodnotou relevance.



Obrázek 13. Systém MEKRES.

Závěr

Navržený model reprezentace znalostí GLIKREM, vycházející ze specifikace GLIF3.5, je univerzálním nástrojem pro formalizaci znalostí uložených ve formě volného textu (např. lékařských doporučení). Reprezentace grafického modelu v XML a vytvoření datového rozhraní ve formě paramodelu (v XML) umožňuje použití modelu v rozličných typech aplikací a napojení na reálná data z různých zdrojů. Pro transformaci vstupních dat i vlastního modelu je použito definice ve formě XSLT souboru.

Hlavní rozšíření a přínosy navrženého modelu jsou následující:

- Formální reprezentace grafického modelu v XML. Výsledkem reprezentace je návrh XSD (*XML Schema Definition*).
- Přesnější definice rozhodovacích kritérií a jejich vyhodnocování v tří hodnotové logice.
- Definice atributu priorita pro snížení nutného počtu vyhodnocení jednotlivých voleb v rozhodovacích krocích.
- Návrh paramodelu (v XML), který je použit jako rozhraní mezi GLIKREM a reálnými daty pacientů v klinických informačních systémech.
- Rozšíření GLIKREM o klíčové atributy. Toto rozšíření je dále využito v systému reprezentace lékařských znalostí (MEKRES).

Poděkování

Vytvořeno s podporou grantu č. AV0Z10300504 a č. 1M06014 MŠMT ČR.

Literatura

- [1] Peleg M., Ogunyemi O., Tu S., et al.: Using features of Arden Syntax with object-oriented medical data models for guideline modeling. *Proc AMIA Symp.* 2001:523-7.
- [2] Peleg M., Tu S.W., et al.: Comparing Computer-interpretable Guideline Models: A Case-study Approach. *The Journal of the American Medical Informatics Association*, 2003 Jan-Feb; 10(1): 52-68.
- [3] Shahar Y., Miksch S., Johnson P.: The Asgaard Project: a task-specific Framework for the application and critiquing of time-oriented clinical guidelines. *Artif Intell Med* 1998;14:29-51.
- [4] Tu S.W., Musen M.A.: A flexible approach to guideline modeling. *Proc AMIA Symp* 1999:420-424.
- [5] Tu S.W., Musen M.A.: From guideline Modeling to guideline execution: Defining guideline-based decision-support Services. *Proc AMIA Annu Symp* 2000:863-867.
- [6] Quaglini S., Stefanelli M., Lanzola G., Caporusso V., Panzarasa S.: Flexible guideline-based patient careflow systems. *Artif Intell Med* 2001;22:65-80.
- [7] Johnson P.D., Tu S.W., Booth N., Sugden B., Purves I.N.: Using scenarios in chronic disease management guidelines for primary care. *Proc AMIA Annu Fall Symp* 2000:389-393.
- [8] Gennari J.H., Musen M.A., Fergerson R.W., Grosso W.E., Crubezy M., Eriksson H., et al.: The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. *Int J Hum Comput Stud.* 2003;58(1):89-123.
- [9] Bury J., Fox J., Sutton D.: The PROforma guideline specification language: progress and prospects. *Proceedings of the First European Workshop, Computer-based Support for Clinical Guidelines and Protocols (EWGLP 2000)*, 2000.
- [10] Ohno-Machado L., Gennari J.H., Murphy S.N., Jain N.L., Tu S.W., Oliver D., et al.: The Guideline Interchange Format: A model for representing guidelines, *Journal of the American Medical Informatics Association.* 1998. ISSN 1067-5027.
- [11] Zeng Q.: Guideline Interchange Format 3. technical specification. Available at: <http://www.glif.org> - Peleg, Boxwala, et al.
- [12] Boxwala A.A., Peleg M., Tu S.W., Ogunyemi O., Zeng Q., Wang D., et al.: GLIF3: A Representation Format for Sharable Computer-Interpretable Clinical Practice Guidelines. *J Biomed Inform.* 2004;37(3):147-61.
- [13] Sordo M., Boxwala A., Ogunyemi O., Greenes R.: Description and Status Update on GELLO: a Proposed Standardized Object-oriented Expression Language for Clinical Decision Support. *Medinfo*; 2004; 2004. p. 164-8.
- [14] Laleci G.B., Dogac A., et al.: SAPHIRE: A Multi-Agent System for Remote Healthcare Monitoring through Computerized Clinical Guidelines. online at <http://www.srdc.metu.edu.tr/webpage/projects/saphire/>.
- [15] Buchtela D., Peleska J., Vesely A., Zvarova J.: Medical Knowledge Representation System, 21st International Congress MIE 2008 in Göteborg, 2008, Sweden.
- [16] Buchtela D., Peleska J., Vesely A., Zvarova J.: Method of GLIF model Construction and Implementation, The XIX International Congress of the European Federation for Medical Informatics. 2005.
- [17] Peleška J., Anger Z., Buchtela D., Šebesta K., Tomečková M., Veselý A., Zvára K., Zvárová J.: Formalization of Medical Guidelines. *European Journal for Biomedical Informatics.* 2005. Prague.
- [18] Veselý A., Zvarova J., Peleška J., Buchtela D., Anger Z.: Medical Guidelines Presentation and Comparing with Electronic Health Record. *International Journal of Medical Informatics* 75 (3-4). 2006. 240-245

Kontakt

Ing. David Buchtela

Centrum biomedicínské informatiky
Ústav informatiky AV ČR, v.v.i.
Pod Vodárenskou věží 2
Praha 8 182 07
Česká republika
e-mail: buchtela@euromise.cz



EJBI 2008

ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 4 (2008), Issue 1

Lithuanian version

www.ejbi.eu



Content

Lithuanian version

- It 2 - 7 Objektinių Technologijų Naudojimas Medicininių Duomenų Saugojimui Ir Paieškai
Vita Speckauskiene, Arunas Lukosevicius

Objektinių Technologijų Naudojimas Medicininių Duomenų Saugojimui Ir Paieškai

Vita Speckauskiene, Arunas Lukosevicius

Biomedical Engineering Institute at Kaunas University of Technology, Lithuania

Santrauka. Straipsnyje analizuojamos objektinėmis technologijomis ir šių technologijų panaudojimas kuriant medicinines paskirties informacinę sistemą (IS). Medicininių duomenų standartų, konceptų ir technologijų spartus vystymasis iškėlė problemą – kaip sukurti sistemą, kuri pakankamai lanksčiai ir gerai veiktų su nuolat besikeičiančiais duomenimis. Po objektinių technologijų analizės: J2EE (Java 2 Platform, Enterprise Edition) ir Microsoft .NET, bei atvirojo kodo objektinių duomenų bazių valdymo sistemų (ODBVS), mes pasiūlėme sistemos architektūrą ir panaudojant .NET technologiją bei ODBVS db4o sukūrėme sistemą. Rezultatai apima ne tik sukurtą sistemą, bet taip pat objektinių duomenų bazių privalumus, bei sistemos integraciją į Nacionalinę elektroninės sveikatos sistemą. Sukurta sistema turėtų padėti medicinos darbuotojams pasiekti svarbią informaciją iš kiekvieno kompiuterio, kiekvienoje ligoninėje.

Reikšminiai žodžiai: objektinės technologijos, objektinė duomenų bazė, informacinė sistema, duomenų saugojimo standartai

Ivadas

Medicininės paskirties informacinės sistemos (IS) yra viena iš pastaruoju metu intensyviai besiplėtojančios e-sveikatos sistemos komponentų. Pagrindiniai šiuolaikinių medicininių IS skiriamieji bruožai yra: 1) saugomi ir kaupiami duomenys yra įvairialypiai; 2) informacija yra surišta sudėtingais semantiniiais ryšiais, kurie, kaupiantis žinioms, neišvengiamai evoliucionuoja; 3) informacijos kiekiai yra dideli, nes apima visą gyvenimą trunkančius pacientų sveikatos įrašus; 4) vartotojų ratas (pav. gydytojai) yra plačiai pasiskirstęs per visus tris sveikatos paslaugų lygius, todėl reikalingos saugios ir patogios kompiuterinio tinklo prieigos dvipusiam

ryšiu su IS; 5) IS turi būti suderintos su Europos Sąjungos (ES) [1] ir tarptautiniais standartais [2, 6]; 6) duomenys turi tarnauti klinikinių sprendimų rengimui (KSR).

Tai kelia gana specifinius reikalavimus IS, jų architektūrai. Taip pat kuriant tokias sistemas reikia ieškoti naujų technologijų [3, 4, 5, 7, 16, 17]. Šiuo metu pasaulyje yra sukurta ir kuriami vis nauji standartai optimizuojantys ir palengvinantys pacientų sveikatos įrašų, taip pat ir vaizdų saugojimą. Tačiau jų apdorojimas ir centralizuotas saugojimas išlieka problema, kadangi ilgą laiką naudotos technologijos nėra pakankamai lanksčios tokiems duomenims saugoti ir apdoroti. Be to, iškilus suderinamumo tarp nacionalinių sveikatos paslaugų užtikrinimo poreikiams bei vystantis kompiuterinėms technologijoms medicinoje didėja poreikis vieningo tarptautinio standarto skirto medicinos duomenų mainams. Kiekvienoje šalyje šie klausimai iki šiol buvo sprendžiami gana skirtingai, todėl šiuo metu egzistuoja atskiri standartai, aprašantys įvairius IS ir duomenų saugojimo aspektus: ASTM, ASC X12, IEEE/MEDIX, NCPDP, HL7 (Health Level 7) [6], DICOM [2] ir kt. Kiekvieną standartą kurianti organizacija, grupė ar komitetas paprastai specializuojasi. Deja, nors standartai sumažina ir gelbsti dirbant su mediciniais duomenimis, tačiau duomenų perdavimas ir kaupimas vis dar išlieka technologinė problema. Ilgą laiką naudojamos technologijos nėra pakankamai lanksčios dirbti su minėtais duomenimis. Duomenys saugomi reliacinėse duomenų bazėse (RDB), nors teigiama [18, 19], kad objektinės (ODB) turi vis daugiau privalumų. Saugios ir lanksčios internetinės prieigos reikalavimas taip pat verčia ieškoti efektyvių programavimo kompiuteriniuose tinkluose technologijų.

Šio straipsnio tikslas yra išanalizavus galimybes pasiūlyti ir sukurti medicininės IS

architektūrą, atitinkančią pagrindinius aukščiau išvardintus bruožus, palyginti ir pasiūlyti atitinkamas IS įgyvendinimo priemones. Šiam tikslui pasiekti analizuojamos objektinės duomenų bazės (ODB), jų veikimui tinkluose palankios programavimo technologijos, duomenų saugojimo metodai, bei analizuojamos ir įgyvendinamos pagrindinės siūlomos sistemos savybės. Taip pat ši sistema ruošiama integracijai į Nacionalinę elektroninės sveikatos sistemą (NESS), kuri šiuo metu yra kuriama nacionaliniu mastu. Deja, kol kas negalime pavaizduoti sistemų integralumo, kadangi NESS dar nėra įdiegta, tačiau pateikėme keletą pastebėjimų.

Metodai

Objektinių duomenų bazių analizė Objektinė duomenų bazė yra nauja sparčiai besivystanti technologija, kurią diegia vis daugiau kompanijų. Tačiau jau atsiranda ir ODB bazių vartotojai, o didėjant jų skaičiui - daugėja ir poreikių. Šios duomenų bazės (DB) ypatingos tuo, kad duomenys jose saugomi lentelėmis, kurios savyje gali turėti kitas lenteles, sukurdamos hierarchiją. Lentelės užpildomos ne tik konkrečiais duomenimis, bet ir objektais. Tokią technologiją labai patogiu pritaikyti medicininių diagnostinių vaizdų saugojimui. Kadangi medicininės informacinės sistemos yra nuolatos evoliucionuojančios, objektinis principas duomenų bazėms suteikia lankstumo ir patikimumo [18, 19]. Taip pat įgalinamas objektų bendravimas duomenų bazėje.

Reliacinės duomenų bazių valdymo sistemos (RDBVS), pritaikytos tik tekstinių ir skaitmeninių duomenų saugojimui. Jos nepajėgios susitvarkyti su neapibrėžtais iš anksto duomenų tipais, kurie susiformuoja šiuolaikinėse objektinėse kalbose. Be to, RDBVS visiškai nepritaikytos objektų sudėtinėms dalims - metodams saugoti.

Informacinė sistema turi apdoroti ne tik tekstus, bet ir skaičiavimo lenteles, vaizdus, diagramas, žemėlapius, garso įrašus, kitokią daugialypės terpės informaciją. Be to, visa ši informacija turi būti organizuota kaip vieninga struktūra, patogi vartotojui. Objektinė technologija galėtų užtikrinti naują informacijos saugojimo ir paieškos technologiją, suteikti informacijos valdymui naujų galimybių. Objektinės duomenų bazių valdymo sistemos (ODBVS) saugo ir sudėtingas struktūras - sudėtinius objektus (composite objects), kurie susideda iš kitų objektų.

1 lentelė. Komercinės ODBVS.

OODBMS	Vendor
Versant Developer Suite and FastObjects	Versant Corporation
TITANIUM	Micro Data Base Systems, Inc.
GemStone/S and GemStone Facets	GemStone Systems, Inc.
Cache	InterSystems Corporation
JADE	Jade Software Corporation
JYD Object Database	JYD Software Engineering Pty Ltd.
VOSS	Logic Arts, Ltd.
Objectivity/DB	Objectivity, Inc.
ObjectStore Enterprise and PSE Pro	Progress Software Corporation

Sudėtiniai objektai saugomi sudarant adresines nuorodas (references) į žemesnio lygio objektus, lygių skaičius gali būti labai didelis. Kiekvienas objektas gali būti įtrauktas į daugelį kitų objektų kaip komponentė. Taip susiformuoja struktūra. Objektinėse DB yra vidinių duomenų tipų, tačiau galima sukurti naujus, reikalingus konkrečiu atveju, duomenų tipus, ši savybė įgalina saugoti daugialypės terpės informaciją: paveikslus, nuotraukas, vaizdo įrašus, kalbos įrašus, muziką ir kitokių tipų informaciją.

Objektinės DB yra dviejų tipų: pasyvios ir aktyvios. Pasyvios ODB saugo objektų struktūrą, bet nerealizuoja objektų elgsenos. Aktyvios ODB leidžia objektams sąveikauti tiesiog duomenų bazėje.

Dauguma siūlomų ODBVS yra mokamos. Lentelėje pateiktos geriausiai žinomos [8].

Vis tik didesnė pasiūla yra nemokamų ODBVS. Sekančioje lentelėje apibūdintos septynios atviro kodo ODBVS [10, 11, 12, 13, 14, 15].

Iš lentelės matome, kad dauguma atviro kodo DB sukurtos J2EE programavimo aplinkai. Tačiau ne visos jos atitinka standartus ir užtikrina pilną ODB būdingą lankstumą. Objektinių duomenų bazių

standartus tvarko ODMG (Object Database Management Group) [9]. Vis tik, analizės metu, nepavyko išvelgti nė vieno trūkumo ODBVS db4o, kurią pasirinkome galutiniam rezultatui įgyvendinti.

J2EE ir .NET technologijų palyginamoji analizė

Tinklo paslaugos (Web service) technologijos šiuo metu yra labai plačiai naudojamos, dėl to, kad jos suteikia daug galimybių programuotojui, jų realizavimas nėra sudėtingas. Pagrindinės vyraujančios platformos yra J2EE ir .NET. Kompanijos siūlančios šias paslaugas yra ir vienos didžiausių konkurenčių. Taigi programuotojui, ketinančiam naudoti vieną iš šių produktų iškyla sunkus uždavinys - kurią pasirinkti. Sun Microsystems siūlo naudoti Java 2 Platform, Enterprise Edition (J2EE), o Microsoft - .NET technologiją.

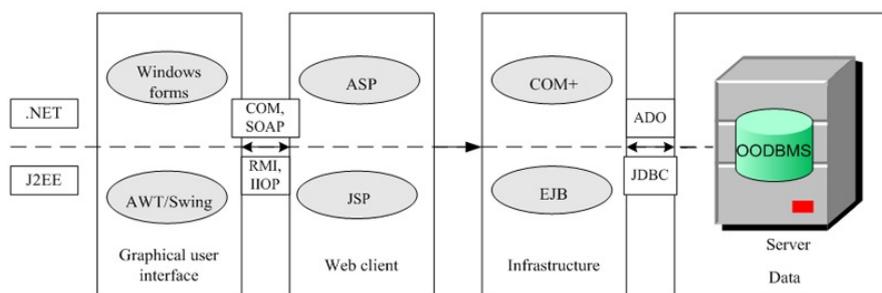
Abi šios technologijos siūlo įrankius tinklo paslaugų kūrimui. Pagrindinis J2EE privalumas yra tas, kad jos įrankius galima naudoti bet kurioje operacinėje sistemoje (OS). Microsoft .NET sukurta Windows operacinėms sistemoms, tačiau šiuo metu kompanija siūlo atitinkamus įrankius ir Linux OS. Abi technologijos naudoja objektines programavimo kalbas (.NET - Visual Basic ir kt., J2EE - JAVA). .NET technologijos privalumas yra tas, kad ji geba atvaizduoti puslapius įvairiais HTML formatais. Nors J2EE technologijoje servletais galima gauti tokį patį rezultatą, tačiau jie reikalauja papildomo programavimo [7]. Dar vienas .NET privalumas yra daugiakalbiškumas, ji palaiko Visual Basic, C#, Fortran ir COBOL programavimo kalbas. J2EE veikia tik JAVA programavimo kalba. Pagrindinės šių technologijų funkcijos pateiktos lentelėje [3, 4, 16].

2 lentelė. Atviro kodo ODBVS

OODBMS	Programming environment	Limitations
db4o	J2EE .NET	Handling of default and null values (noticed during implementation).
SOD2	J2EE C++	Does not have a query mechanism.
ozone	J2EE	Implemented in JAVA and works only in the J2EE environment. Does not have a query mechanism.
ObjectDB	J2EE	Implemented in JAVA and works only in the J2EE environment. Free DB edition does not support a client-server architecture (unless the DBMS is in the same server).
GigaBASE	J2EE C++	Does not support null values.
XL2	J2EE	Does not have data recovery functions. Does not have a query mechanism.
Orient	J2EE, C++	Available for C++ (ODMG 3.0 interface) and Java with the support of Sun JDO 1.0 technology.

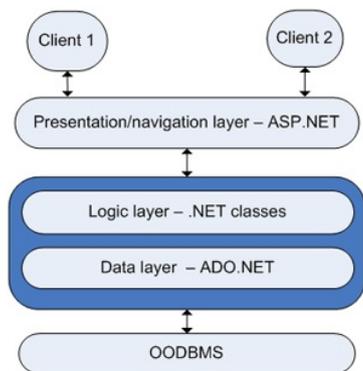
3 lentelė. J2EE ir .NET technologijų suvestinė.

Feature	J2EE	.NET
Relational DB Access	Java Database Connectivity (JDBC) or L/J	Active Data Objects (ADO.NET)
Web Client	JAVA server pages (JSP) and servlets	ASP.NET
Graphical user interface (GUI)	AWT/Swing	Windows forms
Messaging	Java Messaging Service (JMS 1.0)	Microsoft Messaging Queuing (MSMQ)
Web Services Support	Java Web Services Developer Pack (JWSDP)	Built directly into .NET and Visual Studio
Infrastructure	Enterprise Java Beans (EJB)	COM+
Naming and Directory Service	Java Naming and Directory Interface (JNDI)	Active Directory Services Interface (ADSI)
Database storage	-	SQLServer (the official .NET platform database technology)



1 paveikslas. IS architektūra .NET ir J2EE technologijoms.

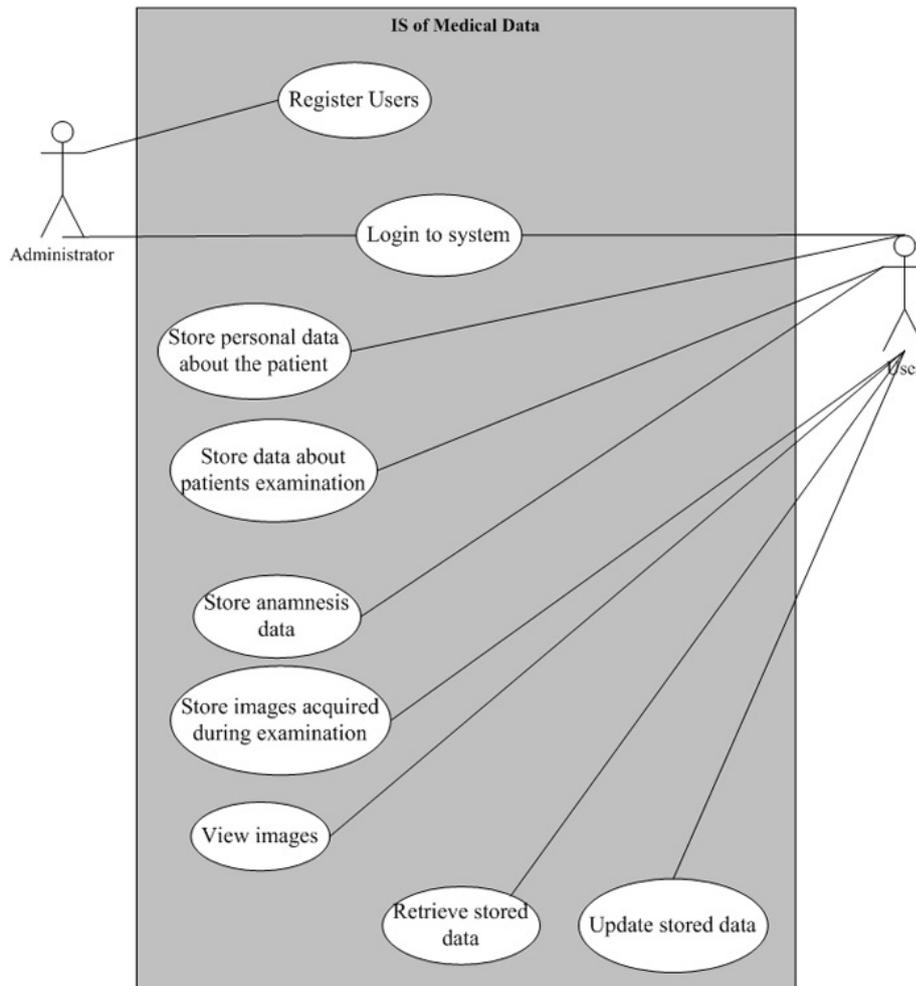
Kaip matome abi technologijos yra ypatingai panašios. Svarbiausia, kad jos įgalina kurti tinklo paslaugas, atvaizduoti duomenis internete. .NET tinklo paslauga gali būti vartotoju J2EE tinklo paslaugoms ir atvirkščiai, kadangi standartai, kuriuos pastarosios naudoja, yra vienodi [7]. Be to tinklo paslaugų programas, galima pakartotinai panaudoti, nesvarbu kokia platforma naudojantis jos sukurtos. Dauguma autorių, dėl technologijų panašumo, nenurodo ypatingų priežasčių, kurios galėtų lemti pasirinkimą [5, 16, 17]. Vis tik galima įžvelgti kai kuriuos .NET privalumus, jos plėtojimo, bei naudojimo galimybes. Tai lėmė technologijos parinkimą galutinio produkto kūrime.



2 paveikslas. Sukurtos sistemos architektūra.

Sistemos architektūra ir funkcionalumas Medicininių duomenų saugojimui ir apdorojimui skirtą IS būtų galima realizuoti naudojantis minėtomis technologijomis. Nepriklausomai nuo šiuo pasirinkimo sistemos architektūra yra analogiška (1 paveikslas). Šių technologijų panašumo privalumas yra tas, kad jei kiltų būtinybė, abi sistemos galėtų bendrauti su DBVS. Kaip minėta ankstesniuose skyriuose, medicininių duomenų saugojimo, gavimo ir paieškos sistemos kūrimui pasirinkome ODVS db4o [10] ir .NET technologiją. Paveiksle pavaizduota sukurtos sistemos architektūra.

Kaip matome sistema susideda iš trijų lygmenų: atvaizdavimo, programos logikos ir duomenų. Bendravimas tarp atvaizdavimo ir duomenų lygmenų yra vykdomas per programos logikos lygmenį. 3 paveiksle pavaizduota kuriamos informacinės sistemos panaudojimo atvejų diagrama.



3 paveikslas. IS panaudojimo atvejų diagrama.

4 lentelė. Programinio kodo pavyzdžiai, vaizduojantys ODB užklausas ir programavimo galimybes.

using com.db4o; using com.db4o.query;	Standard libraries (mainly used)
Connecting to the data base and making a query	
ObjectServer db4oServer = Db4o.openServer("db.yap", 4488); db4oServer.grantAccess("..."); ObjectContainer client = Db4o.openClient("..."); Query query1 = client.query(); query1.constrain(typeof(Vartotojas)); query1.descend("_vartotojas").descend("_pavarde").constrain(p.Text); ObjectSet result1 = query1.execute(); client.close(); db4oServer.close();	Connection to the server Create server client Make a query Set the class to look into Find a particular data Execute query Closing the client Closing the server
Setting and updating new data	
Vartotojas vartotojasn = new Vartotojas(P1.Value,p.Text,'N'); client.set(vartotojasn); Vartotinfo vartotinfo = new Vartotinfo(v.Text,el.Text,g.Text,k.Text,s.Text); vartotinfo.Vartotojas = vartotojasn; client.set(vartotinfo);	Generating new data for class "Vartotojas" Storing the data Generating new data for class "Vartotinfo" Creating a reference Storing the data
Working with "blob" type files	
Db4o.configure().setBlobPath("blobs"); Byla snap = new Byla("..."); client.set(snap); snap.Bylapav.readFrom(new File(File1.Value)); double status = snap.Bylapav.getStatus(); while(status > Status.COMPLETED) { status = snap.Bylapav.getStatus(); }	Setting the blob path Generating new data for class "Byla" Storing the data (also the text data about the file: the file name, extension and size) Preparing to store the blob type file itself Storing the file itself
Byla found = (Byla)result.next(); string pav = found.Bylapav.getFileName();	Querying for the file Retrieving the file

Panaudojimo atvejų diagrama vaizduoja sistemos funkcionalumą (funkcijas, kurias gali atlikti vartotojas ir administratorius).

Rezultatai

Sistema sukurta naudojant aprašytas technologijas. Integracija tarp .NET ir db4o nesukėlė didesnių problemų.

Integracija tarp .NET ir db4o nesukėlė didesnių problemų. Bendravimas tarp sistemos ir duomenų bazės pasiekiamas per standartinį duomenų bazės užklausų mechanizmą, kuris yra standartinėse duomenų bazės bibliotekose. Programinio kodo pavyzdžiai, vaizduojantys ODB užklausas ir programavimo galimybes, panaudotas sukurtose sistemoje, pateikti lentelėje.

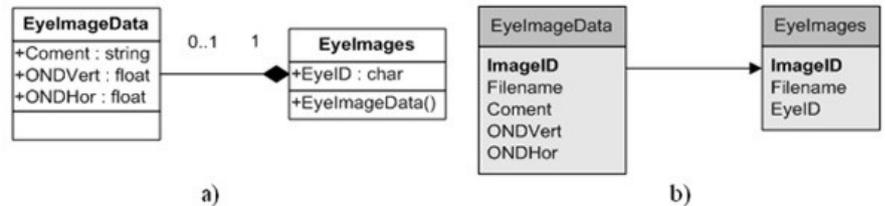
Duomenų bazė yra viena .yap tipo byla, blob tipo byla gali būti bet kokia byla be jokių apribojimų dydžiui ar bylos plėtimui (jei pageidaujama). Taigi duomenų ir bylų įkėlimo mechanizmai labai lankstūs ir patogūs programuotojui.

Metodas sistemos testavimui.

Sukūrus sistemą buvo vykdomas jos testavimas ir atliktas eksperimentas įvedant keletą sistemos pakeitimų. Eksperimento metu siekta įvertinti objektinių technologijų, naudotų IS realizavimui, funkcionalumą, atlikti pakeitimai visuose trijuose sistemos lygmenyse - duomenų bazėje, vartotojo sąsajoje ir programos logikos lygmenyse. Didžiausias dėmesys skirtas parodyti ODBVS privalumus medicininiams duomenims saugoti ir apdoroti. Šiai užduočiai įgyvendinti sukurtas scenarijus realiai egzistuojantis ir praktikoje. Atsiranda pakeitimai funkcinuose reikalavimuose - papildomai reikia įvesti medicininių vaizdų parametrus. Po pirminio testavimo atlikome pakeitimus minėtus scenarijuje ir po to vėl atlikome sistemos testavimą. Sekančios keturios dalys apibūdina testavimo procedūrą, pasikeitimus duomenų bazėje, atvaizdavimo, bei programos logikos lygmenyse.

Testavimas.

Testavimui naudoti duomenys sukaupti Kauno medicinos universiteto klinikose pacientų apžiūros metu. Testavimui naudojome nedidelį kiekį duomenų,



4 paveikslas. Duomenų bazių modeliai: a) ODB b) RDB.

Navigation link

- Prisijungti vartotojas
- Pasirinkti pacientą
- Pasirinkti tyrimą
- Anamnezė
- Oftalmologiniai tyrimai
- Bendrieji oftalmologiniai tyrimai
- Priekinės akies dalies tyrimas
- Akies dugno tyrimas
- Fotosesijos duomenys
- Fotosesijos duomenys vaizdas
- vo disko sritys aprašas
- Kraunamųjų aprašas
- Tinkamies - gyslainės aprašas
- Akių palyginimas
- Ultragarsiniai tyrimas
- Klinikiniai duomenys
- Diagnozė
- Paisika
- Išeiti

Akies dugno fotosesijos duomenys

Vaidzas pasirinktas! Kairė akis Dešinė akis

EKS ir OND vertikalus diametrų santykis

EKS ir OND horizontalus diametrų santykis

Komentaras

Vykdyti

User interface elements

5 paveikslas. Pakitimai vartotojo sąsajoje.

kadangi pagrindinė užduotis buvo įvertinti duomenų saugojimą ODBVS, bet ne atlikti DB veikimo matavimus. Testavome vartotojo funkcijas (pvz.: įkelti/išrinkti informaciją apie pacientą, peržiūrėti paciento nuotraukas), administratoriaus funkcijas (pvz.: vartotojų registracija), taip pat vartotojo sąsają (pvz.: navigaciją, duomenų gavybos rezultatus). Visos procedūros, kaip ir viskas testavimas, atlikti sėkmingai.

Pakeitimai duomenų bazėje.

Tam, kad vaizdžiai parodyti ODB pranašumus, sukūrėme RDB modelį analogišką ODB klasėms. Pasikeitimai stebėti lyginant vaizdines diagramas. Supaprastinta vienos dalies vaizdinė diagrama pateikta 4 paveiksle.

Iš paveikslo matome, kad ODB modelyje yra saugoma mažiau duomenų, negu RDB. Nereikia saugoti identifikatorių, kadangi sąveika tarp klasių yra atliekama operacija ("EyelImageData()") paveiksle 4 a). Laukas "Filename" (4 b paveiksle) yra nereikalingas ODB, kadangi „blob“ (bet kokia daugialypės terpės ar kita byla) tipo

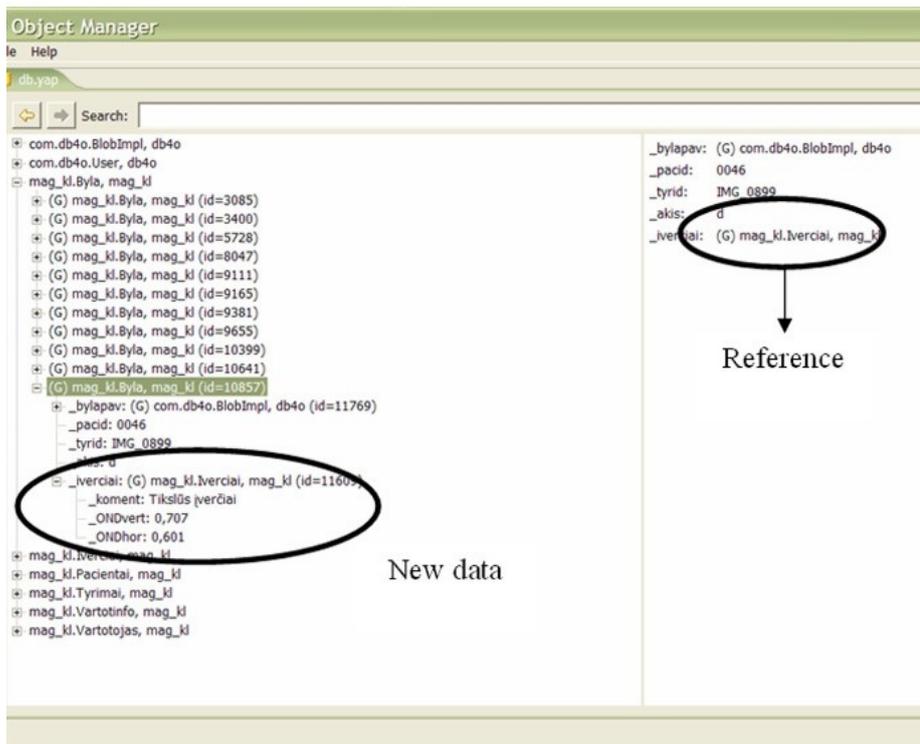
bylos informacija yra automatiškai kaupiama duomenų bazėje. Taigi nėra reikalo kurti papildomus laukus sąryšiu tarp lentelių, be to duomenys nėra dubliuojami, taip pat daugialypės terpės informacijos valdymas yra patogesnis.

Pakeitimai vartotojo sąsajoje.

Vartotojo sąsajos pakeitimai dirbant su RDB ir ODB yra ekvivalentūs, kadangi atvaizduojami tie patys duomenys. 5 paveiksle pavaizduoti šie pakeitimai sukurtos sistemos atveju.

Pakeitimai programos logikos lygmenyje. Pakeitimai šiame lygmenyje yra gerokai sudėtingesni. Abiem atvejams reikiami šie pakeitimai:

- programinis kodas naujiems vartotojo sąsajos elementams,
- testavimas šiems elementams, naudojant papildomus duomenis,
- naujų duomenų įkėlimas.



6 paveikslas. ODB adresinė nuoroda ir nauji duomenys.

Nedetalizuosime kaip ODBVS elgiasi šiuo atveju, tačiau naudojantis ODBVS db4o duomenų peržiūros ir išrinkimo programa „Object Manager“, pastebėsime, kad duomenų bazės struktūros yra automatiškai atnaujinamos, kuomet DB yra sužadinama. Duomenų įkėlimui ir atnaujinimui, pavaizduotam 6 paveiksle, reikalingas toks pats programinis kodas - ir įkėlimui ir atnaujinimui galioja tas pat sakiny. Tačiau, RDB reikalingi du SQL (Struktūrizuota užklausų kalba, Structured Query Language) sakiniai (Insert Into ir Update). Toks sutrumpinimas ODB akivaizdžiai sumažinam programinio kodo kiekį.

Tai pavaizduota 6 paveiksle, kuriame matomi adresinė nuoroda ir nauji duomenys.

Taip pat pastebėjime, kad adresinė nuoroda išskleidžia naujus duomenis, taip išvengiama duomenų dubliavimo.

Baigiant, galime pastebėti, kad neturėtų kilti problemų integruoti šią sistemą į NESS, kadangi NESS šiam tikslui naudoja IHE (Integrating Healthcare Enterprise). Kadangi db4o ODB yra viena byla konvertuojama į XML (eXtensible Markup Language), kuris kartu su HL7 ir CDALevel

2 yra naudojami NESS, neturėtų kilti problemų sistemų bendram darbui. Kiti pastebėjimai ir ODBVS privalumai pateikti sekančiame skyriuje.

Diskusija ir išvados

Tarptautinis bendradarbiavimas dirbant su standartais po truputį veda prie spartesnės produktų kokybės, taip pat ir programinės įrangos plėtotės. Sistemos geba dirbti naudojant skirtingus duomenų tipus, užtikrinant suderinamumą ir lankstumą. Ateityje sukurtą sistemą ketiname integruoti į NESS.

ODBVS analizė parodė, kad ji yra patogi ir pakankamai lanksti kaupti nuolat besikeičiančius duomenis. Naudojant objektines technologijas nesudėtinga įgyvendinti nutolusių DB pasiekiamumą.

Sukurta sistema parodė, kad ji yra pakankamai lanksti ir lengvai keičiama, kadangi naudojamos tinkamos technologijos, kurios tuo pačiu metu yra pajėgios dirbti su mediciniais standartais. Nors sistema galėjo būti sukurta naudojant skirtingas technologijas, mes nusprendėme naudoti .NET, dėl jos plėtojimo, daugiakalbiškumo galimybių, patogios Visual Studio .NET programavimo aplinkos integruotumo.

Eksperimento metu parodyti šie pagrindiniai ODB privalumai prieš RDB:

- galimybė įkelti/atnaujinti duomenis ta pačia komanda,
- sumažėja programinio testo ir klaidos tikimybės,
- esant adresinėms nuorodom tarp objektų nereikia kurti papildomų laukų ryšiams tarp lentelių palaikyti,
- duomenys nedubliuojami,
- lankstesnis daugialypės tarpės informacijos saugojimo mechanizmas,
- pildant DB nereikia kurti ryšių ir lentelių duomenų bazės lygyje.

Eksperimentui atlikti naudojome nedidelį kiekį duomenų, tačiau ateityje galvojame kaupti didesnius kiekius duomenų, ir stebėti laikines DB veikimo charakteristikas.

Taip pat protinga paminėti, kad pateikti rezultatai nereiškia, jog reikėtų atsisakyti RDBVS naudojimo, greičiau tai paskatinimas išbandyti ODBVS - pakankamai lanksčią technologiją sprendžiant skirtingas problemas, ypač susijusias su e-sveikatos sistemomis.

Literatūra

- [1] CEN/TC 251- Standard Development Workgroup site svetainė, prieiga per internetą: <http://www.cen251.org/>
- [2] DICOM duomenų mainų standarto svetainė, prieiga per internetą: <http://medical.nema.org/>
- [3] Estes B. T., Maxime O.: J2EE vs .Net: The choice depends on your needs. 2003; prieiga per internetą: <http://www.computerworld.com/printthis/2003/0,4814,84155,00.html>
- [4] Farley J.: Microsoft .NET vs. J2EE: How Do They Stack Up?. 2000; prieiga per internetą: http://java.oreilly.com/news/farley_0800.html
- [5] Hanson J. J.: .NET versus J2EE Web Services. A Comparison of Approaches; prieiga per internetą: <http://www.webservicesarchitect.com/content/articles/hanson01.asp>
- [6] HL7 - standartus kuriančios organizacijos svetainė, prieiga per internetą: <http://www.hl7.org/>
- [7] Lurie J., Belanger R. J.: Does one Web services platform dominate the other?. 2002; prieiga per internetą: <http://www.javaworld.com/javaworld/jw-03-2002/jw-0308-j2eenet.html>

- [8] Objektinių duomenų bazių kūrėjų ir platintojų svetainė, prieiga per internetą: http://www.service-architecture.com/products/object-oriented_databases.html
- [9] Object Database Management Group svetainė, prieiga per internetą: <http://www.odmg.org>
- [10] ODBVS db4o svetainė, Prieiga per internetą: <http://www.db4o.com/>
- [11] OODBMS ObjectDB site, <http://www.objectdb.com/>
- [12] ODBVS ozone svetainė, prieiga per internetą: <http://ozone-db.org/frames/home/what.html>
- [13] ODBVS SOD2 svetainė, prieiga per internetą: <http://www.tneumann.de/sod/index.html>
- [14] ODBVS XL2 svetainė, prieiga per internetą: <http://www.xl2.net/>
- [15] Objektinės-reliacinės duomenų bazės GigaBASE svetainė, prieiga per internetą: <http://www.garret.ru/~knizhnik/gigabase.html>
- [16] Sessions R.: Java 2 Enterprise Edition (J2EE) versus The .NET Platform Two Visions for eBusiness. 2001; prieiga per internetą: <http://www.objectwatch.com/FinalJ2EEandDotNet.docDotNet.doc>
- [17] Sheil H., Monteiro M.: Rumble in the jungle: J2EE versus .Net. How do J2EE and Microsoft's .Net compare in enterprise environments?. 2002; prieiga per internetą: http://www.javaworld.com/javaworld/jw-06-2002/jw-0628-j2eevsnet_p.html
- [18] Sorli M., Romo J., Stach R., Bredehorst B.: REMOTE: GRD1-2000-25433. 2001, p. 40-42; prieiga per internetą: <http://www.remote-project.org/downloads/REMOTE.D1.1-1.1--2001-07-09.pdf>
- [19] Zamir S.: Handbook of object technology. CRC Press, 1999, 35skr.

Contact**Vita Speckauskiene**

Biomedical Engineering Institute at
Kaunas University of Technology
Studentu str. 65 Kaunas, LT-51359
Lithuania

e-mail: vita.speckauskiene@med.kmu.lt