Knowledge Management for Medical Computational Problem Solving: An Ontological Approach

Charalampos Bratsas[1], Panagiotis Bamidis[1], Paulo Quaresma[2], George Pangalos[3], and Nicos Maglaveras[1]

[1] *Laboratory of Medical Informatics, Medical School, Aristotle University of Thessaloniki, POB 323, 54124 Thessaloniki Greece*
{mpampis, bamidis, nicmag}@med.auth.gr

[2] *Department of Informatics, University of Evora, Evora, Portugal*
pq@di.uevora.pt

[3] *Informatics' Laboratory, General Department of Polytechnics, Aristotle University of Thessaloniki*
pangalos@auth.gr

Abstract. This paper presents an ontological approach that manages the Medical Computational Problem Solving (MCPS) by means of organizing and visualizing their existing knowledge. Medical Computational Problems are medical problems, the solution of which deals with mathematical or statistical models, such as data mining, signal or image processing, etc. Thus to define the MCPS requires the Medical domain of knowledge for the problem definition and additional fields of knowledge for the solution definition. Prototype Med-CPS ontology is proposed to combine, exchange and reuse knowledge representation of the different domains. The structure and the flow of the knowledge that manage Med-CPS ontology by making in-depth analysis of the requirements of a Knowledge Base System that could be associated with this ontology is presented and analysed. The Med-CPS ontology is implemented and tested in protégé.

1 Introduction

Medical algorithms can generally be divided into two main categories, algorithms that refer to medical processes and they are similar to medical prescriptions and algorithms that refer to calculations of parameters and they can facilitate the decision-making. As far as the first category is concerned, these algorithms refer to Diagnostic Problem Solving (DPS) and apply to clinical reasoning strategies and clinical guidelines [1]. Currently, efforts are being made to introduce standardized approaches for DPS representation and sharing [2]. Examples of these representations are Proforma [3], the GuideLine Interchange Format (GLIF) [4], Asbru [5], and EON [6].

The second category of medical algorithms refers to Medical Computational Problem Solving (MCPS) and provides all computational facilities required to solve a Medical Problem. Medical Computational Problems (MCP) are medical problems, the solution of which deals with mathematical or statistical models, such as data mining, signal or image processing, and estimation of parameters. The main characteristic of the algorithms that solve an MCP, is the fact that they are often designed and implemented by non medical scientists. These algorithms may be used as stand-alone, included in practice guidelines or embedded within medical devices. They are implemented in programming platforms known as Problem Solving Environments such as Matlab, MathCAD, Mathematical, SPSS, C++, java , etc.

The main problem in exposing an MCPS is that the existing information about MCPS is scattered and poorly organized. A Knowledge Base System (KBS) can assist algorithm researchers, who are either medical scientists or medical informatics scientists, in their research efforts by providing an infrastructure for organizing and visualizing intellectual landscapes. In order to demonstrate how algorithm researchers can build a KBS associated with MCPS an ontological approach was chosen in this piece of work. More specifically, the classification of MCPS involves the definition of a domain ontology, which semantically describes and associates MCPs, algorithmic solutions, algorithmic implementations, and provides the schema for the construction of a Knowledge Base. The paper further argues on the importance of the users roles and, therefore, provides a framework to model them trough specific technical implementation details.

The remainder of this article is structured as follows: firstly, the requirements that the ontology model has to meet are identified. Secondly, the ontology model construction method is presented, followed by a description of the Med-CPS ontology. Finally, use of the ontology is illustrated and discussed through suitably chosen examples.

2 Identifying the Requirements

The design of the KBS through ontologies inherently involves an in depth analysis of the following questions: to whom this KBS is referred to and what kind of assistance does it provide? What are the goals to be achieved by the KBS? What special requirements must be taken into consideration?

Different users have different needs and the KBS has to provide solutions for each need. Initially, the knowledge authors of the MCPS that record knowledge have to be supported by the system. The knowledge domain that refers to MCP is Medicine.

Thus, in order to fully define an MCP, the contribution of the Medical Researchers (practitioners or other medical scientists) is required. Nevertheless, for the solution of a MCP other fields of knowledge are also prerequisites, such as Statistics, Mathematics, Informatics, Physics etc. Therefore, for the description of the algorithmic solution the contribution of both Medical Researchers and Medical Algorithm Researchers (that hold the additional knowledge) is necessary.

Also, the design of the system has to consider the needs of practitioners or plain users that will search solutions for an MCP.

A KBS associated with MCPS should be able to answer the following questions: Is there a solution for an MCP? Which are the inputs and the outputs of the algorithm that solves the MCP? Is the algorithm solving the MCP implemented? Do similar or related MCP exist? Are there any publications or references for the particular MCPS? Tracking the answers to these questions leads to the outline of the MCPS description characteristics.

Thus, an emerging key attribute of a MCP is the incorporation of medical terminology in its description. Therefore, the medical term description using a term-thesaurus that contains commonly accepted definitions, synonyms and semantic types was adopted so as the system adheres to standard terminologies frequently encountered in these concepts.

3. Building MCPS-KBS via ontologies

Ontologies provide a structure framework for modeling the concepts and the relationships of some domain of expertise. Ontologies support the creation of repositories of domain-specific reference knowledge, or knowledge for communication and sharing among people and applications [7]. In order to create a MCPS Knowledge Base (KB) a prototype - Medical Computational Problems Solving (Med-CPS) - ontology model was designed and implemented by use of the Ontology Web Language (OWL). An OWL, in its abstract syntax, contains a sequence of annotations, axioms, and facts, which convey the main concepts of an OWL ontology and provide information about classes, properties, and individuals in the ontology [8]. The tool Protégé [9] with its OWL plug-in [10] was used to build the Med-CPS Ontology. Protégé is an open source ontology editor.

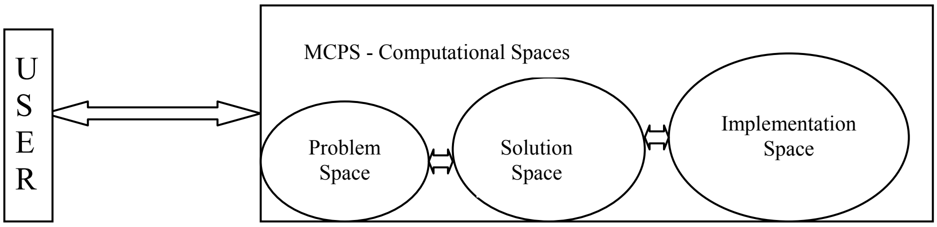

Fig 1. Med-CPS Ontology basic Models

The user space and three computational spaces – medical problem, solutions, and implementation – are proposed in order to provide a detailed accurate description of the Med-CPS Ontology (Fig 1). The problem space consists of medical computational problems; the solution space is composed of algorithm solutions, whereas implementation model objects assist in carrying out the solutions. The user model objects consist of the users and their different capabilities. A high level MCPS KB abstraction can be modelled as a collection of these objects, their intra-relations within each space and their inter-relations across the spaces.

In order to achieve concept consensus, a standardized medical terminology the Unified Medical Language System (UMLS) was utilised, as developed by the US National Library of Medicine since 1986 [11]. The intention is to help health professionals and researchers use biomedical information from different sources, by mapping many existing terminologies within a unifying framework. UMLS is comprised of three knowledge bases: the Metathesaurus, a large repository of concepts, the Semantic Network, a limited network of 134 semantic types, and the Specialist Lexicon which corresponds to lexical resources [11].

To describe MCPS references in a structured manner, the BibTex OWL- Ontology was adjusted into Med-CPS Ontology. The BibTex Ontology is a method of marking up bibliographic data, primarily for use in LaTeX documents, but also useful for generic bibliographic storage [12].

4 Med-CPS OWL Ontology

In this section the architecture of the Med-CPS ontology and the way it meets the requirements mentioned above are presented. The description of the ontology is consisted of two parts; the first part concerns the users and the second the computational Spaces of the Med-CPS ontology MCP.

4.1 Users of Med-CPS Ontology.

The use of a MCPS- KBS serves two purposes; the organization of knowledge pertaining to the MCPS and the mining of the input knowledge.

As a result, the system of knowledge management has two different groups of users, with distinctive roles; those who produce the knowledge (Knowledge Authors) and those who mine it (Simple Users). By using the Med-CPS ontology the groups of users are congregated to different classes and subclasses that include different properties for the different abilities of the users, as demonstrated in figure 2.

Particularly, Knowledge Authors are the ones who can describe the MCPS (all three Computational Spaces). Thus, in Med-CPS Ontology the KnowledgeAuthors class includes the Object Property *describes,* which relates the individuals of KnowledgeAuthors class with the individuals of the MCP Class or Algorithm Class or Implementation Class. Simple Users have only got the authority to search for solutions to a particular MCP. Thus, the SimpleUsers Class includes the SearchMCP_Description, which is a DataType Property (free text).

In many cases, in order to define a problem and its solution, different domains of knowledge are required. An MCP always refers to Medical Domain of Knowledge, whereas the algorithm can also refer to other fields of knowledge. For example, the MCP "recover missing parts from ECG for monitoring congestive Heart Failure Patients" is a medical problem that has been defined by Users that were familiar with the Medical Domain of Knowledge but they were incapable of solving the problem or describing the solution. The particular MCP has come up from a citizens' health care system, where heart failure patients communicate from their homes via a variety of interfaces and transmitting their ECG. Due to the unsatisfactory level of ECG transmission, the ECG signal had missing parts and, as a result, the practitioners had difficulties in ECG processing. This problem was resolved by Users who are familiar with another domain of Knowledge via the "Filling Time Series - FiTS" algorithm [13].

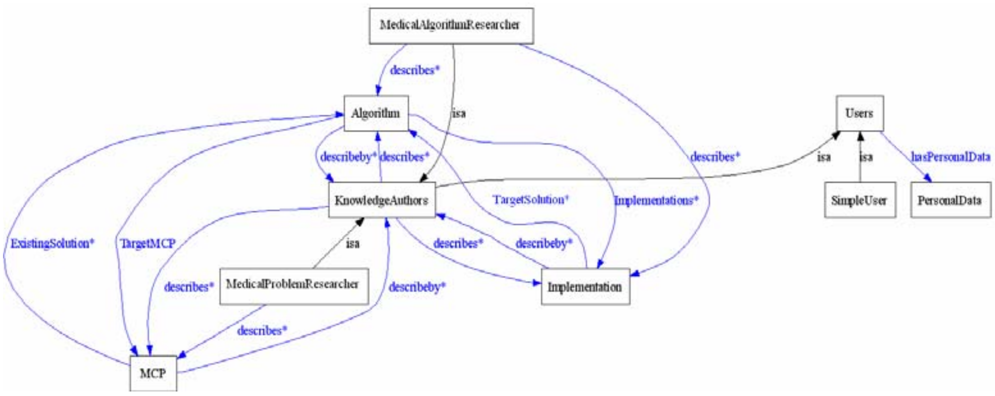

Fig 2. Users Model in Med-CPS Ontology. Hierarchy of Users and their distinguished roles.

Taking into account the hypothesis of the previous paragraph, the KnowledgeAuthors can be categorised to Knowledge Problem Researchers (the ones that define the MCP) and the Knowledge Algorithm Researchers (the ones that define the solutions of the MCP). In the Med-CPS ontology, the KnowledgeProblemResearcher class and the KnowledgeAlgorithmResearcher subclasses of KnowledgeAuthor Class are defined by the use of the unionOf restriction. Analytically, a Knowledge Author can be a Medical Problem Researcher, or a Medical Algorithm Researcher or both. OWL restrictions' component, namely allValueFrom, was employed in order to define the different abilities of these Users. As a result, when the property *describes* is included in the MedicalProblemResearcher class, its values must be of a certain type of MCP ($\forall$ describes.MCP) and when it is included in the MedicalAlgorithmResearcher class, its values must be of a certain type of Algorithm or Implementation ($\forall$ describes.(Algorithm $\sqcup$ Implementation)).

4.2 Computational Space - Med-PCS Ontology

The MCPS- KBS should incorporate knowledge that is able to manage information about MCPs and their solutions. All the information required for that scope are organized via Med-CPS ontology and are presented in OWL-Structure providing Knowledge Objects of the KBS.The Knowledge Objects are formulated into classes of the computational spaces of the Med-CPS ontology and the attributes of the Knowledge Objects are the properties of these classes (Fig 3).

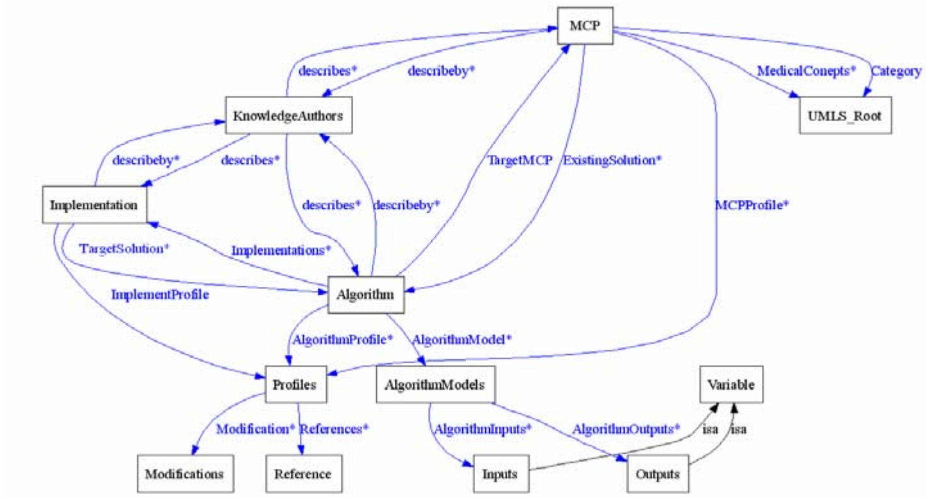

Fig 3. Partial visualization of the Computational Spaces of Med-CPS Ontology.

Table 1 shows the three major classes; MCP, Algorithm, Implementation of the Med-CPS computational spaces, as well as their key properties that associate these classes with others. In additional, the significance of each association is given. In the following paragraph additional Med-CPS Ontology characteristics are given.

Table1: The Main Knowledge object of Computational Spaces of Med-CPS Ontology.

Knowledge Objects as domain Classes	Object Properties that links the individuals between two classes	Linking Classes	Meaning
MCP classs - main class of MCP space	Describeby	KnowledgeAuthor	By whom the MCP is described
	ExistingSolutions	Algorithm	Which is MCP's solution
	MCPProfile	Profile*	Name, Language, Date, Description, Modification, References
	MedicalConcepts and Category	UMLS_ROOT	Every Medical Concept or Category description are following the UMLS description
Algorithm Class – main class of solution Space	Describeby	KnowledgeAuthor	By whom the algorithm is described
	TargetMCP	MCP	The MCP which the algorithm solves
	AlgorithmProfile	Profile*	Name, Language, Date, Description, Modification, References
	Model	AlgorithmModel	The input and output variable information and the algorithm pseudo code
	AlgorithmImplementation	Implementations	Which is the executable program
Implementation Class – main class of implementation Space	Describeby	KnowledgeAuthor	By whom the implementation is described
	TargetSolution	Algorithm	Which algorithm is implemented
	ImplementationProfile	Profile*	Name, Language, Date, Description, Modification, References

*(in the linking classes): indicates the existence of a particular class into the three computational spaces (1st column).

Medical Concepts can be used as MCP Keywords. For example, the Medical Concepts for the MCP "Estimate the risk mortality after Myocardial Infarction" are the following: Mortality Vital, Myocardial Infarction, which are also the Keywords of the problem. The UMLS medical concepts definitions are automatically received from UMLS-KS (UMLS Knowledge Server) into the UMLS_Root class. The description of each concept is formulated as: a Concept Unique Identifier, a definition, broader relationships, narrower relationships, semantic type and synonyms. The connection with UMLS-KS takes place through the UMLS-plug-in of protégé [14].

The algorithm Class also includes the Data Type Property ProblemSolvingStrategy, a free text that includes keywords of problem Solving Strategies that is adopted by an algorithmic solution (Statistical, or Data Mining or Image Processing etc).

The inputs and outputs of the algorithm are described by the individuals of the Variable class of Input and Output classes, which are subclasses of the former, as figure 3 illustrates. A description of the Variable includes the following features: Name, Data Type, Measurement, and Description. The Input and Output descriptions include the characteristics of a variable plus the PreConditions and the PostConditions, respectively

In the Implemention Class, includes also the following Properties: Programming Language, ExecutionEnviroment, ImplementationResources.

In the relations among the three computational Spaces the inverse property is used. More specifically, the TargetProblem property is stated to be the inverse of the property and the ExistingSolutions and the TargetSolution property is stated to be the inverse of the property Implementations. The first relation is used to associate the Problem Space with the Solution Space and the second relation is used to associate the Solution Space with Implementation Space. So, if X is related to Y by the TargetProblem property or TargetSolution property, then Y is related to X by the ExistingSolutions property or Implementations property respectively. Finally the Reference class is the main class of the BibTex Ontology.

5. Med-CPS Ontology Employment

In order to evaluate the Med-CPS Ontology by the use of Protégé-3 twenty-three cardiology algorithms and their MCPs were described. These formulated descriptions provided a prototype CMPS-KB. The major of them (twenty) were originated by the Institute of Medical Algorithms. For more than eight years, the Institute has collected and processed over 4,000 algorithms, including about 200 cardiology-specific algorithms that are available on the following URL www.medal.org [15].

Some typical such MCPSs include:

- Calculation of the body mass index (BMI)
- Calculation of anatomic areas and volumes from echocardiographic data
- Application of the Framingham data for predicting risk of cardiovascular disease
- Preoperative risk assessment prior to cardiac or noncardiac surgery.
- Identification of prognostic factors for short and long term prognosis following myocardial infarction.

The other three algorithms were developed and implemented in the Medical Informatics Laboratory of Aristotle University of Thessaloniki. The MCPs of these algorithms arised during the project research called Citizen Health System CHS [16], where the Patients communicate with a system via a variety of interfaces, like public telephone, internet or a mobile device (through WAP) and submit values of their vital parameters and their ECGs with the help of electronic micro devices (transmission to the contact centre along with some yes/no answers to simple questions regarding mostly the occurrence of certain symptoms). The first two implemented algorithms “Uncertainty Rule Generation (URG)” and “Adaptive - URG” [17] are referred to the MCP: “identify the possible relations among the measurements of a Heart Failure Patient and his/her symptoms”. The last implemented algorithm “FiTs” is referred to the MCP “recover missing parts from ECG for monitoring congestive Heart Failure Patients” [13].

The Med-CPS ontology in collaboration with a tool which includes components such as authorization of the users, search knowledge engine, and knowledge editor constitutes a KBS MCPS.

A number of the scenarios may be devised to evaluation such a system. In this case these scenarios are linked with two postgraduates programs run at Medical School of Aristotle University namely, MSc Medical Informatics and MSc in Medical Research Methodology. The former mainly attended by students of exact technological fields, while the latter of Medical Students. Through these pools of users the tow category Authors of KBS MCPS can be achieved. So, a pool of medical trainees, who provide the knowledge of a specific medical domain, are requested to check for availability of algorithms and a pool of postgraduate students taking courses in processing, simulation, data mining, etc, are requested to address whether specific algorithms resolve medical problems, which are taught in medical sciences courses.

6. Conclusions and future work

Medical CPS or their description can be found on the web or in documents; however, in most of the cases they are not structured and related to the problems that are able to solve. In this paper, a Medical Problem-Centered collaborative knowledge management architecture is proposed that was designed using ontologies. The Med-CPS Ontology uses four models to create formal knowledge objects that answer the MCPS researchers' needs; Firstly, the users model which corresponds to the different abilities of the users. Secondly, the problem model which corresponds to well-defined medical computational problems and it is the heart of the Med-CPS Ontology. Thirdly, the solution model, which corresponds to existing algorithm solutions and finally, the implementation model, a practical view of the MCPS research domain that corresponds to existing implementations.

The representation of the Med-CPS Ontology has proved that most of the needs of MCPS KBS are covered employing an ontological approach. The advantages of this approach are:

1. The Knowledge Object (formalized) and content objects (free-formatted) were constructed using ontologies. Thus, the prospect to organize the existing information about MCPS in a structured manner was exploited.
2. Knowledge Objects that represent MCP can be used in order to achieve an efficient search process for algorithmic solutions of particular MCP. In particular, MCP is related with the existent algorithmic solutions. Furthermore, the medical concepts and the category formulated by a UMLS play a significant role during the search process as Keywords.
3. Knowledge Objects that represent solution implementations can be used to mediate interactions between users and conceptual algorithmic solution. They are useful for both academic and corporate industrial applications.
4. Access rights that are defined by the User Model are enforced for content objects, so as to allow a personal workspace for each user, which can be used to store contact object and to exchange information with other members.

Most of the requirements of the MCPS - KBS were met by using the ontological approach and the Med-CPS Ontology. The unanswered questions, still under investigation are: the determination of a new MCP and the interrelation of MCPs. In order to overcome these obstacles, the construction of an algorithm that defines vectors by the individuals of the MCP Model classes is researched. Specifically, the Pearson Correlation or the Cosine Based Similarities or the Euclidean Distance among these vectors can provide the semantic distance of the MCPs.

Clearly, and due to its complexity, all modules have aspects that may be improved such as the reengineering of Med-CPS ontology, the description of further MCPS, and the creation of an algorithms search engine, where medical researchers may formulate queries in natural language and of a system that interprets and answers by taking into account Med-CPS ontology.

The final aim of this project is to build a Medical Computational Problem Solving Knowledge Portal. This Portal will constitute Problem Solving Environment (PSE) which will provide all the computational facilities necessary to solve a target class of Medical Problems. Thus, Medical Scientists and Algorithm Scientists will be able to organize and perform MCP Solutions. This portal could be used in clinical research or more broadly in Medical research, by providing existing solutions of MCPs. Moreover it could be utilized as a Problem Based Learning tool or as a collaborative learning tool, by providing the knowledge of the MCPS.

Acknowledgement

This work was supported by PROMESIP and IRAKLITOS both funded by the Greek Ministry of Education (EPEAEK I-II).

References

1. Jose F. Arocha, Dongwen Wang, Vimla L. Patel.: Identifying reasoning strategies in medical decision making: A methodological guide. Journal of Biomedical Informatics 38. (2005) 154-171
2. Paul A. de Clereq, Arie Hasman, Johannes A. Bolm, Hendrikus H.M Korsten .:The application of ontologies and problem-solving methods for development of sharable guidelines. Artificial Intelligence in Medicine 22. (2001). 1-22.
3. Fox J, Johns N, Rahmanzadeh A. Disseminating medical knowledge: the *Porforma* approach. Artificial Intelligence in Medicine 14. (1998). 157-81.
4. Patel VL, Allen VG, Arocha JF, Shortli EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. J Am Med Inform Assoc 1998; 5: 467-83.
5. Shahar Y, Miksch S, Johnson P. The asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. Artificial Intelligence in Medicine 14. (1998). 29-51.
6. Tu SW, Musen MA. A flexible approach to guideline modeling. Proc AMIA Symp (1999):420-4.
7. Davies J, Fensel D, Harmelen F.: Towards the semantic web, Ontology-driven Knowledge Management , Wiley & Sons Ltd, (2003).
8. World Wide Web Consortium. OWL Web Ontology Language Semantics and Abstract Syntax.W3C Recommendation 10 Feb, 2004.
9. N. Noy, M. Sintek, S. Decker, M. Crubezy, R. Fergerson, and M. Musen. Creating semantic Web contents with Protégé-2000. IEEE Intelligent Systems, 2(16):60–71, 2001.
10. H. Knublauch, M. A. Musen, and A. L. Rector. Editing description logics ontologies with the Protégé OWL plugin. In International Workshop on Description Logics, Whistler, BC, Canada, 2004.
11. Unified Medical Language System UMLS http://www.nlm.nih.gov/research/umls/
12. BibTex Definition in Web Ontology Language (OWL) Version 0.1 (2004). hhtp://visus.mit.edu/bibtex/0.1/
13. Konias, N. Maglaveras, I. Vlahavas, “Predicting Missing Parts in Time Series Using Uncertainty Theory”, Biological and Medical Data Analysis: 5th International Symposium, ISBMDA 2004, J. M. Barreiro, F. Martin-Sanchez, V. Maojo, et al. (Ed.), Springer-Verlag, 3337, pp. 313-321, Barcelona, Spain, November 18-19, 2004, 2004.
14. UMLS plug-in. http://protege.stanford.edu/plugins/umlstab/documentation.html
15. Svirbely J, Sriram MG. MEDAL, a compendium of medical algorithms for access over the Internet. Proc AMIA Symp. 1999;p1172S
16. Maglaveras, N., Koutkias, V., Chouvarda, I., Goulis, D.G., Avramides, A., Adamidis, D., Louridas, G., Balas, E.A.: Home Care Delivery through the Mobile Telecommunications Platform: The Citizen Health System (CHS) Perspective, International Journal of Medical Informatics 68, (2002) 99-111
17. Konias, S., Giaglis, G.D., Gogou, G., Bamidis, P.D., Maglaveras, N.: Uncertainty Rule Generation on a Home Care Database of Heart Failure Patients. In Proc. of Computers inCardiology, IEEE Comp. Soc. Press, (2003), 30, 765-768