Semi Automatic Ontology Based Knowledge Extraction and Search for Medical Algorithms From Web Documents

Charalampos Bratsas, Lab of Medical Informatics, Aristotle University of Thessaloniki, Greece
mpampis@med.auth.gr
Evangelos Kaimakamis, Lab Of Medical Informatics, Aristotle University Of Thessaloniki, Greece
Panagiotis Bamidis, Lab Of Medical Informatics, Aristotle University Of Thessaloniki, Greece
Paulo Quaresma, University Of Evora/Informatics Department, Evora, Portugal
Nicos Maglaveras, Lab Of Medical Informatics, Aristotle University Of Thessaloniki, Greece

Introduction: KnowBaSICS-M (Knowledge Base System for Intelligent Computational Search in Medicine) is a Knowledge Base System providing basic help to clinical doctors and algorithm researchers, when attempting to reach the algorithmic solutions of Medical Computational Problems (MCPs). It is a tool for providing unstructured and widely scattered medical information related to computational algorithms by enabling authors to insert new MCPs and algorithmic solutions into the Knowledge Base (KB). KnowBaSICS-M also links a knowledge extraction tool with MCP ontology. The extraction tool extracts knowledge from particular online documents that refer to algorithmic solutions of MCPs (like www.medal.org) by matching it with the given classification structure of MCP Ontology.

Methods: MCP Ontology has proved that most of the needs of MCP KBs are covered by employing an ontological approach by using Ontology Web Language (OWL). The code development was based on open-source development platforms, tools and technologies, like the 'Protege' ontology editor, the 'Eclipse' open- source Integrated Development Environment, the Unified Medical Language System (UMLS), a medical lexical knowledge source and a set of associated lexical programs. One of the most important functions is the search process. For effective searching, Information Retrieval algorithms estimating the similarity among MCPs and requested MCPs were used, while automatic extraction of terms as keywords for the Quested MCP was performed through the UMLS Metathesaurus.

Results: Knowledge Author Users using the advance Java Client can have full access and manage the MCP Ontology and the clinical user can search for algorithmic solutions and implementations by describing the MCP in a natural language upon using the Web Client. For the search process the Similarity Threshold is set at around 47%, since the Precision and Recall characteristics are approximately 81% and 72% respectively, numbers that are considered quite satisfactory. Discussion: The high rates of precision and recall characteristics of the search processes shown by the system may be attributed to the relatively small total number of MCPs present in the current version of the KB and also to the precise and careful MCPs' description exhibited by our test users. Though our initial experiments are showing promising results, a more extensive evaluation is needed, in order to complete and extend these first observations. The overall performance of KnowBasics-M in terms of usability, user friendliness and precision / recall features is satisfactory. Thorough use of such an application is expected to enhance task automation, cost containment and quality services in medical care, while at the same time medical research and high quality medical education by means of focused problem based learning are going to be benefited at a considerable rate.

References:
1. C. Bratsas, P. Bamidis, P. Quaresma, G. Pangalos and N. Maglaveras, An Internet Application to Manage an OWL Ontology for Medical Computational Problem Solving, Proceedings of MedNet 2005 - 10th World Congress on the Internet in Medicine, A. Engin and P. Niederer (eds.), Special Issue: Technology and HealthCare, IOS Press, (ISSN 0928-7329), 365 - 366, Prague, Czech Republic, 2005
2. C. Bratsas, P. Bamidis, P. Quaresma and N. Maglaveras, Similarity Estimation among OWL Descriptions of Computational Cardiology Problems in a Knowledge Base, Proceedings of CinC 2005 -- Computers in Cardiology, IEEE, 101-104, Lyon, France, 2005.
3. Ch Bratsas, P Quaresma, G Pangalos, N Maglaveras, Using Ontologies to Build a Knowledge Base of Cardiology Problems and Algorithms, In Proc. of Computers in Cardiology, IEEE Comp. Soc. Press, 2004; 31: 609-612.
4. Medscape URL: http://www.medscape.com/viewarticle/523074_4