

Web-based CBR System for Support Medical Diagnosis

Alexander Stamenov and Ivan Koychev¹

Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski"
astamenov@gmail.com, koychev@fmi.uni-sofia.bg

Introduction

From the early days of development of Artificial Intelligence, there was a strong in-terest in applications in the area of medicine. The interest was strong enough to form a separate branch in the early '80s entitled Artificial Intelligence in Medicine (AIM). DXplain [1] is an illustration of system from this early period. It is an internal medi-cine expert system developed at the Massachusetts General Hospital that is still in use at a number of hospitals and medical schools, mostly for clinical education purposes.

Rather than using an expert system approach or another rule-inferring paradigm we decided to employ a Case-based Reasoning (CBR) methodology [4]. Storing and searching among past cases has an advantage in complex domains where it is difficult to create a global theory that explains most of the existing cases.

MediCase - a Web-based CBR Medical Diagnostic System

The MediCase is a web-based, CBR medical diagnosis system. The choice of a web based architecture is implied by the inherent distributed character of the Internet and thus satisfying a desire for a broader availability.

Another important factor making the existence of web based software possible is the great availability of technologies such as Ajax [2], capable of converting a web browser into a thin client. As most web based software the system has three main lay-ers: *client layer* – inside a web browser, *server layer* and *database layer* – in our case we can call it *case-base layer*.

The client side of the application is built around the Java based Google Web Tool-kit framework (GWT) [3]. GWT is making possible the creation of rich AJAX capa-ble user interfaces relatively easy.

In the case-base layer cases are kept as pairs of attribute lists – one describing the problem and one its solution. The main flow of interaction with the system incorpo-rates the general cycle of CBR: Retrieve-Reuse-Revise-Retain [4].

After a physician is logged in he is navigated to the main query screen. On that screen the query case is created. This is done by populating with values the attributes needed for the query, changing the weights of their relevance and selecting similarity function where appropriate. Further more the MediCase will suggest to the physicians some appropriated adaptations of the retrieved cases. They can take actions to further adapt the proposed solution if it doesn't

¹ Also associated with Institute of Mathematics and Informatics - BAS.

justify his or hers expert opinion. The final adapted solution for the query case can be returned to the dataset of cases carrying a tag describing its adaptation process (automatic or manual).

Another set of use cases of interaction with the system are those regarding its ad-ministration. The datasets of cases can be added or augmented by batch imports of large sizes or when adding a single adapted case by the main flow of interaction.

MediCase provides the user with opportunity to choose the similarity measure be-tween cases. The first prototype implements only Euclidian and Manhattan distances. Both are used in their weighted forms giving another level of flexibility. A change in the weight of an attribute regulates its importance to the similarity of the cases. Going even further into refining the similarity metric MediCase allows different similarity functions to be associated with each attribute, according to its type. For example, for a nominal attribute, matrixes of similarity between values can be defined and used in-stead of the generic ones provided by the system.

The kNN algorithm is known for its completeness but its speed is affected by the size of the number of cases and the number of dimensions in which the cases are de-scribed. To improve the scalability during case retrieval a *kd-tree* [5] can be used. A *kd-tree* is a useful indexing structure in the form of a binary tree. It can store a finite set of points from a k-dimensional space. Each node defines a hyperplane dividing one of the ‘k’ dimensions.

Conclusions

The MediCase system aims at providing the physician with suggestions for similar cases and in this way helping her/him in giving more precise diagnosis.

An important direction for future development of the system is to make it more “social”. This could be accomplished by providing the users with an online environ-ment for discussion over cases. The quality control of the new case will be maintained by discussion in the community and/or peer reviewing, of the contributed cases.

Implementation of kd-Ball trees in addition to the classical kd-Trees is also consid-ered for further enrichment of the search module.

References

1. Coiera, E.: Clinical Decision Support Systems. Guide to Health Informatics 2nd edition. Hodder Arnold (2003)
2. Garrett, J. J.: Ajax: A New Approach to Web Applications, <http://www.adaptivepath.com>
3. Google Web Toolkit, <http://code.google.com/webtoolkit>
4. Aamodt, A., Plaza E.: Case-Based Reasoning: Foundational Issues, Methodological Varia-tions, and System Approaches. Artificial Intelligence Communications. IOS Press, vol. 7, 1, pp.39--59 (1994)
5. Moore, A.W.: An introductory tutorial on kd-trees. Technical Report No. 209, Computer Laboratory, University of Cambridge, (1991) http://www.ri.cmu.edu/publication_view.html?pub_id=2818