Medical case-based reasoning for detecting inborn metabolic defects

Töpel T, Neumann J, Hofestädt R
Bioinformatics Department, Bielefeld University, Germany
thoralf.toepel@uni-bielefeld.de

Introduction Inborn defects of metabolism are an important field of molecular genetic and clinical research. Together with the fast progress of biotechnological research, the number of diseases that are clearly due to genetic defects increases too. On the basis of modern newborn screening methods, a part of these inborn diseases can be diagnosed and treated at an early stage. In addition, molecular genetic investigations are used by proving the assumed genetic defect for a higher quality of the diagnosis and to enable a precise therapy. These investigations can be supported particularly by methods of computer science.
Expert or knowledge-based systems are already part of successful artificial intelligence research. Most of these systems use rules, frames or clauses to formalise stored information. But an intricate process of knowledge elicitation is required that is often referred to as the knowledge elicitation bottleneck. Other problems in developing knowledge-based decision support systems are the required special skills and the difficult maintenance of the implemented system. Health professionals require their expert knowledge by contact with similar problems over many years. A professional keeps in mind the resulting conclusions, applied methods and practices particularly in context of solved tasks. According to this, the case-based reasoning (CBR) approach uses old experiences to understand and solve new problems. The CBR system remembers a previous situation similar to the current one and can adapting a previously successful solution to that problem. Because of this, the case-based reasoning approach is especially qualified for use in medicine. CBR is mainly applied for diagnostic and therapeutic tasks and has proved its suitability for medical knowledge-based systems [1].
In this contribution we will present a medical case-based reasoning module to support the diagnosis of inborn defects of metabolism. This tool is a component of the Rare Metabolic Diseases Database RAMEDIS [2] and accessible at http://www.ramedis.de/cbr/. The collection of relevant patient data in electronic case reports and its analysis by case-based methods will be helpful for diagnosis and therapy of inborn metabolic diseases.
System Architecture and Methods The design and development of our system is characterised by focusing to a high comparability and a comprehensive description of the stored cases on the one hand. Otherwise, the system realises low response times and intuitive use. Base of our system is a data model that contains a large number of pre-defined attributes like a controlled vocabulary. A set of these attributes describes a case. Within its description are varied information about main data, laboratory parameters, symptoms, therapy information and molecular genetic data. The controlled vocabulary is maintained by a system administrator that adds new, not existing attributes on demand. This procedure ensures the prevention of using homonymous or synonymous medical terms in the case description.
The basis of our system is a relational database system that stores the cases in form of patient case reports. These reports are collected by health professionals, which submit the data of their currently treated patients with appropriate diseases. An alternative way to expand the case base is the collection of relevant medical publications. Scientific researches can use our system by a web-based graphical user interface that is connected to the CBR module, receives user requests and presents its results. Furthermore, the database is used by other software applications. The medical case-based reasoning module controls the CBR functions and executes queries on the database. Our system implements a case-match CBR that searches for similar cases and leaves the adaptation of the proposed solution to the user.
The implemented retrieval phase of the CBR cycle selects a number of potential solution candidates from the case base of RAMEDIS. When this phase uses a lot of cases, the processing may result in a high computational complexity. Therefore, a pre-selection of cases is useful. This is done by using the MAC/FAC model [3] that pre-selects in a first step (many are called) a number of cases from the case base by relational retrieval. The typically used indexing mechanisms and queries of this pre-selection are based on the structures and methods of relational database systems. A case becomes an element of the candidates set if there is at least one attribute of the case that matches to the problem situation (partial identity). In the following phase (few are chosen), the similarity of the pre-selected cases and the current problem situation is calculated. This calculation bases on the weighted average of the single similarities and on different distance functions depending on the attribute types like Euclid distance or distance table.
Results As described before, the access to the system is enabled by a graphical user interface via WWW at http://www.ramedis.de/cbr/. This interface is based on dynamically generated web pages that can be presented by every common web browser like Mozilla Firefox or Internet Explorer. The user selects a number of features from the controlled vocabulary to describe the problem situation most suitable. Afterwards, the retrieval module runs the case-match CBR algorithm on the case base and calculates a list of similar cases. Detailed information on single similarities for each feature are given to explain the calculation process to the user. With this list, the user can select solution candidates and use diagnosis and therapy information to consider the suitability of the selected case in comparison to the current problem situation.
Conclusion Medical case-based reasoning is useful in supporting diagnosis and therapy, because the CBR approach is especially suited for processing medical knowledge based on experiences. This type of knowledge is hard to describe in form of clauses and rules. Because of this, we applied methods of CBR to existing medical data on inborn metabolic diseases that is stored in the RAMEDIS database for mutations and their phenotypes [2]. The data derives from health professionals and medical publications. So far the case base contains more than 700 case reports that are characterised by altogether 4200 symptoms, 25500 laboratory values. Our system prototype is used intensively by clinical partners that prove the applicability of our system. We plan to extend our system to cover more characteristics and to integrate cases from related internet databases.
This work was supported by the German Ministry of Education and Research in the German Human Genome Project (Project "Modelling of gene regulatory networks for linking genotype-phenotype information") by the grants 01KW9912 and 01KW0202.

References

[1] R. Schmidt and L. Gierl. Case–based reasoning for medical knowledge-based systems. In A. Hasman, H. Prokosch, B. Blobel, J. Dudeck, G. Gell, and R. Engelbrecht, editors, *Medical Infobahn for Europe*, pages 720–725. IOS Press, 2000.
[2] T. Töpel, U. Scholz, U. Mischke, D. Scheible, R. Hofestädt, and F. Trefz. Supporting genotype-phenotype correlation with the rare metabolic diseases database Ramedis. *In Silico Biology*, 2(3):407–414, 2002.
[3] K. D. Forbus, D. Gentner, and K. Law. MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19(2):141–205, 1995.