Multiple Sclerosis Diagnoses - Case-Base Reasoning Approach

V. Kurbalija, M. Ivanović, Z. Budimac
Department of Mathematics and Informatics, University of Novi Sad, Trg D. Obradovića 4, 21000 Novi Sad, Serbia
e-mail:{ kurba, mira, zjb }@im.ns.ac.yu

M. Semnic
Institute of Neurology, Medical School, H. Veljkova 1-3, 21 000 Novi Sad, Serbia
e-mail: semnicmr@EUnet.yu

Abstract:

Case-Based Reasoning (CBR) is the area of the artificial intelligence in which the new problems are solved by adapting the solutions of the previously successfully solved similar problems. Medicine is a suitable domain for application of CBR because the knowledge of medical experts consists of mixture of textbook knowledge and experience, which consists of cases. Architecture of the system for the diagnoses of multiple sclerosis disease, based on CBR is proposed in this paper.

1. INTRODUCTION

Case-Based Reasoning (CBR) has become successful technique for knowledge-based systems in different domains. This promising technique is based on use of previous experience in form of cases to better understand and solve new problems in particular domain. The main idea is an assumption that, for many particular domains, similar problems usually have similar solutions.

Medicine is rather suitable domain for application of CBR because the knowledge of experts consists of mixture of textbook (objective) knowledge and experience (subjective) knowledge. These two sorts of knowledge can be clearly separated, and represented in an appropriate manner. Objective knowledge can be represented in forms of rules or functions, while subjective knowledge is contained in cases. CBR has been mainly applied in medicine for diagnostic and partly for therapeutic tasks.

In this paper architecture for CBR decision support system for the diagnoses of multiple sclerosis disease is described and the current state in the implementation is presented. The rest of the paper is organized as follows. In section 2 foundations of CBR technology, are mentioned. Section 3 explains some problems of using CBR technology in medical domain. In section 4 the architecture of realized system and one simple example are shown, while section 5 concludes the paper.

2. FOUNDATIONS OF CASE-BASED REASONING TECHNOLOGY

Generally speaking, case-based reasoning is applied for solving new problems by adapting solutions that worked for similar problems in the past [3,12]. Basic scenario for mainly all CBR applications looks as follows: In order to find a solution of an actual problem, one looks for a similar problem in an experience base, takes the solution from the past and uses it as a starting point to find a solution to the actual problem.

In CBR systems experience is stored in form of cases. The case is a recorded situation where problem was totally or partially solved, and it can be represented as an ordered pair (problem, solution). The whole experience is stored in case base. Each case represents some previous episode where the problem was successfully solved.

The main problem in CBR is to find a good similarity measure – the measure that can tell in what extent the two problems are similar. In the functional way similarity can be defined as a function sim : $U \times CB \rightarrow [0\ ,\ 1]$ where U refers to the universe of all objects (from a given domain), while CB refers to the case base (just those objects which were examined in the past and saved in the case base). The higher value of the similarity function means that these objects are more similar [12].

The main advantage of this technology is that it can be applied to almost any domain. CBR system does not try to find rules between parameters of the problem; it just tries to find similar problems (from the past) and to use solutions of the similar problems as a solution of an actual problem. So, this approach is extremely suitable for less examined domains – for domains where rules and connections between parameters are not known. In some extent – medicine is one such domain, especially the diagnosis of rare or disregarded diseases, like multiple scleroses. The second very important advantage is that CBR approach to learning and problem solving is very similar to human cognitive processes – people take into account and use past experience to make future decisions.

3. SOME OPEN PROBLEMS OF CBR SYSTEMS

Although CBR technology has many positive aspects, there are still a lot of open problems in building commercial systems. Current commercial systems are mostly oriented towards acquisition and retrieval of cases and, its simple adaptation and evaluation. The well-known open problems appearing in CBR systems are:

1) **Retrieval/selection**
2) **Memory organization**
3) **Matching**
4) **Adaptation/validation**
5) **Forgetting**

Although all of these problems are important, here just problems important for medical domain will be discussed. Some authors [15] think that the main problem of CBR medical oriented systems is the **adaptation** task. In medicine, adaptation can be a serious problem, because cases often consist of an extremely large number of features. Some adaptation techniques, typical for medical domains [15], are: *Focusing on retrieval, Combination with rule-based systems* and *Generalized cases.*

The next problem in medical case-based reasoning systems is avoidance of infinitive growth of case base i.e. problem of **Forgetting**. Although every case in the diagnoses process is important, it might be redundant for the system, because maybe the system already has the knowledge to propose that diagnoses, or maybe some similar or more general cases already exists in the system. Maintenance of the system could be implemented using case properties [4,14].

For solving the problem of **Memory Organisation,** the special memory structure was used - Case Retrieval Net (CRN) [2, 11, 13]. Case Retrieval Nets are memory model that has been developed for efficient retrieval in large case bases. This net has nodes for each case and nodes for each value of every feature - these values are called *information entities*. There exists an acceptance arc between every two corresponding information entities. Also there exists a relevance arc from information entity nodes to the case nodes. All the arcs in the net are weighted with appropriate values.

4. "CaBaGe" – A DECISION SUPPORT SYSTEM BASED ON CBR

During the research in the area of CBR [2,5,6,7], one natural demand came out: it would be very useful to have some basic, core system (framework) that can produce decision support systems in different domains. "CaBaGe" (Case Base Generator) [8,9,10] was the direct result of these intentions.

The main characteristic of this system is that it is domain independent. The input for the system is the description of the database and the database from any domain. On the basis of those data, system creates Case Retrieval Net (CRN) and it is capable to solve new problems (or to propose solutions) from a domain of the input database.

The system reads the data from two input files. In the first input file ("Case Pattern File") the description of the case is stored, while the second file ("Case Base File") contains the list of all already solved cases.

Together with the reading of the case base file, system creates case retrieval net (CRN). The broad structure of CRN is given in the figure 1. Two main parts of the CRN are array of attributes and list of solutions. The array of attributes is created using case pattern file, while the list of the solutions is created from case base file – solution is the last value from every case. Every attribute consists of its name and the list of values. Every value for every attribute represents one information entity because the information entity is an ordered pair *(attribute, value)*. Every value (or information entity) contains the list of arcs to the solution nodes. The arc is given with its weight and a pointer to the solution node. Weights of the arcs between the information entity node *e* to the solution node *c* represent the value of the relevance function between *e* and *c*. These weights are calculated as a number of cases (from the case base file) that contain the information entity *e* and whose solution is *c*.

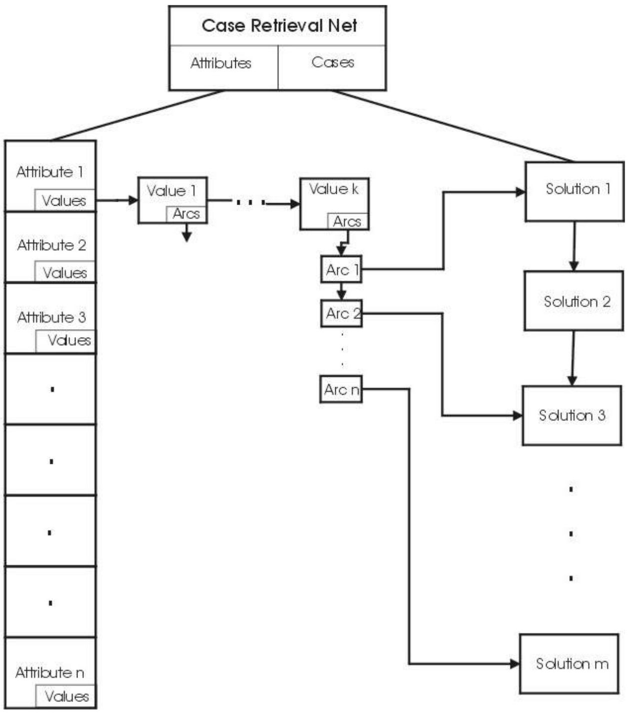

Figure 1. Structure of CRN

After creating the CRN the system expects from the user to enter the current problem (query). Since the query and the case have the same structure (a set of the information entities) the user has to enter the values of some (or all) attributes in a form. In order to better describe the problem, the user should enter all known values of the attributes although it is not necessary. The form contains one more field for every attribute – importance. The importance is the value from the interval (0,1), and describes how much is the user sure in the validity of the value of the attribute he is entering. Value 1 means that he is 100% sure that the data are valid, while the value 0 means that he doesn't know the value of that attribute at all.

After entering the query, the system searches for the possible solution in the following way: *The information entities (attribute, value) that occur in the query are initially activated with the value of importance. The activation is propagating trough the arcs to the solution nodes, by multiplying the value of the activation of the information entity node and the weight of the corresponding arc. Final activation of the solution nodes is calculated by summing all gained activations.*

4.1. "The application of "CaBaGe" system in multiple sclerosis diagnoses

Multiple sclerosis (MS) is a chronic disease of unknown cause that affects the nervous system (brain, spinal cord and peripheral nerves) in form of multifocal lesions of myelin nerve sheet. It causes damage of vision, muscle strength, sensation, coordination, speech, and bladder control and may affect cognitive functions. MS usually appears in young adulthood (more in females) and Vojvodina is region with high incidence of disease.

Table 1. Structure of the case

Data gained from conversation with the patient	Data gained from medical checkup	
Name	Ambliopy	Hemihypoaesthesia
Family name	Loss of visual acuity	Hyperaesthesia
Age	Diplopia	Dysarthria
Residence	Nystagmus	Nasality of speech
Nationality	Central facial palsy	Scanning speech
Education	Peripheral facial palsy	Euphoria
Family history	Trigeminal neuralgia	MRI evidence of demyelination
Weakness in 1 or more extremities	Paresis n. XII	9 T2 hyperintense lesions
Cramps in extremities	Lhermit sign	1 or more infratentorial
Paroxysmal pain	Spasticity	1 or more juxtacortical lesions
Chronic pain	Monoparesis	3 or more periventricular
Facial pain	Hemiparesis	VEP latences
Dizziness	Triparesis	CSF cell count
Gait difficulties	Quadriparesis	Total CSF protein
Visual loss for one or both eyes	Subclonus	IgG
Ocular pain	Clonus	IgG/albumine
Double vision	Babinski	Link index
Increased frequency of voiding	Plantar response absent	De novo synthesis IgG in CNS
Hesitancy of voiding	Abdominal reflexes absent	MC index of synthesis Ig
Urge incontinence	Dysmetria	Oligoclonal bands
Bowel disfunction	Intention tremor	Blood barrier dysfunction
Impotence	Ataxia	Diagnosis
Fatigue	Romberg test	
Depression		
Difficulty in sustaining attention		
Memory disturbances		
Number of relapses		

After completing the "CaBaGe" system, next problem was extracting the data from previous diagnoses of MS. This problem is very complex because the data must be retrieved from the medical histories of the patients from the last 10 years. By analysing the diagnoses problem of MS disease the appropriate case structure with characteristic features is extracted (see. Table 1). The case has 72 different features representing the most important observations in the diagnoses process of MS disease. The last row in the table represents the correct diagnosis for the corresponding cases.

At this moment the database is modelled and colleagues from the Institute of Neurology are extracting and entering the data. The system is tested with some randomly generated data, but the structure of the data is final. In the rest of this chapter the functionality of the system will be shown, using the test data.

The structure of a case is described in case pattern file “MSDiagnoses.key”. This structure corresponds to the case structure in the Table 1. Just the first five features are of the string type while all of the rest features are of "int", "bool" or "float" type. Some of these features represent the Boolean type (string type with just two values: TRUE/FALSE). The test cases are given in the case base file "MSDiagnoses.cbr". The values of the features must be given in the correct order according to the case pattern file. Last value in the description of the case is always the solution of the case (Diagnoses). These values are determined dynamically during the parsing of the case base file and can be: *Definitive MS, Probable MS, Possible MS* or *No MS*.

In the Figure 2, the main window of the system is shown. At the beginning the system expects that user enters the paths of the case pattern and case base file. If the paths are good, these two files are loaded and the appropriate case retrieval net is created. Also, the dynamically created form will appear in the middle of the window.

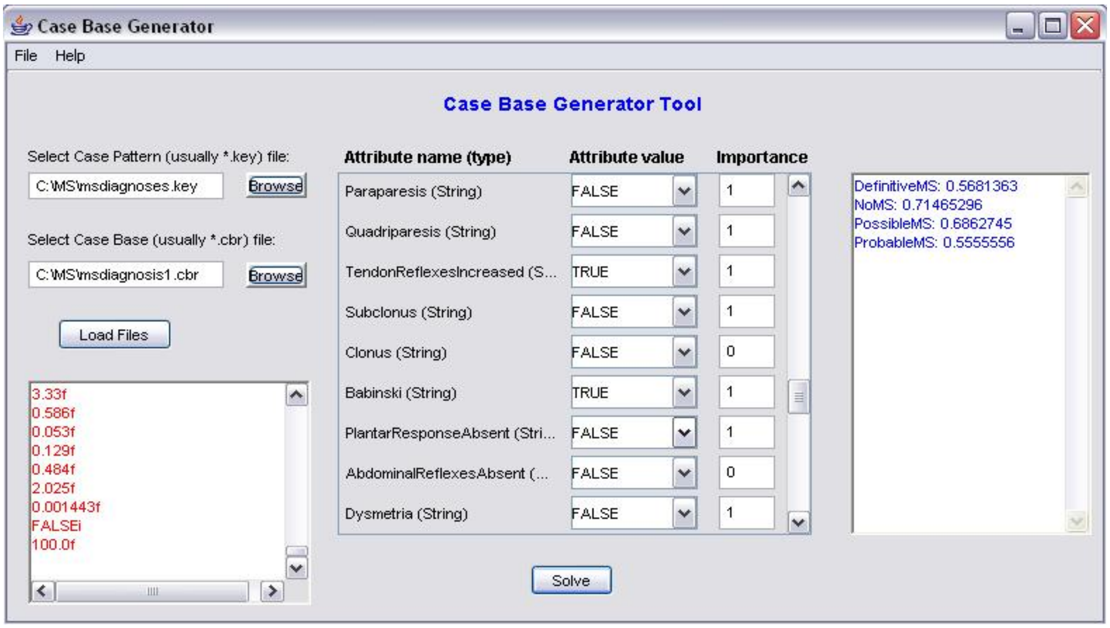

Figure 2. The main window of the "CaBaGe"

After loading files and creating CRN, the system expects, the user to enter current problem i.e. the known values of the attributes. If some values are not known then the user should enter the zero value in the corresponding importance field. When the entering the data into the form is finished, the process of the spreading activation is started. After finishing the process of retrieval and finding similar cases in the case base, the all possible solutions and their activations, has been shown to the physician. The solution with the highest number, i.e. similarity, is the “suggested solution”. In presented example (Fig. 2) it is most possible, that for this patient, the diagnosis is "No Multiple Sclerosis".

5. CONCLUSION AND RELATED WORK

Case-based reasoning technique seems to be suitable for medical knowledge-based systems [1,3,15]. Main intention of proposed decision support system is to help medical expert in making hard decisions. Also, the important advantage of this system is to reduce the gap between experienced experts and beginner physicians because this system manipulate with all previous knowledge in order to make a final decision.

Some further work might be in the direction of creating a rule based system for the diagnoses of MS disease which will simulate a textbook knowledge of physicians. Then the case based system together with the rule based system will completely simulate the physician's process of decision making. Also the system will help the experts in comparing several criteria for diagnoses of MS disease: Schumacher Criteria, Poser Criteria, McDonald Criteria… This comparison might help physicians in reducing the time for diagnoses or the number of medical checkups needed for reliable diagnoses.

References

[1] Aamodt, A., Plaza, E.: Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches, AI Commutations, pp. 39-58. 1994.

[2] Budimac Z., Kurbalija V., Case-Based Reasoning – A Short Overview, Proceedings of Second International conference on Informatics and Information Technology “Molika 2001”, Molika, Macedonia, December 20-23, 2001, pp. 222-234.

[3] Case-Based Reasoning at AIAI, http://www.aiai.ed.ac.uk/project/cbr/

[4] Iglezakis, I. (2001), “The Conflict Graph for Maintaining Case-Based Reasoning Systems”, 4th International Conference on Case-Based Reasoning (ICCBR 2001), pp. 263-276, Vancouver, Canada, July/August 2001.

[5] Ivanović M., Kurbalija V., Budimac Z., Semnic M., Role of Case-Based Reasoning in Neurology Decision Support, Proceedings of the Fifth Joint Conference on Knowledge-Based Software Engineering "JCKBSE 2002", Maribor, Slovenia, September 11-13, 2002, pp. 255-264.

[6] Kurbalija, V.: On Similarity of Curves – project report, Humboldt University, AI Lab, Berlin, 2003.

[7] Kurbalija V., Ivanović M., Case-Based Reasoning – A Powerfull Artificial Intelligence Approach, Journal of Electrical Engineering 12/s, Slovak Centre of IEE, Vol. 53, 2002, pp.20-24 (Proceedings of Conference of Applied Mathematics for undergraduate and graduate students “SCAM 2002”, Bratislava, Slovak Republic, April 19-20, 2002).

[8] Kurbalija V., Ivanović M., CaseBaseGenerator for Intelligent Systems, Novi Sad Journal of Mathematics – NSJOM, vol. 35, no. 1, 2005, pp. 25-40.

[9] Kurbalija V., Ivanović M., Budimac Z., CBR – Case-Based Reasoning, infoM 5/2003 - Journal of Information Technology and Multimedia systems, pp. 14-20

[10] Kurbalija V., Implementation of Decision Support Systems using Case Based Reasoning Technology, Master Thesis, Novi Sad, 2006.

[11] Lenz M.: Case Retrieval Nets Applied to Large Case Bases, draft, Lenz’s home-page.

[12] Lenz M., Bartsh-Spoerl B., Burkhard HD., Wess S.: Case-Based Reasoning Technology: From Foundations to Aplications, Springer Verlag, October 1998.

[13] Lenz M., Burkhard H.D., Bruckner S.: Applying Case Retrieval Nets to Diagnostic Tasks in Technical Domains, Lenz’s home-page.

[14] Reinartz, T., Iglezakis, I., Roth-Bergofer, T., (2000), “On Quality Measures for Case Base Maintenance”, 5th European Workshop (EWCBR 2000), pp. 247-260, Trento, Italy, September 2000.

[15] R. Schmidt, L. Gierl, Case-based Reasoning for Medical Knowledge-based Systems, Internet source.