ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology
Volume 1, Issue 4, June 2012

AN ENHANCEMENT IN RELEVANCE KNOWLEDGE DISCOVERY MODEL FOR MEDICAL REASONING USING CBRM

Prof.S.Anithaa
anithabennett@yahoo.co.in
VIT university Chennai campus

Abstract: *Healthcare technology is entering in to a new evolutionary phase. The medical community has an obligation to the public to provide the safest, most effective possible healthcare. In this context Knowledge sharing is crucial for better patient care in the healthcare industry, but it is challenging for physicians to exchange their clinical insights, outcomes and valuable experiences, particularly with regard to make decisions in diagnosing disease at early stages. The aim of our study is to facilitate knowledge sharing and information exchange in this area by means of a knowledge-based system. We propose a knowledge-based system, which automatically models each physician's experience and experts in health care industry. This is done by collecting as many as possible instances occurred during past histories about the patient symptoms, treatments, and relevant medications. This process can be referred as Case Based Reasoning Model(CBRM) and will be analyzed from a statistical perspective to form a authenticated interactive knowledge sharing process for making right decisions at right time.*

Key Application Areas includes : Improving the Quality of Patient Care *Identifying high-risk patient groups with combinations of symptoms and/or risks, Identifying the need for prophylactic measures to prevent outbreak of disease, Improve patient care through efficient prescribing of drugs by identifying duplication or over-prescribing of drugs, and also identifying potential drug interactions in contraindicated drugs, Search for statistical data regarding patient-disease patterns, classifying them based on age, gender, geographical locations, food groups, etc., by identifying common factors among patients with similar diseases. Identifying the need for diagnostic tests in specific patients, leading to effective dispensing of health care measures.*

Revenue Generation and Saving Time *Lowering the cost and effort involved in clinical Research and Development through automated reviews. Identifying the need for specific diagnostic tests in specific patients, leading to effective dispensing of health care measures and eliminating unnecessary tests which leads to saving valuable time and cost considerably.*

Keywords: *Case Based Reasoning Model, Health care, Knowledge based systems, Knowledge sharing.*

I. INTRODUCTION

Medical diagnosis and decision-making involves interplay between vast numbers of medical knowledge resources . This can range from implicit knowledge held by healthcare workers to experiential and data-induced knowledge. Systems that can simultaneously access and combine relevant information from these various knowledge resources are crucial to the diagnostic and prognostic processes and subsequently to the efficient treatment of patients. From a decision support viewpoint, healthcare workers need complete, contextually-relevant information that is consistent with the patient's current medical state and that is appropriately presented at the correct level of abstraction. In this research we have drawn on medical knowledge management initiatives that promote the collection, integration and distribution of a single medical modality [1]. This allows us to build encapsulated patient profiles that are used both to effectively store patient data and also for the purposes of comparison with new patient profiles for diagnosis and treatment. The system we are developing also has beneficial repercussions from a healthcare modeling view point [2], as explanatory models from amassed patient data can easily be created that can identify trends as well as comparing diagnosis, treatments and departments.

II. RESEARCH ISSUES AND PROPOSED CONTRIBUTION

Clinical data mining has three objectives: understanding the clinical data, assist healthcare professionals, and develop a data analysis methodology suitable for medical diagnosis[3]. Our work deals with the insight of health care professionals such as radiographer or physician decisions, as they diagnose and treat patients during illness. Since the health care professionals interacts with the patient profile (patient age, gender, medical history, examination and laboratory findings, medical imagery like X-rays etc are integrated as encapsulated profiles) in the course of a diagnosis the system analyses their actions. This enables to capture human expertise and proficiency in diagnosing and treating any particular disease. This proficiency of diagnosing disease and relevant treatment is referred as an expert knowledge and is stored in knowledge base for future reference. This information can then be used to filter, retrieve and display the most relevant similar patient case histories from huge repositories of patient data and can further be used for diagnostic comparison with new patient complaints, symptoms or other clinical evidences. The information of new patients can be entered in to the system and previous diagnoses, treatments and outcomes for similar patients can be instantly accessed by physicians. The application can save much physician time by avoiding manual scanning of patient records and retrieved information can be interactively explored and can used to guide health care professionals towards appropriate and relevant information regarding diagnoses and treatments.

Effective use of this knowledge base of previous case histories is made possible by the application of Case-Based Reasoning (CBR) techniques. CBR is a well established method for building medical systems [4], and one of the intuitively attractive features of CBR in medicine is that the concepts of patient and disease lend themselves naturally to a case representation. Also medical practitioners logically approach diagnosis from a case-based standpoint (i.e., previous specific patient interactions are as strong a factor as individual symptoms in making a diagnosis). There are three main advantages in our approach. First by reusing collective knowledge in support of similar patient cases the time required to diagnose or treat a new patient can be significantly reduced. Second, the approach facilitates knowledge sharing (remote or otherwise) by retrieving potentially relevant knowledge from other experiences. Finally from a knowledge management perspective, contextual expert knowledge relating to particular cases may now be stored and reused as an additional resource for support, training and preserving knowledge assets.

III. CASE BASED REASONING

Case Based Reasoning (CBR) is a recognized and well-established method for the health science. The health science domain offers the CBR community worthy challenges and is driving CBR research forward by offering a variety of complex tasks, which are difficult to solve with other methods and approaches. [5] The origin of CBR can be traced to Yale University and the work of Schank and Abelson in 1977 [6]. Earlywork exploiting CBR in the medical domain was performed by Koton [7] and Bareiss [8] in the late 1980s. The CBR is inspired by human reasoning, i.e., solving a new problem by applying previous experiences adapted to the current situation. A case (an episodic experience) normally contains a problem, a solution, and its result. The CBR is an appropriate method to explore in a medical context where symptoms represent the problem, and diagnosis and treatment represent the solution. With its capability of incrementally collecting, reusing and sharing the knowledge implicitly embedded in previously experienced situations, case-based reasoning (CBR) [9] is currently recognized as a very well suited reasoning methodology in medical applications.

IV. CBRM CASE BASED REASONING MODEL

Life cycle of CBRM is a combinational model of statistical view and CBR view. This combinational model gives knowledge for health care professionals to make decisions as shown in figure 1. The CBR view has four main steps (4R's) retrieve, reuse, revise, and retain. In the retrieval step, a new problem is matched against the previous cases in the case library. Domain knowledge is used to determine the similarity between the new case and the case available in domain knowledge, and the degree of similarity leads to make decision about the newly arrived case. The most relevant solutions available in the knowledge base are proposed to solve the newly arrived problem with respect to some adaptations if necessary. The selected solution is revised before it is reused. Then, the new problem and its solution are retained in the case library for future use.

CBR systems typically needs to undergo preprocessing and filtering prior to Case formulation. For example, if the clinical data are collected from sensor signals, images, free-text sources, etc., then the system may require feature extraction, feature mining, indexing, weighting, etc

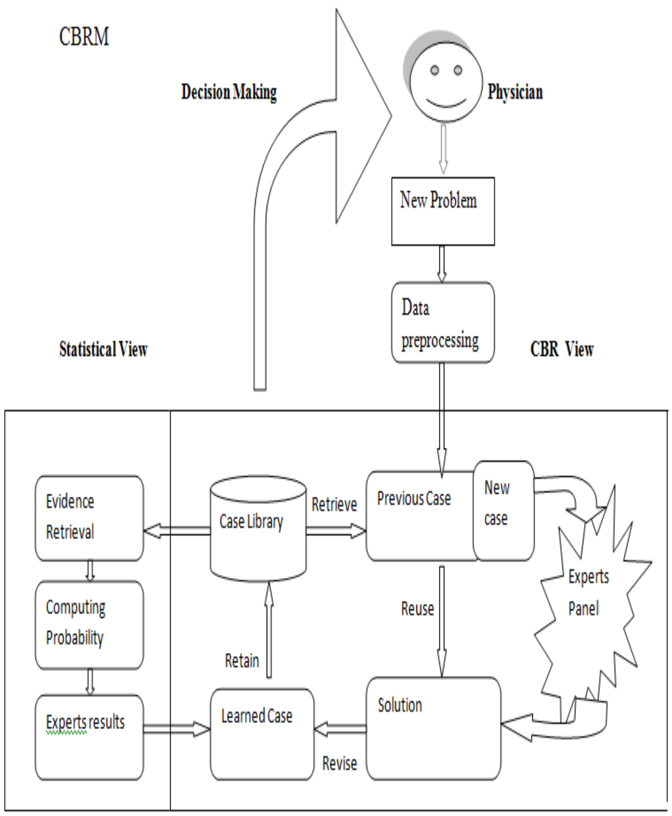

Figure 1:CBRM (Case Based Reasoning Model)

In the medical domain, the health care professionals, clinicians or doctors may start their practice with some initial experiences (solved cases). Over a period of time, they use that initial experiences as past experiences to solve a new problem. This may involve some adjustment of the previous solutions to solve the new problem. Thus, a new experience (case) has been created, which enriches the clinician's / doctor's set of experiences. In fact, this is how the traditional CBR cycle works. So, the CBR is a reasoning process, which is medically accepted and also getting increasing attention from the medical domain.

V. RELEVANCE COMPUTATION

The CBRM accepts new cases from health care professionals and preprocess them in order to remove noise, inconsistent data. The outcome of preprocessing is sent to the CBR model which attempts to **retrieve** the most similar case from case library. The system will carry out similarity computation test by weighted average using nearest neighbor algorithm. This algorithm will calculate similarity between new problem and similar cases in case library. The nearest case with highest *k* score will be used as similar case to solve new problem. The similarity relationship between cases are grouped into a set of independent attributes such as gender, age, marital status, blood pressure, depression, smoke, stress, anxiety etc. For each attributes, a relevance metric is defined to measure the similarity between two items. For example, two cases with same gender will get the maximum similarity rating while for attributes that greatly dissimilar, will get low rating. The output of each metric is an integer value. The degree of similarity is expressed between 0 (not similar) and 1 (very similar) . In this paper, we have created two sample cases for the similarity retrieval purpose in case database. Figure 5 shows the similarity comparison between New Case and Sample 1 while Figure 6 shows the similarity comparison between New Case and Sample 2.

New Case

Age	24
Sex	Male
Marital Status	N
Blood Pressure	Y
Depression	Y
Smoke	Y
Stress	Y
Anxiety	Y
Food Habit	V
Alcohol	Y

Figure 2: Summary of new problems entered by Health care professionals and labeled as 'New Case'

Sample 1

Age	32
Sex	Male
Marital Status	Y
Blood Pressure	Y
Depression	Y
Smoke	N
Stress	Y
Anxiety	N
Food Habit	NV
Alcohol	Y

Figure 3 shows the summary sample1

Sample 2

Age	45
Sex	Male
Marital Status	Y
Blood Pressure	Y
Depression	Y
Smoke	Y
Stress	Y
Anxiety	N
Food Habit	V
Alcohol	N

Figure 4 shows the summary sample2

New Case VS Sample 1

New Case		Similarity rating	Sample 1	
Age	24	0.8	Age	32
Sex	Male	1.0	Sex	Male
Marital Status	N	0.1	Marital Status	Y
Blood Pressure	Y	1.0	Blood Pressure	Y
Depression	Y	1.0	Depression	Y
Smoke	Y	0.1	Smoke	N
Stress	Y	1.0	Stress	Y
Anxiety	Y	0.1	Anxiety	N
Food Habit	V	0.1	Food Habit	NV
Alcohol	Y	1.0	Alcohol	Y

Figure 5: shows New Case VS Sample 1 with similarity measures

New Case		Similarity rating	Sample 2	
Age	24	0.7	Age	45
Sex	Male	1.0	Sex	Male
Marital Status	N	0.1	Marital Status	Y
Blood Pressure	Y	1.0	Blood Pressure	Y
Depression	Y	1.0	Depression	Y
Smoke	Y	1.0	Smoke	Y
Stress	Y	1.0	Stress	Y
Anxiety	Y	0.1	Anxiety	N
Food Habit	V	1.0	Food Habit	V
Alcohol	Y	0.1	Alcohol	N

Figure 6: New Case VS Sample2

The Followings are similarity calculation (relevance computation) by weighted average using nearest neighbor algorithm for both sample cases (weight 5 for high importance and weight 1 for low importance):

Similarity computation: New Case and Sample 1

= 1/10 * [(1 * 0.8) + (1 * 1.0) +(1*0.1)+ (5 * 1.0)+(5*1.0)+(5*0.1)+(5*1.0)+(5*0.1)+(5*0.1)+(5*1.0)]

=1/10*(0.8+1.0+0.1+5.0+5.0+0.5+5.0+0.5+0.5+5.0)

=1/10*(23.4)

= **2.34**

Similarity computation: New Case and Sample 2

= 1/10 * [(1 * 0.7) + (1 *1. 0) +(1*0.1)+ (5 * 1.0) +(5*1.0)+ (5*1.0) +(5*1.0)+ (5 * 0.1) + (5* 1.0)+(5*0.1)]

= 1/10*(0.7+1.0 +0.1+5.0+5.0+5.0+5.0+ 0.5 +5.0 + 0.5)

=1/10*(27.8)

= **2.78**

The result of similarity computation will be used to determine the adoption of similar case's solution as solution to new problem. The acceptance level is predetermined by expert committee. From the similarity calculation, the system will choose solution from Sample 2 as solution for New Case (2.78 > 2.34). The system will **reuse** previous solution to solve users' current requirement. However, if level of acceptance is rejected, the system will send current problem case to the expert committee. After received solution from committee, the system will make proposal to users.. The proposed solution will be compared and **revised** against the actual result. If the result is accepted, the system will learn new case and **retain** the case in case library for future use. If the evaluation is rejected, the system will close the case and users will be advised to make consultation with expert committee. Finally the solution was verified with the statistical methods and the outcome is updated in case library. At the heart of this model lies a knowledge discovery method that permits relevance knowledge to be automatically extracted from existing structured and unstructured data that are available in preprocessing stage. Not only this model can handle data format diversity, high dimensionality, and relative importance of the data source, but it is also capable of incrementally updating the existing indexing structure when new cases are added to the system. The techniques outlined in this paper will yield significant benefit in the improvement of decision support tasks and provide a better insight into the process of medical reasoning

VI. CHALLENGES OF CASE BASED REASONING

There are a huge number of medical applications which offers challenges for the CBR researchers and drive advances in research. Some of the important research issues are given in the following.

1)Feature extraction is becoming complicated in the recent medical CBR systems due to a complex data format like sensor data, images, time series data , and data with free text format.

2) Feature selection and weighting are two other important factors for which many CBR systems depend on expert knowledge. Cases with hidden features could also affect the retrieval performance.

3) The component that plays a central role in the CBR systems is the case data base or case library. A case base can be considered as concrete knowledge of a model consisting of specific cases. The cases stored in a case library should be both representative and comprehensive, so as to cover a wide spectrum of possible situations

VII. CONCLUSION AND FUTURE WORK

The current system in hospitals whereby doctors enter patient information using paper charts is cumbersome, time-consuming and does not facilitate knowledge sharing. Different types of information, including imagery, are stored in different locations and valuable time is often lost trying to correlate data in order to diagnose and treat patients. This system can address such issues by providing doctors with instant access to information that will allow them to make critical decisions and prognoses with greater speed and efficiency. It facilitates knowledge sharing and supports effective communication about the most effective ways in which to treat patients by linking similar patient case histories using case-based reasoning techniques.[10]. The concept of CBRM is a service oriented model which retains previously solved cases and experts' knowledge in case library. The system can propose wellness solution that is personalized to users based on the adaptation of previous solution in similar cases. The proposed CBRM uses a combination of Case Based Reasoning and Statistical Models. Furthermore, the advantage of this model over the other model is, in other CBR models the limitation was that the system will not be able to propose solution to new case that has no similar solution in Case Library. But in our Model the system will build new case and send to the expert committee for confirmation on the solution and verified with the statistical methods. Future study will be focused more on preprocessing of input data to the CBRM which results in further improvement of the system.

References

[1] Jadad, A. R., Haynes, R.B., Hunt, D. & Browman, G.P. "The Internet and Evidence-Based Decision Making: A Needed Synergy for Efficient Knowledge Management in Health Care,," Canadian Medical Association Journal, Vol. 162(3):362-5, 2000.

[2] Ivatts, S. & Millard, P.H. "Health care modelling - Why should we try?," British Journal of Health Care Management, Vol. 8(6), pp.218-222, 2002.

[3] Clinical Data Mining: a Review J. Iavindrasana et al, yearbook 2009.

[4] Nilsson, M. and Sollenborn, M., "Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development," In Proceedings of the 17th International FLAIRS Conference, pp.178-183, 2004.

[5] Shahina Begum et al. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 41, NO. 4, JULY 2011

[6] R. C. Schank and R. P. Abelson, Scripts, Plans, Goals and Understanding. Erlbaum, Hillsdale, New Jersey, 1977.

[7]P. Koton, "Using experience in learning and problem solving. Massachusetts Institute of Technology," Ph.D. dissertation, Lab. Computer Science, Dept. Electr. Eng. Comut. Sci., MIT/LCS/TR-441, Cambridge, 1989.

[8] R. Bareiss, "Examplar-based knowledge acquisition: A unified approach to concept, classification and learning,"Academic Press Professional, Inc., San Diego, CA, 1989.

[9] Kolodner JL. Case-based reasoning. San Mateo, CA: Morgan Kaufmann; 1993.

[10] David Wilson,et al, 3rd International IEEE Conference Intelligent Systems, September 2006

Author:

Prof.S Anithaa
VIT University
Chennai campus
To my credit I have published more than 12 papers in National and International journals and conferences.
My area of interest are Data mining, Database management systems, Knowledge Engineering