Case Based Reasoning Technology for Medical Diagnosis

Abdel-Badeeh M. Salem

***Abstract*—**Case based reasoning (CBR) methodology presents a foundation for a new technology of building intelligent computer-aided diagnoses systems. This Technology directly addresses the problems found in the traditional Artificial Intelligence (AI) techniques, e.g. the problems of knowledge acquisition, remembering, robust and maintenance. This paper discusses the CBR methodology, the research issues and technical aspects of implementing intelligent medical diagnoses systems. Successful applications in cancer and heart diseases developed by Medical Informatics Research Group at Ain Shams University are also discussed.

***Keywords*—**Medical Informatics, Computer-Aided Medical Diagnoses, AI in Medicine, Case-Based Reasoning.

I. Introduction

CBR receives increasing attention within the AI community [1]. CBR is an analogical reasoning method providing both a methodology for problem solving and a cognitive model of people. CBR means reasoning from experiences or "old cases" in an effort to solve problems, critique solutions, and explaining anomalous situations [2]. It is consistent with much that psychologist have observed in the natural problem solving that people do. People tend to be comfortable using CBR methodology for decision making, in dynamically changing situations and other situations were much is unknown and when solutions are not clear.

There are two styles of CBR; problem solving style and interpretive style. Problem solving style can support a variety of tasks including planning, diagnosis and design (e.g. Medicine [3] and Industry [4]). The interpretive style is useful for (a) situation classification, (b) evaluation of solution, (c) argumentation, (d) justification of solution interpretation or plan and (e) the projection of effects of a decision of plan. Lawyers and managers making strategic decisions use the interpretive style [5, 6]. CBR has already been applied in a number of different applications in medicine. Some real CBR-systems are: CASEY that gives a diagnosis for the heart disorders [2], GS.52 which is a diagnostic support system for dysmorphic syndromes, NIMON is a renal function monitoring system, COSYL that gives a consultation for a liver transplanted patient [7] and ICONS that presents suitable calculated antibiotics therapy advised for intensive care patients [8].

Abdel-Badeeh M. Salem is a professor with the Department of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt (phone: +202-0122182645, e-mail: absalem@asunet.shams.edu.eg).

In this paper we focus our discussion around the usage of CBR technology for medical diagnosis. Section 2 presents the main concepts and techniques of CBR methodology. Sections 3 discuss the knowledge engineering tasks in developing CBR-Based systems. Sections 4 and 5 present the CBR-based systems for diagnosis of cancer and heart diseases. The final section contains the summary and conclusion.

II. Case-Based Reasoning Methodology

CBR refers to a number of concepts and techniques that can be used to record and index cases and then search them to identify the ones that might be useful in solving new cases when they are presented. In addition, there are techniques that can be used to modify earlier cases to better match new cases and other techniques to synthesize new cases when they are needed [2].

Fig. 1 shows the CBR methodology adapted from [9]. Following this diagram, the algorithm of interpreting and assimilating a new case can be summarized in the following processes:

1. Assign Indexes: where the features of the new case are assigned as indexes characterizing the event.
2. Retrieve: where the indexes are used to retrieve a similar past case from the case memory. (The past case contains the prior solution).
3. Modify: where the old solution is modified to conform to the new situation, resulting in a proposed solution.
4. Test: where the proposed solution is tried out. It either succeeds or fails.
5. Assign and Store: If the solution *succeeds*, then assign indexes and stores a working solution. The successful plan is then incorporated into the case memory.
6. Explain, Repair and Test: If the solution fails, then explain the failure, repair the working solution, and test again. The explanation process identifies the source of the problem. The predictive features of the problem are incorporated into the indexing rules knowledge structure to anticipate this problem in the future. The failed plan is repaired to fix the problem, and the revised solution is then tested.

To perform the CBR processes, the following knowledge structures (KSs) are very essential:

1. Indexing Rules KS: These rules identify the predictive features in the input that provides appropriate indexes into the case memory.
2. Case Memory KS: Case memory is the episodic memory, which comprises of database of experience.

3. Similarity Rules KS: If more than one case is retrieved form episodic memory, the similarity rules (or metrics) can be used to decide which case is more similar to the current situation.
4. Modification Rules KS: If no old case is going to be an exact match for a new situation, the old case must be modified to fit. We require knowledge about what kinds of factors can be changed and how to change them.
5. Repair Rules KS: Once we identify and explain an expectation failure, we must try to alter our plan to fit the new situation. Again we have rules for what kinds of changes are permissible.

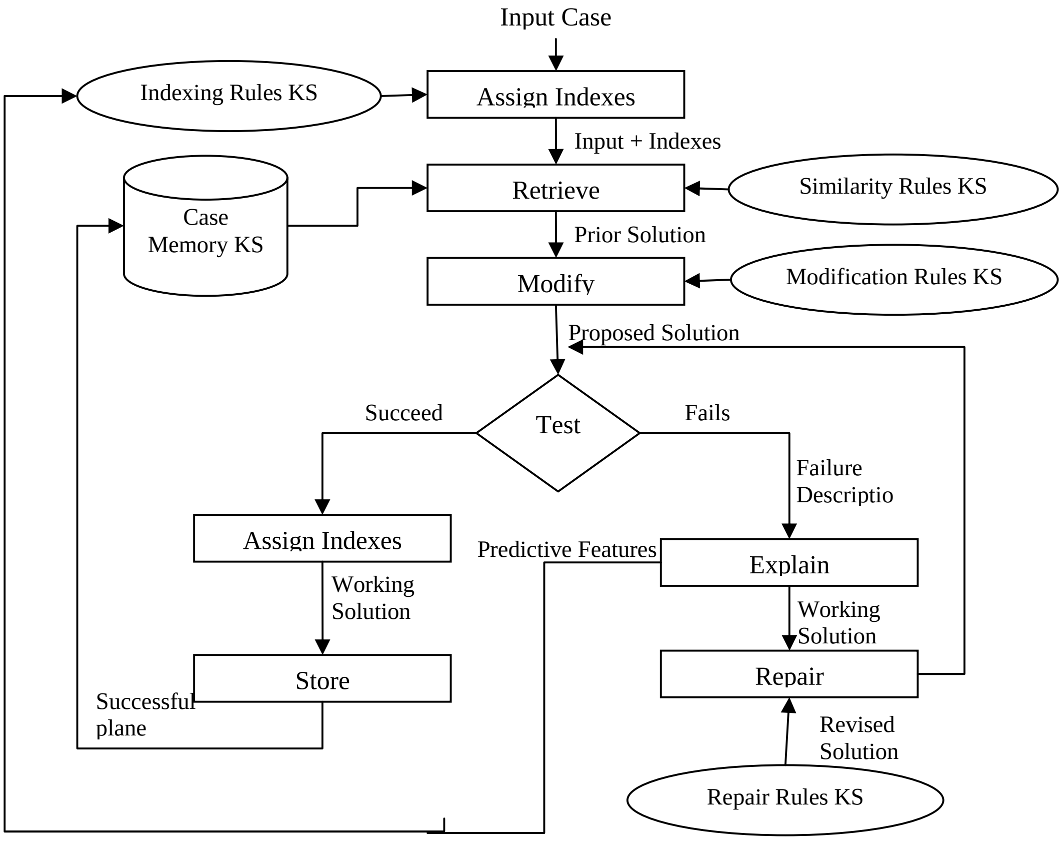

Fig. 1 CBR Methodology. Boxes represent processes and ovals represent KS

III. Knowledge Engineering Tasks in Developing CBR-Based Systems

Accordingly the methodology of developing CBR-based systems in specific domain can be summarized in the following steps:

1. The system will search its Case-Memory for an existing case that matches the input problem specification.
2. If we are lucky (our luck increases as we add new cases to the system), we will find a case that exactly matches the input problem and goes directly to a solution.
3. If we are not lucky, we will retrieve a case that is similar to our input situation but not entirely appropriate to provide as a completed solution.
4. The system must find and modify small portions of the retrieved case that do not meet the input specification. This process is called "case-adaptation".

5. The result of case adaptation process is (a) completed solution, and (b) generates a new case that can be automatically added to the system's case-memory for future use.

In what follows a brief discussion of the knowledge engineering tasks and issues which are crucial in developing the CBR-based systems, namely: (1) case representation, (2) case indexing, (3) case storage and retrieval, (4) case adaptation (5) learning and generalization and (6) the selection of the CBR tools/shells which suitable for the appropriate task is presented.

1. Case Representation: Determining the appropriate case features is the main knowledge engineering task in CBR systems. The case is a list of features that lead to a particular outcome (e.g. the information on a personal-loan form and

whether or not the loan was granted; the information on a patient history and the associated diagnosis). The task involves; (a) defining the terminology of the domain and (b) gathering representative examples of problem solving by the expert. Representations of cases can be in any of several forms; predicate representations, frame representations and representations resembling database entries. Depending on what is included in a case, the case can be used for a variety of purposes: (1) Cases that include a problem-description AND solution can be used in deriving solutions to new problems, (2) Cases that include a problem-description AND outcome can be used in evaluating new situations and (3) Cases that have a specified solution can be used in evaluating proposed solutions and anticipating potential problems before they occur.

2. Case Indexing Process: The CBR system derives its power from its ability to retrieve relevant cases quickly and accurately from its memory. Figuring out when a case should be selected for retrieval in similar future situations is the goal of the case *indexing process*. Building a structure or process that will return the most appropriate case (from the case memory) is the goal of the *retrieval process*. Case indexing process usually falls into one of three approaches: nearest neighbor, inductive, and knowledge-guided or a combination of the three [2].

3. Case Memory Organization and Retrieval: Once cases are represented and indexed, they can be organized into an efficient structure for retrieval. Most case memory structures fall into a range between *purely associative retrieval*, where any or all of the features of a case are indexed independently of the other features and *purely hierarchical retrieval*, where case features are highly organized into a general-to-specific a concept structure. Nearest-neighbor matching techniques are considered associative because they have no real-memory organization. Discrimination nets are more of a cross between associative and hierarchical because they have some structure to the net but greater retrieval flexibility because they have a greater number of links between potential indexing features. Decision trees are an example of purely hierarchical memory organization. The type of memory organization is related to the amount of knowledge available to perform indexing and the retrieval needs of the system. If flexibility is required because one case library is being used for several retrieval tasks, a more associative approach is often used. When the retrieval task is well defined, a hierarchical approach is used because of the advantages in retrieval time the hierarchical approaches have over associative approaches.

4. Case Adaptation: It is difficult to define a single generically applicable approach to perform case adaptation, because adaptation tends to be problem specific [2, 10 and 11]. Most existing CBR systems achieve case adaptation for the specific problem domains they address by encoding adaptation knowledge in the form of a set of adaptation rules or domain model. Adaptation rules are then applied to a retrieved case to transform it into a new case that meets all of the input problem’s constraints. More recent applications have successfully used pieces of existing cases in memory to perform adaptations. In problem domains where it is difficult to codify enough rule-like knowledge to let board adaptation be done, using pieces of cases is the best, if not the only alternative. And, even if cases can’t be adapted by the computer, at least the system has provided the human “adapter” with a significant starting point.

5. Learning and Generalization: as cases accumulates, case generalization can be used to define prototypical cases that embody the major features of a group of specific cases, and those prototypical cases can be stored with the specific cases, improving the accuracy of the system in the long run. In addition, inductive-case analysis research is being done to build domain theories in areas where even the experts don’t understand how the underlying processes in their domain work.

6. CBR - Tools and Shells: The availability of a commercial CBR shells in the market helps the knowledge engineers to overcome some of the problems they currently face in designing and maintaining large knowledge-base learning systems using rule based tools. The most comprehensive evaluation of commercial CBR tools from Europe and the USA has been published by Watson [12]. Most of the tools can be integrated with other model-based tools. This property suggests that including CBR as an element within our applications is now great. The most of the currently available CBR shells are: CBR-Express, CasePoint, ART*Enterprise, CasePower, Esteem, Expert Advisor, ReMind, CBR/text, Eclips, ReCall, RATE-CBR, S3-Case, INRECA, and CASUEL.

In conclusion, the technology of CBR directly addresses the following problems found in rule-based technology.

1. Knowledge acquisition: The unit of knowledge is the case, not the rule. It is easier to articulate, examine, and evaluate cases than rules.
2. Performance: A CBR system can remember its own performance, and can modify its behavior to avoid repeating prior mistakes.
3. Adaptive Solutions: By reasoning from analogy with past cases, a CBR system should be able to construct solutions to novel problems.
4. Maintaining: Maintaining CBR system is easier than rule-based system since adding new knowledge can be as simple as adding a new case.

IV. CBR-Based System for Diagnosis of Cancer Diseases

Cancer is a group of more than 200 different diseases; it occurs when cells become abnormal and keep dividing and forming either benign or malignant tumors. Cancer has initial signs or symptoms if any is observed, the patient should perform complete blood count and other clinical examinations. Then to specify cancer type, patient needs to perform special lab-tests.

This section presents a summary of the CBR-based expert system prototype for diagnosis of cancer diseases developed by Medical Informatics Group at Ain Shams [13]. The main purpose of the system is to serve as doctor diagnostic assistant. The system provides recommendation for controlling pain and providing symptom relief in advanced cancer. It can be used as a tool to aid and hopefully improve the quality of care given for those suffering intractable pain. The system is very useful in the management of the problem, and its task is to aid the young physicians to check their diagnosis. Fig. 2 shows the architecture of the CBR-based system. The system's knowledge base is diverse and linked through a number of indices, frames and relationships. The bulk of this knowledge consists of actual case histories and includes 70 cancer patient cases; some are real Egyptian cases and some from virtual hospitals on the internet.

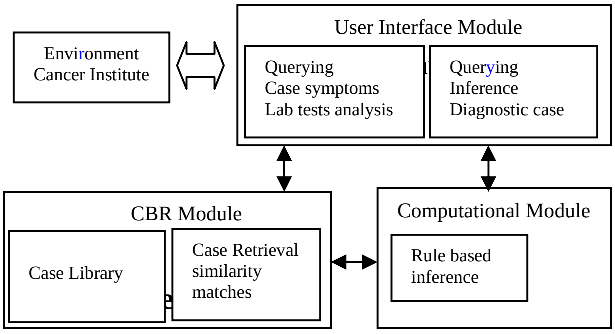

Fig. 2 Architecture of the CBR-based system for cancer diagnosis

The system consists of three main modules; user interface, case base reasoning module and computational module all are interacted with the main environment of cancer diseases. The user is cancer expert doctor; the interaction is through menus and dialogues that simulate the patient text sheet containing symptoms and lab examinations. Computational model uses rule-based inference to give diagnostic decision and new case is stored in case library. Patient cases are retrieved in dialogue with similarity matches using the nearest neighbor matching technique.

Frames technique is used [10] for patient case indexing, storage and retrieval. The patient case will include age, sex and weight occupation, pathologic, medical history family, physical exams and treatments. Example of an Egyptian liver cancer case description of old woman was given as shown in Fig. 3.

Patient: 65-years old female not working, with nausea and vomiting.
Medical History: cancer head of pancreas
Physical Exam: tender hepatomgaly liver, large amount of inflammatory about 3 liters, multiple liver pyogenic abscesses and large pancreatic head mass.
Laboratory Findings: total bilrubin 1.3 mg/dl, direct bilrubin 0.4 mg/dl, sgot (ast) 28 IU/L, sgpt (alt) 26 IU/L.

Fig. 3 Egyptian liver cancer case

The initial diagnostic process is done through firing of rules in the Rule-Based inference. These rules encode information about patient's symptoms and pathological examinations. Fig. 4 shows a real sequence of diagnosis for Hodgkin's cancered patient.

Initial situation: one day, the patient felt a tender lump in her neck
Several days after, the lump was still there
The patient called the doctor; he looked at the skin area in her neck
Doctor and patient dialogue: Have you had fever? No
Have you had night sweats? No
Have you had general itching? No
Have you lost weight? No
How about fatigue? Yes
Decision 1: the patient made blood tests and chest x-rays
Result: a tab finding reveals that patient chest x-ray shows an enlarged mass in chest
Decision 2: a Biopsy for the node in the neck
Final Result: The patient is developing a Malignant cancer of the body lump system a Cancer called HodgKin's disease.

Fig. 4 Real sequence of patient diagnosis

V. CBR-Based System for Diagnosis of Heart Diseases

Heart disease is a vital health care problem affecting millions of people. Heart disease is of 25 different ones: e.g. left-sided heart failure, right-sided heart failure, angina pectoris, myocardial infraction and essential hypertension. The system is able to give an appropriate diagnosis for the presented symptoms, signs and investigations done to a cardiac patient with the corresponding certainty factor. It can be used to serve as doctor diagnostic assistant and support the education for the undergraduate and postgraduate young physicians [14].

In this system the knowledge is represented in the form of frames and the case memory contains 110 cases for 4 heart diseases namely; mistral stenosis, left-sided heart failure, stable angina pectoris and essential hypertension. Each case contains 207 attributes concerning both demographic and clinical data. After removing the duplicate cases, the system has trained set of 42 cases for Egyptian cardiac patients. Statistical analysis has been done to determine the importance values of the case features. Two retrieval strategies were investigated namely; induction and nearest neighbor approaches. The results indicate that the nearest neighbor is better than the induction strategy. Cardiologists have evaluated the overall system performance where the system was able to give a correct diagnosis for thirteen new cases.

VI. Summary and Conclusion

CBR is appropriate methodology for all medical domains and tasks for the following reasons: cognitive adequateness, explicit experience, duality of objective and subjective knowledge, automatic acquisition of subjective knowledge, and system integration. CBR presents an essential technology of building intelligent CBR systems for medical diagnoses that can aid significantly in improving the decision making of the physicians. These systems help physicians and doctors to check, analyze and repair their solutions. The physicians inputs a description of the domain situation and his (her) solution and the system can recalls cases with similar solutions and presents their outcomes to the physician. Also attempts to analyze the outcomes to provide an accounting of why the proposed type of solution succeeded or failed.

The paper presents briefly two CBR-based systems for diagnosis of cancer and heart diseases developed by Medical Informatics Research Group at Computer Science Department of Ain Shams University, Cairo, Egypt.

On the other hand, the convergence of CBR systems, grid computing, and web technologies is enabling the creation and implementation of the intelligent internet-based training technology for the young physicians. Such technology will provide a unique opportunity to distribute training across multiple sites as well as deliver healthcare services reducing both costs and time.

REFERENCES

[1] Greer, J. Proceedings of AI-ED 95, World Conference in Artificial Intelligence in Education, Association for Advancement of Computing in Education (AACE), Washington, DC, (1995).

[2] Kolodner, J. Case-Based Reasoning, Morgan Kaufmann, San Mateo, (1993).

[3] Silvana, Q., Pedro, B., and Steen, A. Proceedings of 8th Conference on Artificial Intelligence in Medicine in Europe, AIME, Cascais, Portogal, Springer, (2001).

[4] Hinkle, D. and Toomey, C., Applying Case-Based Reasoning to Manufacturing, AI Magazine, pp. 65-73, (1995).

[5] Rissland, E.L. and Danials, J.J., A Hybrid CBR-IR Approach to Legal Information Retrival, Proceedings of the Fifth International Conference on Artificial Intelligence and Law, (ICAIL-95), pp. 52-61, College Park, MD, (1995).

[6] Salem, A.M. and Baeshen, N., Artificial Intelligence Methodologies for Developing Decision Aiding Systems, Proceedings of Decision Sciences Institute, 5th International Conference, Integrating Technology and Human Decisions: Global Bridges into the 21st Century (D.I.S. 99 Athens), Greece, pp.168-170, (1999).

[7] M. Lenz, S Wess, H Burkhard and B Bartsch, Case based reasoning technology: from foundations to applications, Springer 1998.

[8] B. Heindl. Et al,: A Case-Based Consiliarius for Therapy Recommendation (ICONS) computer-based advise forv calculated antibiotic therapy in intensive care medicine, computer methods and programs in biomedicine 52, pp 117-127, 1997.

[9] Salde, S. Case-Based Reasoning: A Research Paradigm, AI Magazine, Vol. 12, No. 1, 42-55, (1991).

[10] Voss, A. Towards a Methodology for Case Adaptation, Proceedings of the 12th European Conference on Artificial Intelligence, Budapest, Hungary, pp. 147–157, (1996).

[11] Abdel-Badeeh M. Salem , Bassant M. El Bagoury, A Case-Based Adaptation Model for Thyroid Cancer Diagnosis Using Neural Networks, Proceedings of the sixteenth international FLAIRS Conference, AAAI Press, pp.155-159, (2003).

[12] Ian W., Applying Case-Based Reasoning: Techniques for Enterprise Systems, Morgan Kaufmann, California, (1997).

[13] Salem A.B.M, Roushdy M., and El-Bagoury, B.M., An Expert System for Diagnosis of Cancer Diseases, Proceedings of the 7th International Conference on Soft Computing, MENDEL, pp. 300-305, (2001).

[14] Abdel-Badeeh M. Salem and Rania A. HodHod, A Hybrid Expert System Supporting Diagnosis of Heart Diseases, Proceedings of IFIP 17th World Computer Congress, TC12 Stream on Intelligent Information Processing, Kluwer Academic Publishers, Montreal, Quebec, Canada, pp. 301-305, (2002).