Integrating Case-Based Reasoning, Rule-Based Reasoning and Intelligent Information Retrieval for Medical Problem-Solving

Isabelle Bichindaritz, Emin Kansu and Keith M. Sullivan

Clinical Research Division
Fred Hutchinson Cancer Research Center
Seattle, Washington 98109
ibichind@fhcrc.org

Abstract

This paper proposes a multimodal reasoning framework for the cooperation of case-based reasoning, rule-based reasoning and information retrieval to solve problems. Functionally, it offers via the WWW knowledge-support assistance to clinicians responsible for the long-term follow-up of stem-cell post-transplant patient care . In this domain, no single knowledge source can address the range of problems met, and the cooperation of different knowledge sources, such as cases, clinical pathways and practice guidelines (expressed as rules and/or documents), provides better coverage of both the variety and complexity of the different problems encountered. In particular, case-based reasoning enhances the system by conferring an ability to learn from experience, and thus improve results over time.

Introduction

The system presented here is a *knowledge-based computerized decision-support system* on the World-Wide Web (WWW). It permits to assist its users to perform a set of clinical tasks using a general framework for reasoning from knowledge sources of varied quality, which means that their knowledge is based upon varied evidence.

This computerized decision-support system is applied to the *long-term follow-up* (LTFU) of patients having undergone a *bone-marrow* or *stem-cell transplant* (BMT) at the Fred Hutchinson Cancer Research Center (FHCRC) in Seattle, after their return to their home community. This paper focuses on the integration of *case-based reasoning* (CBR), *rule-based reasoning* (RBR) and *information retrieval* (IR) in this system, from a reasoning viewpoint.

This reasoning framework performs its inferences based upon the different knowledge elements contained in the system's knowledge-base, and which are represented by documents, rules and cases. In this domain, no single knowledge representation format is sufficient to represent knowledge. Nevertheless, in order for this system to reason from all of these sources, a common general representation language and common reasoning steps are proposed.

After introducing the functionality of the system, from a user viewpoint, this paper presents in sequence case-based reasoning, rule-based reasoning, information retrieval, and then the system multimodal reasoning as it is performed. It also discusses this system's approach to CBR integration and compares it with other systems.

System Functionality

This system, called *CARE-PARTNER*, proposes to implement the concept of evidence-based medical practice and to monitor its application. *Evidence-based medical practice* emphasizes the performance of medical practice based upon proven and validated knowledge. Available on the WWW both to the LTFU expert and to the home-town, non-expert, primary physician, CARE-PARTNER assists in the clinical tasks of diagnosis, treatment planning, information search and preventive care. The users present their problems to CARE-PARTNER, and CARE-PARTNER proposes a solution for them (see Fig. 1).

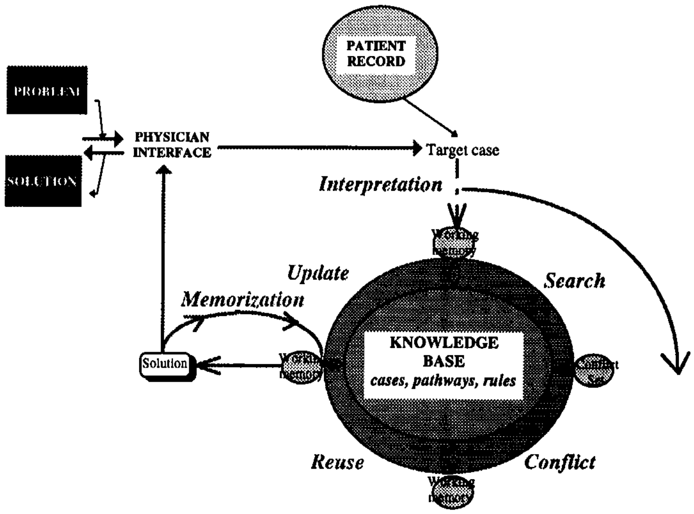

Fig. 1. The CARE-PARTNER system functionality.

An example of the system's input and output is provided on figures 2 and 3 respectively.

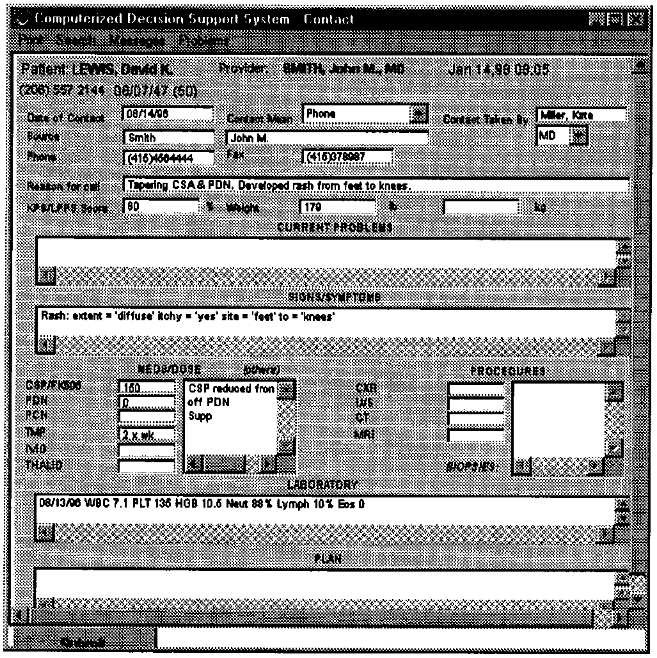

Fig. 2. A case to solve as it is entered to the system.

CARE-PARTNER is a multimodal reasoning system that generates case-specific advice by reusing elements of its knowledge-base, each of which is represented differently depending upon its type:

1. Practice monographs: *practice monographs* are treatises on a certain domain, for example 'bone-marrow transplant', and represented as documents.
2. Practice guidelines: a *practice guideline* is the description of a standardized set of rules to follow for a problem-solving situation (mostly a diagnosis or a treatment). A practice guideline is a compendium of many rules, represented as documents and rules.
3. Practice pathways: a *practice pathway* covers the same type of knowledge elements as a guideline, but it has been created by a group of FHCRC experts by the means of a knowledge acquisition program. It is handled by the system as a single rule, but internally can be represented as a complex of many rules.
4. Practice cases: a *practice case* is a sample of a problem-solving situation. Some are complex, solved by an expert. These are represented as cases.

Case-Based Reasoning

As a CBR system, this system performs a *reasoning* (R_{CBR}) by reusing the elements of a set of previously performed reasoning processes, called a case-base, or *memory* (M), in order to perform a set of *cognitive tasks* (T), such as *diagnosis* or a *treatment planning.*

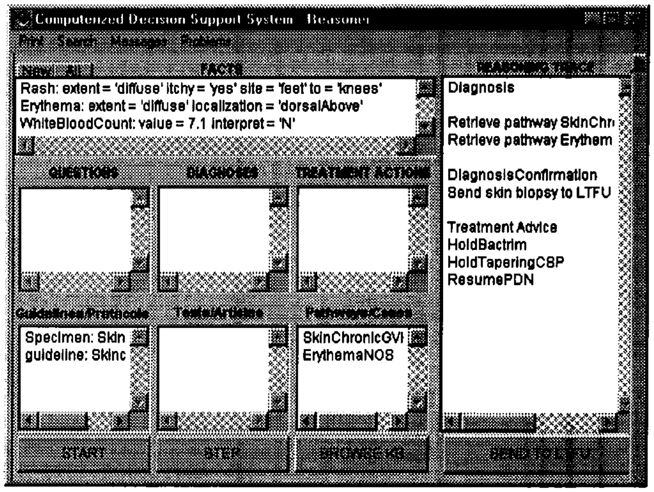

Fig. 3. An output screen of the system.

The *general cycle* of case-based reasoning is a series of inferences grouped in steps (Fig. 4) (Aamodt and Plaza 1994). This cycle starts by the presentation of a *target case* to solve c_t, the form of which depends upon the cognitive task performed. In the *interpretation step* R_i, this initial case is transposed from an external representation language, that of the application domain, into an internal representation language, that of the memorized cases. Thus this interpretation is a translation between two representation languages. Moreover, the indices associated with this case are calculated by an indexing function, which is often the same as the interpretation function. Then, the representation of the case by its indices is used by the system in order to compare this case with the memorized cases, during the *retrieval step* R_t. This step is a search through the memory for the cases similar to the new case c_t, and leads to a set of extracted cases. In the *reuse step* R_u, a choice is made in favor of the case(s) which will serve as the basis of the reuse in order to propose a processing for the target case. The reasoning R proceeds with a *revision step* R_v, during which the proposed processing may be improved. Following, the repaired case may be added to the memory, and the memory enriched by this new experience, during the *memorization step* R_m.

$$R_{CBR} = < R_i, R_t, R_u, R_v, R_m >$$

The memory M of the present system is also its *knowledge-base*, and so contains elements, called *entities*, different from cases.

The system performs a *set of cognitive tasks*, such as *diagnosis*, *planning* or *interpretation.* Indeed, more generally, a target case c_t expresses the terms of a problem to be solved by the system:

$$c_t = <S_i, S_f, Goal>$$

with S_i being the initial situation, or premises of the problem, S_f being the final situation of the problem, and *Goal* being the result expected from the system (and so associated with the task *t*, which is a solution type). For different task types, a target case may be represented differently, for instance:

- diagnosis: $c_t = <S_i, S_f, Goal>$ with $S_f = \{\}$ and $Goal = diagnosis$;
- planning: $c_t = <S_i, S_f, Goal>$ with S_f = *target situation* and *Goal* = *plan* with the initial state satisfied by S_i and the final state satisfied by S_f.

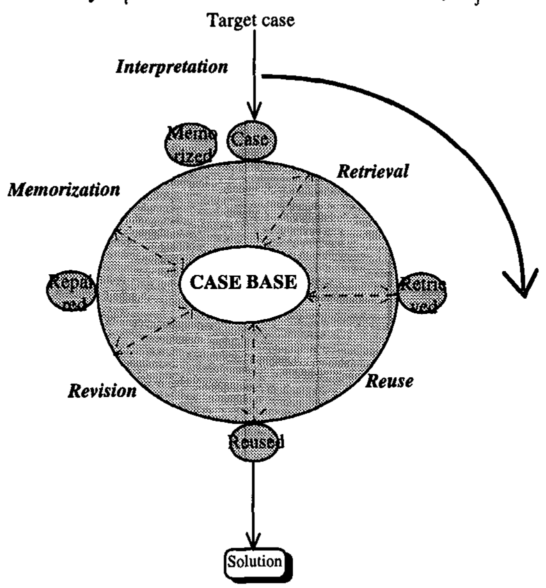

Fig. 4. The case-based reasoning cycle.

A memorized case c_s comprises, in addition to the representation of a target case, a solution *Sol*.

Rule-Based Reasoning

Certain knowledge entities, such as pathways and guidelines, are expressed as rules, and so this system is also a RBR system. As a RBR system, this system performs *rule-based reasoning* (R_{RBR}) by firing the elements of a set of rules, (the rule-base) in order to solve a problem as for CBR. Each rule consists of *a condition* and an *action* (*if condition then action*).

The *general cycle* of rule-based reasoning (Fig. 5) starts by the presentation of a *target problem* to solve c_t. In the *interpretation step* R_i, this initial problem is transposed from an external representation language, that of the application domain, into an internal representation language, that of the rules. Then, the representation of the problem is used by the system in order to match it with the condition parts of the rules, during the *pattern-matching step* R_c. This step is a search through the set of rules for the ones that the problem matches and leads to a set of applicable rules, called the *conflict set*. In the *conflict resolution step* R_r, a choice is done in favor of the rule which will be fired first. The reasoning R_{RBR} goes on with a *production step* R_p, during which the actions in the *action* part of the fired rule are performed. These actions modify the problem description, by adding new representation elements, and/or by modifying others. Following, the elements in the condition part of the rule are removed from the problem description, while the results of the performance of the action part are added to the problem description, and linked to the set of rules, in the working memory, during the *update step* R_n.

$$R_{RBR} = < R_i, R_c, R_r, R_p, R_n >$$

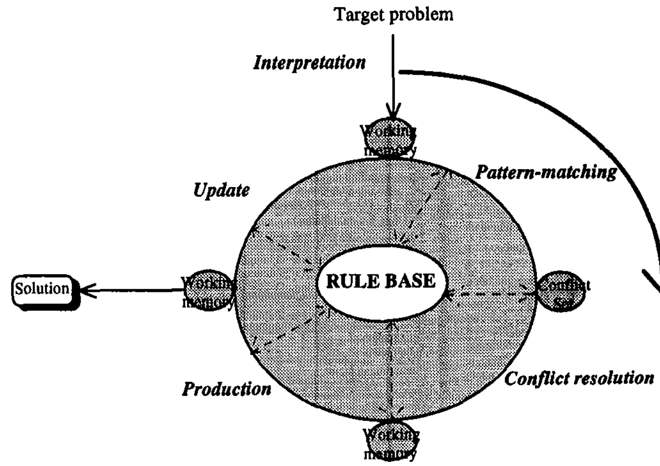

Fig. 5. The rule-based reasoning cycle.

The working memory *W* of the production part of the system contains the updated description of the problem, and so evolves during the reasoning process. Updates are done only in this part of the system, and do not modify the knowledge-base. This is why this memory is a working memory, and contrasts with the CBR memory *M*, which is a long-term memory.

The system performs a *set of cognitive tasks*, such as *diagnosis*, *planning* or *interpretation*. More generally, a target problem c_t consists of the terms of a problem to solve by the system, and has the same description as for CBR.

A rule *r* comprises a condition part C_f and an action part A_i. C_f is matched during $R_{RBR,i}$ with S_i , then with the following S_j (corresponding to the evolution of the working memory) until S_f is satisfied, or *Goal* if it is empty:

$$r_i = <C_i, A_i>$$

Information Retrieval

Certain knowledge entities, such as monographs, articles and guidelines, are expressed as documents, and this system also retrieves pertinent documents during problem-solving. As an IR system, this system performs a *reasoning* R_{IR} by retrieving and presenting to the user a set of documents answering a *query*. Each document consists of *a set of indices* and a *text*.

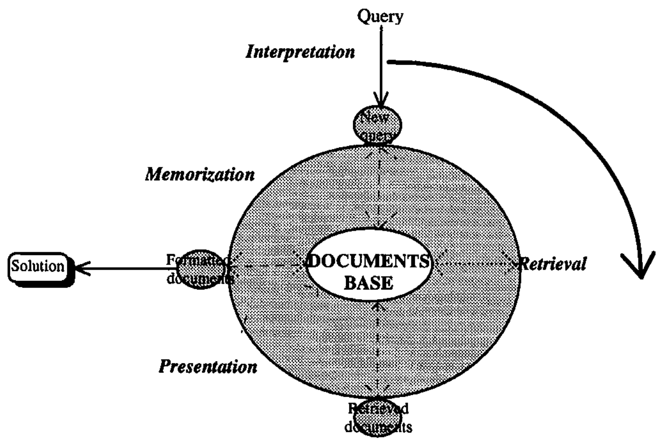

Fig. 6. The information retrieval cycle.

The *general cycle* of information retrieval (Fig. 6) starts by the presentation of a *query* to satisfy q_t. In the *interpretation step* R_i, this initial problem is transposed from an external representation language, that may be the natural language, into an internal representation language, that of the indexes. The indexes, which are the sub-set of the query representation used for the search through the document-base, are used by the system in order to match them with the indexes parts of the documents, during the *retrieval step* R_r. The reasoning R_{RI} goes on with a *presentation step* R_p, during which the retrieved documents are formatted according to the user's needs. In some systems, the user can refine the results by submitting a sub-query to the system, and thus start a new reasoning cycle. Also, some systems keep a trace of the user's search process, in a memorization step R_m:

$$R_{RI} = < R_i, R_p, R_p, R_m >$$

A document d_i comprises an index part I_i and a text part T_f:

$$d_i = <I_i, T_i>$$

The Multimodal Reasoning Cycle

The multimodal reasoning proceeds through different steps (Fig. 7) that are a merge of the reasoning steps of both CBR, RBR and IR.

The knowledge-base of the system is composed of cases, rules and documents, all expressed in a common object-oriented representation language, presented in (Bichindaritz and Sullivan 1998).

This system is capable of handling the wide variety of problems that physicians can face when they care for patients, and the first task of the system is to determine the nature of the problem to be solved. This first step, called the screening step R_Δ, classifies the problem as an information retrieval task, or a patient problem-solving task, or yet another task. Only patient problem-solving tasks are addressed here. An information retrieval task can either be performed alone, or called by the system during problem-solving.

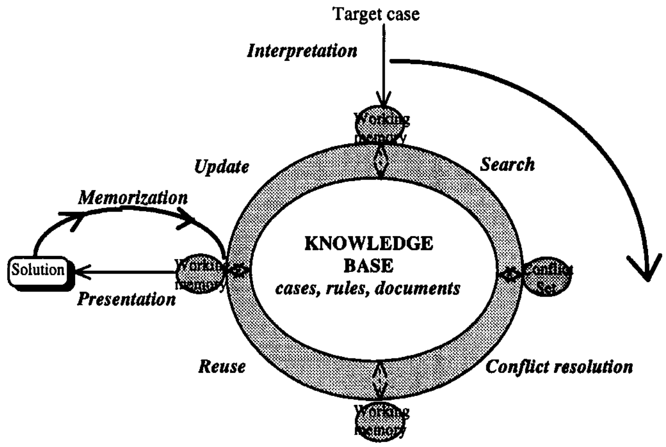

Fig. 7. The multimodal reasoning cycle.

The patient problem-solving task follows several steps:

1) Interpretation (R_i): given the description of a patient problem, the system constructs, by interpretation, the initial situation expressed in the knowledge representation language of the system, mainly an abstraction (Bichindaritz and Conlon 1996). Let c_c be the target patient case to solve.
2) Knowledge search (R_s): the knowledge-base is searched in parallel for applicable rules, pathways and cases, and documents. So the pertinent search methodologies, pattern-matching and case-based retrieval are used in parallel. The result is a conflict set containing both cases, pathways and rules. Let CS be this conflict set: $CS = \{ c_i, r_j, p_k \}$ where the c_i are cases, the r_j are rules and the p_k are pathways. The documents retrieved are not used in the subsequent steps, only when the solution is proposed.
   Conflict resolution (R_r): the first criteria to choose the entity to reuse is the number of problem description elements matched. For equality in this number, a priority order is used, giving preference to the rules, then the pathways, then the most similar case.
3) Reuse (R_u): if the selected entity is a rule, it is fired, if it is a case, its solution is adapted, and if it is a pathway, it is applied (which is comparable to the firing of a rule). The applied knowledge elements may lead either to a solution proposal for the target problem, presented with the documents, or to the generation of new elements in the problem description, also called here the working memory.
4) Update (R_n): the current case, corresponding to the working memory, in the process of being solved is updated, the knowledge representation elements used by firing a rule are marked as used (they can be used

later for the case-based reasoning). The solved problems are removed from the description of the problem. If a solution for the problem answers all the problems in the list, then the processing is stopped and the solution is proposed to the user. Otherwise, the reasoning cycle restarts at the step R_s.

5) Memorization (R_m): when the solution is complete, it is memorized with the target case solved.

The output of the system to the user is a user-friendly formatting of the retrieved entities. Different colors are used to differentiate between guidelines (which contain highly authoritative advice approved by a large scale committee), practice pathways (consisting of useful, but less stringently reviewed expert advice) and cases (which are less certain still, because they represent the opinion of a single expert and are unreviewed). The links existing in the knowledge-base between the retrieved entities are also given to the user. In addition, when case-based reasoning is used, the rules applied are presented to explain the results.

Discussion

Different methods have been described to achieve an integration of case-based reasoning and reasoning methodologies such as rule-based reasoning or model-based reasoning (MBR).

Such "defined" systems can be categorized into two major groups:

- In the first group, RBR is the main reasoning process, and CBR is a heuristic to improve the RBR. In essence, CABARET (Rissland and Skalak 1989) resorts to CBR when the applicable rules are contradictory. In a similar way, ANAPRON (Golding and Rosenbloom 1991) performs RBR, but before firing a rule, checks that the problem to solve is not an exception to this rule. In this case, it resorts to case-based reasoning.
- In the second group of systems, CBR is the main reasoning process and RBR or MBR are used to take advantage of a partial domain model available for one part or another of the reasoning process. Examples of this cooperation are the ALEXIA system (Bichindaritz and Seroussi 1992), where a physio-pathological model is used to abstract indices during the interpretation step. ARCHIE (Goel et al. 1991) judges the pertinence of cases during the retrieval step by using qualitative models of the architecture domain. KRITIK (Bhatta and Goel 1993) uses a qualitative model during the adaptation step, and CASEY (Koton 1988) uses rules to assess the equivalence between features in the retrieval and the adaptation step.

In the present system, it cannot be put forward that either CBR, RBR, or any other reasoning type predominates. Another example of such system is the GREBE system (Branting, 1991). When both methods are applicable, one is chosen based on the degree of certainty of the inferences performed by one method or the other.

We report a system which attempts to foster a closer cooperation between the methodologies. This system separates the reasoning steps of each methodology, such as CBR, RBR or IR. It allows partial reasoning processes and the results are obtained at the end of each reasoning step. If GREBE runs in parallel the complete reasoning cycles, in our system, namely CARE-PARTNER, only the *Search* step can be run in parallel with either the case-based retrieval of CBR, or the pattern-matching of RBR, or the retrieval of IR. It would also be possible to apply a case-based retrieval to the rule-base, or a pattern-matching to the case-base.

Conclusion

The system presented here is applied to a complex medical domain. This research involves a three year project, with an evaluation encompassing both outcomes and cost results scheduled for each system component.

It supports the knowledge of its users at both the expert level, and the non-expert level.

The integration of CBR and other types of reasoning, such as RBR and IR, permits the system to reuse all the types of knowledge available for medical problem-solving.

One of the difficulties of the domain has been the conversion of the textual guidelines into structured knowledge elements, namely rules, that the system can reuse during its reasoning process. We would like to study ways of reducing the length and complexity of translating between the textual format and the system structured format, for example by performing this task semi-automatically, or by writing the guidelines directly in the system knowledge structured format. This would permit to add other types of functionality to the system, such as automatically updating the knowledge from the scientific literature, which becomes more and more fastidious as the volume of the literature increases regularly. It is a promising future direction.

Acknowledgments

This work is supported in part by grant R01HS09407 from the AHCPR (Agency on Health Care Policy and Research).

References

Aamodt, A., and Plaza, E. 1994 Case-Based Reasoning: Foundational Issues, Methodological Variations, and System, *AI Communications* 7(1) 39-59.

Bhatta, S.R., and Goel, A.K. 1993. Model-Based Learning of Structural Indices to Design Cases. In Proceedings of the Workshop of IJCAI-93, Reuse of Designs: an interdisciplinary cognitive approach, A1-A13. Chambery, France.

Bichindaritz, I., and Seroussi, B. 1992. Contraindre l'analogie par la causalité, *Techniques et Science Informatiques* Volume 11, n 4: 69-98.

Bichindaritz, I. 1994. Apprentissage de concepts dans une mémoire dynamique: raisonnement à partir de cas adaptable à la tâche cognitive. Ph.D. diss., Université René Descartes.
Bichindaritz, I., and Conlon, E. 1996. Temporal knowledge representation and organization for case-based reasoning. In Proceedings TIME-96, 24-29. IEEE Society Press.
Bichindaritz, I., and Sullivan K.M. 1998. Reasoning from Knowledge Supported by More or Less Evidence in a Computerized Decision Support System for Bone-Marrow Post-Transplant Care. In AAAI Spring Symposium on Multimodal Reasoning, (in press). Stanford, Calif.: AAAI Press.
Branting, K.L. 1991. Exploiting the Complementarity of Rules and Precedents with Reciprocity and Fairness. In Proceedings of a Workshop on case-based reasoning (DARPA), 130-132. San Mateo, Calif.: Morgan Kaufmann.
Goel, A.K., Kolodner, J.L., Pearce, M., Billington, R. and Zimring, C. 1991. Towards a Case-Based Tool for Aiding Conceptual Design Problem Solving. In Proceedings of a Workshop on case-based reasoning (DARPA), 109-120. San Mateo, Calif.: Morgan Kaufmann.
Golding, A.R., and Rosenbloom, P.S. 1991. Improving Rule-Based Systems through Case-Based Reasoning. In Proceedings of AAAI-91, 22-27. San Mateo, Calif.: Morgan Kaufmann.
Koton, P. 1988. Reasoning about evidence in causal explanations. In Proceedings of AAAI-88, 256-261. San Mateo, Calif.: Morgan Kaufmann.
Rissland, E.L., and Skalak, D.B. 1989. Combining Case-Based and Rule-Based Reasoning : A Heuristic Approach. In Proceedings of IJCAI-89, 524-530. San Mateo, Calif.: Morgan Kaufmann.

Appendix

1. **Integration Name/Category**: CARE-PARTNER (CBR / RBR / IR).
2. **Performance Task:** most clinical tasks, such as diagnosis, treatment planning, prediction, prevention, information search.
3. **Integration Objective:** each methodologies permits to reuse different knowledge elements, CBR for cases, RBR for rules and IR for documents. So the integration objective is to have a coverage of the problems as complete as possible. It is also motivated by the aim of the system, which is to implement the concept of evidence-based medical practice. This practice emphasizes the reuse of guidelines, very formal entities represented as rules. To achieve this goal, it has been necessary to add rules to the cases. Documents are also used because it is not possible to transform all the scientific literature into rules.
4. **Reasoning Components:** CBR, RBR, IR.
5. **Control Architecture:** CBR, RBR and IR are used in a unified way.
6. **CBR Cycle Step(s) Supported:** Pre-processing, Retrieval, Reuse, Revision, Retention, Post-processing.
7. **Representations:** cases, rules and documents, respectively for the CBR component, the RBR component and the IR component.
8. **Additional Components:** a knowledge-acquisition component permits the input of pathways to the system in its knowledge representation language.
9. **Integration status:** Applied.
10. **Priority Future Work:** Evaluation (a controlled clinical trial is scheduled to start for CARE-PARTNER in July) where both patients and physicians outcomes will be measured for one year.