Case-Based Reasoning in CARE-PARTNER: Gathering Evidence for Evidence-Based Medical Practice

Isabelle Bichindaritz, Emin Kansu, and Keith M. Sullivan

Clinical Research Division
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue N., D5-360
Seattle, Washington 98109-1024
{ibichind, ekansu, ksulliva}@fhcrc.org

Abstract. This paper presents the CARE-PARTNER system. 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. CARE-PARTNER aims at implementing the concept of evidence-based medical practice, which recommends the practice of medicine based on proven and validated knowledge. From an artificial intelligence viewpoint, it proposes a multimodal reasoning framework for the cooperation of case-based reasoning, rule-based reasoning and information retrieval to solve problems. The role of case-based reasoning is presented in this paper as the collection of evidence for evidence-based medical practice. Case-based reasoning permits to refine and complete the knowledge of the system. It enhances the system by conferring an ability to learn from experience, and thus improve results over time.

1 Introduction

Evidence-based practice in medicine encourages to emphasize the performance of medical practice based upon proven and validated practice. Since scientific knowledge is based upon evidence, the medical science grounds its knowledge upon evidence. Thus evidence-based practice is matched to knowledge-based practice in this article. The system presented here, CARE-PARTNER, is a *computerized knowledge-support system* on the World-Wide Web (WWW). It means that this system is an evolution of the well-known computerized decision-support systems that supports the quality of the knowledge of both its own knowledge-base and of its users. In concrete terms, CARE-PARTNER assists its users in the performance of their clinical tasks using a general framework for reasoning from knowledge sources of varied quality. This means that their knowledge is based upon varied evidence, but also that the evidence associated to each piece of knowledge can vary through time.

This system is applied to the *long-term follow-up* (LTFU) of patients having undergone a *stem-cell transplant* (SCT) at the Fred Hutchinson Cancer Research Center (FHCRC) in Seattle, after their return in their home community [11]. From an

B. Smyth and P. Cunningham (Eds.): EWCBR'98, LNAI 1488, pp. 334-345, 1998.

artificial intelligence viewpoint, it is an interesting example of the application of case-based reasoning (CBR) integrated with such other artificial intelligence (AI) methodologies as knowledge acquisition (KA), machine learning (ML), rule-based reasoning (RBR) and intelligent information retrieval (IR) methods. CARE-PARTNER's knowledge-support function at the point-of-care is performed by three main factors. Firstly the knowledge-base of CARE-PARTNER is composed of several knowledge-sources, of varied quality: monographs, scientific literature, clinical practice guidelines, pathways and cases. Secondly, this knowledge-base is made available on the WWW to the home-town practitioners who follow-up the transplant patients after their discharge from the FHCRC. It actively helps the monitoring of these patients' care and the effective application of the knowledge at the point-of-care. Thirdly the system introspectively studies its results and learns from its experiences, thus becoming more and more competent, through case-based reasoning. *Case-based reasoning* (CBR) gives the system the ability to learn from experience and to refine its knowledge automatically. This article focuses in presenting the role of case-based reasoning in the context of the application.

The second section defines the concept of evidence-based medical practice and criteria for knowledge quality. The third section presents the knowledge-base of the system. The fourth section shows how case-based reasoning is integrated with rule-based reasoning and information retrieval in the system. The fifth section emphasizes learning from experience. The sixth section provides a discussion of the advances of this system and its implementation is described in the seventh section, just before the conclusion.

2 Evidence-Based Medical Practice

This system proposes to implement the concept of *evidence-based medical practice* in the context of stem-cell post-transplant care. In medical information systems, as in other scientific fields, the quality of practice may be characterized by two main dimensions of the knowledge upon which it is grounded:

– **Reliability:** *the most reliable* knowledge is the one generated by the consensus of a *world-wide committee of experts (practice principles)*. Then, by decreasing order of reliability, we find the knowledge generated by *a committee of experts (practice guidelines)*, then by *a group of experts (practice pathways)*, then by *one expert (practice case)*, then by a person who is not an expert.
– **Certainty:** the *certainty* of knowledge is associated to the *proof(s)* that have permitted to validate it. We can present some knowledge submitted to *world-wide controlled clinical trials* as *almost certain*. Then, by decreasing order of certainty, we find knowledge submitted to *controlled clinical trials*, then *uncontrolled trials*, then knowledge *grounded upon an individual's experience*, and finally on no evidence.

	<Nausea	importance('M')
OR	Emesis	importance('M')
OR	DiarrheaNOS	importance('M')

…,
ColonoscopyNOS result(‘Abnormal’) importance(‘C’) order(1)
AND ColonBiopsy finding(‘ColonCGVHD’) importance(‘N’)order(1)
AND MicrobialCultureNOS site(‘colon’) importance(‘C’) order(1)
AND VitalSignsNOS result(‘Abnormal’) importance(‘M’) order(1)
AND SchirmerTearText result(‘Low’) importance(‘H’)
…>

Fig. 1. Extracts from a ColonChronicGVHD diagnosis pathway.

In a young medical domain such as SCT, and in a pioneer center for this therapy such as FHCRC, the knowledge used for patient problem-solving belongs to almost all of the types quoted earlier:

1. **Practice guidelines:** a *practice guideline* is composed of “systematically developed statements designed for practitioners and patients that will be helpful in making clinical decisions on the prevention, diagnosis, treatment and management of selected conditions”. Practice guidelines are created by specific committees from the scientific literature, complemented by the professional judgment of the committee’s members.
2. **Practice pathways:** a *practice pathway* covers the same type of knowledge elements as a guideline, but specialized in the management of diagnosis and treatment related to LTFU. It has been created by a group of LTFU experts exclusively for the CARE-PARTNER system, and by the means of a knowledge acquisition program.
3. **Practice cases:** a *practice case* is a sample of a problem-solving situation, some being complex, solved by an expert. This problem-solving example may have needed to resort to the expertise and agreement of several experts, but it is essentially a real patient problem-solving situation, and not a standardized, prototypical one as a practice guideline or a practice pathway.

3 CARE-PARTNER Knowledge-Base

The system *knowledge-base* is a network of *entities* (such as the practice guidelines, pathways and cases), where entities are *nodes*, connected by *links*, or *relationships*. The links are those defined in the Unified Medical Language System (UMLS) [14], that provides an ontology (concepts and links) for the medical domain in general. This ontology has been refined for the application domain, but the set of 50 links provided between the concepts was sufficient. A knowledge acquisition (KA) workshop has been created for CARE-PARTNER. Its access is restricted to the knowledge-base creation and maintenance team of LTFU.

The elements of the representation language are those of semantic networks:

- A domain ontology, which is the set of *class symbols* (also called concepts in the UMLS) C, where C_i and C_j denote elements of C, and the set of *relationship symbols* (also called relations in the UMLS) R, where R_i and R_j denote elements of R. The classes are organized in a polyhierarchy of classes, and several main categories can be described, such as *Functions*, *Diseases*, *Morphology*, *Topography* for instance. SNOMED International v. 3.4 has been used to codify the ontology, whenever possible. Also, many classes describe *Events*, *Time* and *State* concepts [17].

- A set of *individual symbols* (also called instances) I, where *i* and *j* denote elements of I. Among these, some refer to instances of classes, others to numbers, dates, and other values. Instances of a class C_i are noted aC_i.
- A set of *operator symbols* O, permits to form logical expressions composed of classes, instances and other values, and relationships. Pathways, guidelines and cases are expressed this way, and such a composition permits to represent complex entities in a structured format. The set of operators comprises the following: ∧ (AND), ∨ (OR), ¬ (NOT), ATLEAST *n*, ATMOST *n*, EXACTLY *n*, ≥, ≤, <, >, =

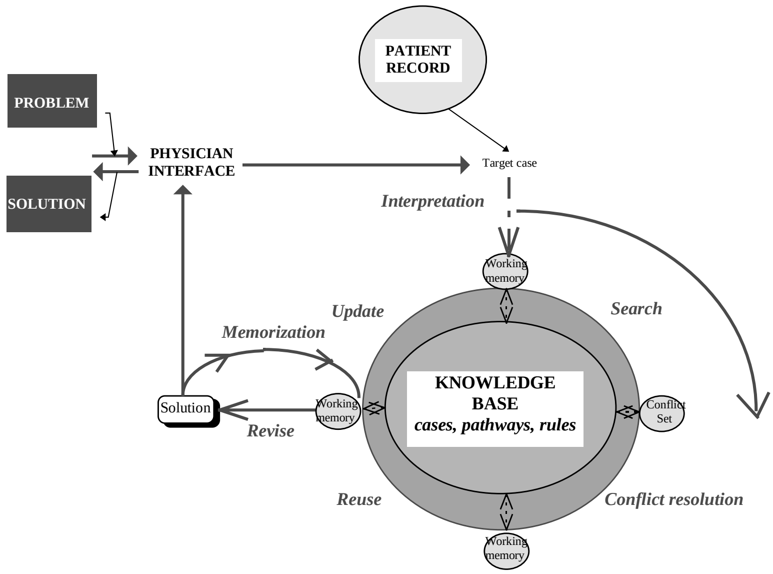

Fig.2. CARE-PARTNER reasoning cycle.

In this representation language, the attributes of a class are represented via the relationships. A binary relationship, the arguments of which are an instance and a value mean that a certain attribute of a class has or gets a certain value.

Guidelines and pathways are expressed as rules <*condition*, *action*>, and cases are expressed as <*problem situation*, *solution*>, where *condition* and *problem situation* have the same representation, and *action* and *solution* also have the same representation, either a composition of instances with operators, or relationships between instances and values:

$$\textit{problem situation} = \Theta\ aC_i$$
$$\Theta\ R_j\ (aC_{j,1}, aC_{j,2}, \ldots aC_{j,k})$$
$$\textit{solution} = \Theta\ aD_i$$

$$\Theta\ R_j\ (aD_{j,1}, aD_{j,2},\ldots aD_{j,k})$$

with $\Theta \in$ **O**, the default value being $\vee$ for pathways, and $\wedge$ for cases and rules. Cases are expressed using only $\wedge$.

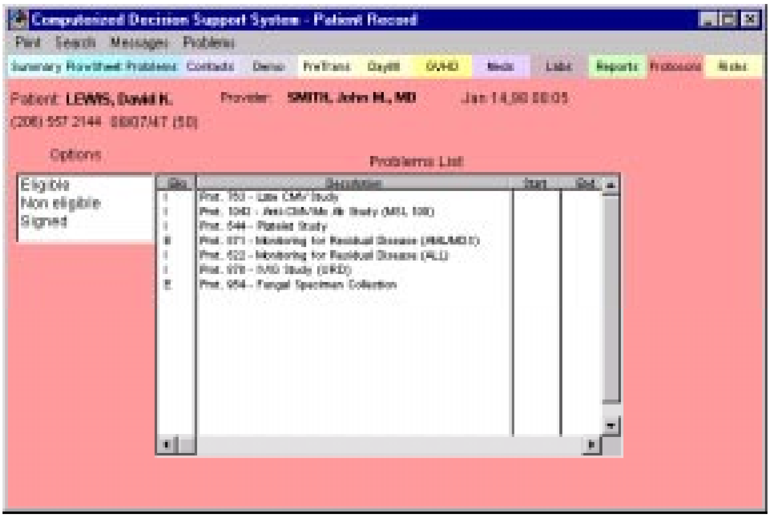

Fig. 3. The protocols page of the electronic patient record notebook on the WWW.

Although the representation language is the same for all the entities, the categories of knowledge used in problem situations and in solutions for instance are quite different. In a medical domain as this one, problem situation elements are mostly instances of the *Functions* hierarchy, mainly *SignsAndSymptoms* and *Findings*. This depends also upon the task to perform.

There are two main categories of guidelines: diagnosis guidelines, and treatment guidelines. Guidelines have been defined by the standard practice committee of FHCRC and/or LTFU and/or an institution. Their transcription into a form suitable for the system has been performed by the team members. Pathways have been defined directly in the system knowledge representation language, by LTFU experts, through the knowledge acquisition workshop developed for this project (Fig. 1 shows a pathway example). Patients cases have been transcribed from the patients files and existing database records into a form suitable for the system by the team domain experts. For all the entities, a textual description is provided in addition to the formalized representation.

4 CARE-PARTNER Reasoning Framework

From the end-user viewpoint, CARE-PARTNER is a knowledge-support system on the WWW. The end-user may be either a LTFU clinician (physician or nurse) or a home-town physician or nurse.

CARE-PARTNER is securely protected by a comprehensive security plan following the organizational and technical recommendations of [10].

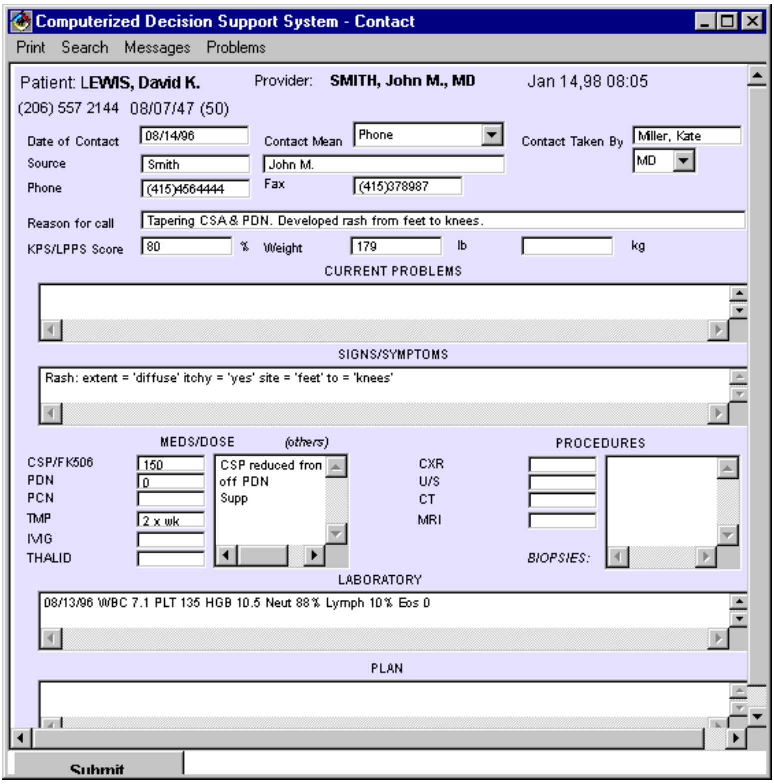

Fig.4. The contact screen permits a user to submit a problem to CARE-PARTNER on the WWW.

4.1 General Presentation

As a classical decision-support system, CARE-PARTNER is “an active knowledge system which uses two or more items of patient data to generate case-specific advice” [2]. This definition outlines the three main components of CARE-PARTNER on the WWW (Fig. 2):

- **the knowledge-base**, represented by the large lower circle on Fig. 2. It is presented in the preceding section.
- **the electronic patient record**, represented by the small upper circle on Fig. 2. It is a notebook gathering all long-term follow-up information available about the patient (Fig. 3 shows one page of it).
- **the system reasoning**, represented by the succession of arrows, starting by the problem submitted by the user on the top-left, and finishing by the solution provided by the system on the top-left.

So the user submits problems to CARE-PARTNER, and CARE-PARTNER provides solutions, that he/she takes as advice.

4.2 Multimodal Reasoning Framework

The system reasoning is a multimodal reasoning [7,8], designed for the cooperation of case-based reasoning, rule-based reasoning and information retrieval.

CBR solves a problem by adapting the solution of a memorized problem judged similar to the problem to solve. This methodology is used in CARE-PARTNER to reuse practice cases, but also pathways. RBR is used to reason from the guidelines and the pathways, both represented by rules. Since the knowledge-representation language is the same for all the entities, CBR and RBR can both be applied to the reuse of all knowledge-base entities. This capability is applied only to pathways in CARE-PARTNER at this time, since they go between the two.

The reasoning starts with the presentation to the system of a new problem to solve. This system is capable of handling the wide variety of problems that physicians can face when they take care of patients, and the first task of the system is to determine the nature of the problem to solve. 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 the patient problem-solving task is addressed here. For this task, the reasoning of the system proceeds through the following steps [1]:

1. **Interpretation (R_i):** given the description of a patient problem on the WWW (Fig. 4), the system constructs, by interpretation, the initial situation expressed in the knowledge representation language of the system. Abstraction is the main reasoning type used here, and in particular temporal abstraction to create trends from time-stamped data [5]. Numerical values are abstracted into qualitative values. Let c_c be the target patient case to solve:

$$c_c \quad = \quad \Theta \ aC_i$$
$$\Theta \ Rj\,(aC_{i,1},\ aC_{i,2}, \ldots\ aC_{i,k})$$

   For example Fig. 4 represents a problem submitted to CARE-PARTNER. It is described by the physician in a set of sub-problems, signs and symptoms, here a *diffuse itchy rash from feet to knees*, current medication, procedures and laboratory results. The context of this rash is the tapering of the immunosuppressive therapy. The laboratory results are interpreted as *Normal* (value 'N').

2. **Knowledge search (R_s):** the knowledge-base is searched in parallel for applicable rules, pathways and cases. So the pertinent search methodologies, pattern-matching and case-based retrieval are used in parallel. The result is a conflict set containing at the same time 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.

   In our example, the trigger *rash* permits to retrieve two pathways, only partially matched to the problem, through case-based retrieval, *Erythema* and *SkinChronicGVHD (Graft Versus Host Disease)*. The second diagnosis is ranked first, since the patient is already treated for *AcuteGVHD*, through immunosuppressive therapy. One document is retrieved: the guideline for

SkinChronicGVHD. When the diagnosis will be confirmed, its rules will be activated.

3. **Conflict resolution (R_r):** given the priority given in the system to the reuse of knowledge based upon proven practice, the following hierarchy of reuse is followed:

   I. reuse rules
   II. reuse pathways
   III. reuse cases

   Nevertheless, the first criteria to choose the entity to reuse is the number of problem description elements matched. The entities are ranked by decreasing number of matched problem description elements with the target case to solve, comparing on one hand all the rules of *CS*, on another hand all the pathways of *CS*, and on another hand each case separately.

   For equality in this number, the priority order is used, giving preference to the rules, then the pathways, then the most similar case. In our example, where the rash and the antecedent of GVHD are the main symptoms, the choice will be between two pathways, and will prefer the *SkinChronicGVHD*.

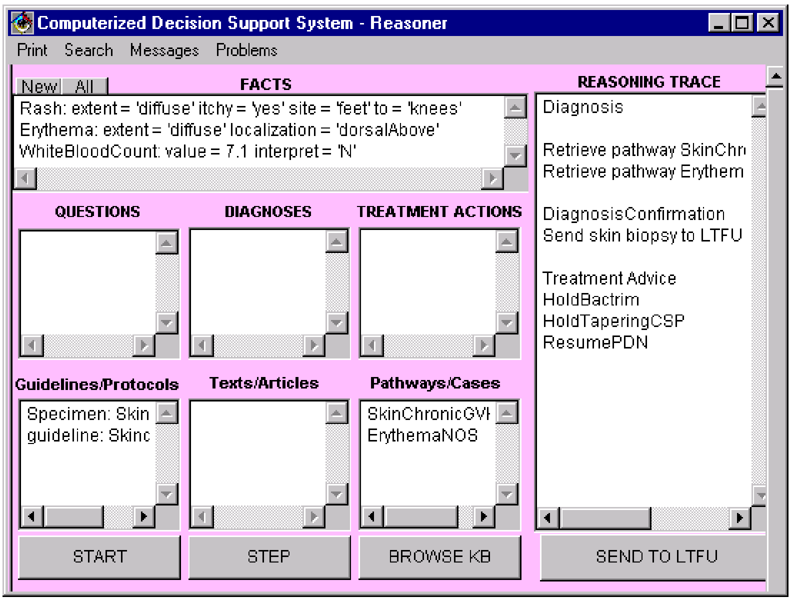

Fig. 5. The result screen provides advice and pertinent documents on the WWW.

4. **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, or to the generation of new elements in the problem description, also called here the working memory.

In our example, the treatment associated to the *SkinChronicGVHD* pathway is proposed. The adaptation takes the form of a customization of the treatment plan, for instance by proposing to *HoldTaperingCSP*, suited only to patients already tapering their immunosuppressive therapy.

5. **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). What are updated are the problems dealt with. 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.
6. **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 frames are used to differentiate between guidelines, practice pathways, cases, scientific literature and other documents. The treatment advice proposed on Fig. 5 is given to the home-town practitioner only after confirmation by the LTFU specialist (but the other information on the screen is given beforehand). The links existing in the knowledge-base between these retrieved entities may be browsed by the user.

In our case example, a second reasoning cycle permits IR to retrieve the *SpecimenShippingInstructionsForSkinBiopsy*, after the treatment plan *SendSkinBiopsyToLTFU* has been added to the working memory.

5 Learning From Experience

The memorization step (Fig. 2 and previous section) is when learning from experience occurs. Several stages of learning are performed [4]:

- **Case learning:** a new case is created with the problem situation and its solution, as proposed by the system. Future follow-up contacts will modify information in this case depending upon the reported results of the solution. Since this case is memorized, it will be possible for the system to reuse it in a subsequent problem-solving episode.
- **Positive feedback:** when the following contacts provide positive feedback about the previous solution, it is marked as validated in the memorized case.
- **Negative feedback:** when the following contacts provide negative feedback, the previous solution is marked as non-valid, and the LTFU experts can give to the system a better solution (if they know one). This will permit to the system to learn from this mistake, by studying the differences between this case and the knowledge elements it has reused by error. Most of the time, the mistake comes from an incompleteness in the knowledge-base, and the addition of the new case will complete this lack of knowledge.

6 Discussion

Different methods have been used to achieve the cooperation of case-based reasoning and other formalisms of knowledge representation such as rule-based reasoning or model-based reasoning (MBR).

Two main categories of such systems can be described:

- The systems for which RBR is the main reasoning process, and CBR is mainly a heuristic to improve the rule-based reasoning. CABARET [16] resorts to CBR when the applicable rules are contradictory. In a similar way, ANAPRON [13] performs a rule-based reasoning, 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.
- The systems for which 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 [6], where a physio-pathological model is used to abstract indices during the interpretation step. ARCHIE [12] judges the pertinence of cases during the retrieval step by using qualitative models of the architecture domain. KRITIK [3] uses a qualitative model during the adaptation step, and CASEY [15] uses rules to assess the equivalence between features in the retrieval and the adaptation step.

In the present system, it cannot be said that either the CBR or the RBR, or another reasoning type, is more important. Another example of such system is the GREBE system [9], in which RBR is not only used to constrain the CBR. When both methods are applicable, one is chosen based on the degree of certainty of the inferences performed by one method or the other. This system, applied to the law domain, uses argumentation for weighing the relevance of a previous case to a new case to solve. Thus arguments in this domain play a role similar as evidence in CARE-PARTNER.

The present system attempts to perform a closer cooperation between the methodologies, by separating the reasoning steps of each of them, and allowing to run in parallel only certain steps, such as the *Search* step. Partial results of each methodology can be used during different reasoning cycles. Thus the integration of CBR with RBR is here closer than previously. In addition, IR is also used in the system via its common reasoning steps.

7 Implementation

CARE-PARTNER is deployed on the WWW. For reasons of portability, it has been developed mainly in Java, and for reasons of efficiency in C++. The ease of integration of native code (C++ DLLs) in Java has permitted to reuse several components, such as a RETE-based object-oriented inference engine, an objected-oriented case-based reasoner and an intelligent information retrieval engine, all capable of interaction with large databases. The architecture of CARE-PARTNER is a three-tiered client-server architecture, and uses the remote method invocation (RMI) as a communication protocol between the client and the server. The client consists of pure Java code, and communicates via RMI with Java servlets on the Java-programmed Web-server. The Java servlets on the server call native C++ for the inference-intensive parts of the reasoning process, such as case-based retrieval, rule-based reasoning and information search. A database server, in this system IBM's DB2, is the repository for all documents, patient cases and pathways. The ontology also resides in the database, with the whole knowledge-base, but is partly incorporated in persistent objects on the server, accessible by the different components via RMI.

8 Conclusion

CARE-PARTNER is a knowledge-support system on the WWW. It provides an example of a complex medical application running on the Internet and integrating case-based reasoning. The definition of a common representation language for the knowledge entities, such as cases and rules, and common reasoning steps for CBR, RBR and IR, have permitted this integration. This system is now about to be brought in randomized controlled clinical trial to assess its effect on physician satisfaction and quality of care, patient outcome and quality of life, and cost of care.

The perspectives of evolution for CARE-PARTNER are to develop the predictive problem-solving component, not yet sufficiently addressed, and to prepare the system for a continuous update of its knowledge from the evolution of the scientific literature. We would like also to refine the rules from the cases, and to learn new rules, thus enriching the pathways. Machine learning methods will permit to reach this goal.

Acknowledgments

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

References

1. Aamodt, A., Plaza, E. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System, *AI Communications* 7(1) (1994) 39-59
2. van Bemmel, J.H., Musen, M.A. Handbook of Medical Informatics. Springer-Verlag (1997)
3. Bhatta, S.R., Goel, A.K. Model-Based Learning of Structural Indices to Design Cases. In Proceedings of the Workshop of IJCAI-93, Reuse of Designs: an interdisciplinary cognitive approach. Chambery, France (1993) A1-A13

4. Bichindaritz, I. Apprentissage de concepts dans une mémoire dynamique: raisonnement à partir de cas adaptable à la tâche cognitive. Ph.D. diss., Université René Descartes (1994)
5. Bichindaritz, I., Conlon, E. Temporal knowledge representation and organization for case-based reasoning. In Proceedings TIME-96, 24-29. IEEE Society Press (1996)
6. Bichindaritz, I., Seroussi, B. Contraindre l'analogie par la causalité, *Techniques et Science Informatiques* Volume 11, n 4: (1992) 69-98
7. Bichindaritz, I., Sullivan, K.M. 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. Stanford, Calif: AAAI Press (1998) 85-90
8. Bichindaritz I., Kansu E., Sullivan K.M. Integrating Case-Based Reasoning, Rule-Based Reasoning and Information Retrieval for Medical Problem-Solving. In AAAI Workshop on CBR Integrations (in press).
9. Branting, K.L. Exploiting the Complementarity of Rules and Precedents with Reciprocity and Fairness. In Proceedings of a Workshop on case-based reasoning (DARPA). San Mateo, Calif.: Morgan Kaufmann (1991) 130-132
10. Clayton, P.D., et al. For the Record. Protecting Electronic Health Information. National Academy Press (1997)
11. Sullivan K.M., Siadak M.F. Stem Cell Transplantation. In Johnson FE, Virgo KS, Edge SB, Pellegrini CA, Poston GJ, Schantz SP, Tsukamoto N (eds.). Cancer Patient Follow-Up. St Louis: Mosby Year-Book Publications (1997) 490-501
12. Goel, A.K., Kolodner, J.L., Pearce, M., Billington, R., Zimring, C. 1991. Towards a Case-Based Tool for Aiding Conceptual Design Problem Solving. In Proceedings of a Workshop on case-based reasoning (DARPA), San Mateo, Calif.: Morgan Kaufmann (1991) 109-120
13. Golding, A.R., Rosenbloom, P.S. Improving Rule-Based Systems through Case-Based Reasoning. In Proceedings of AAAI-91, San Mateo, Calif.: Morgan Kaufmann (1991) 22-27
14. Humphreys B.L., Lindberg, D.A.B., Schoolman H.M., Barnett G.O. The Unified Medical Language System: An Informatics Research Collaboration. J Am Med Inform Assoc 1 (1998) 1-11
15. Koton, P. Reasoning about evidence in causal explanations. In Proceedings of AAAI-88, San Mateo, Calif.: Morgan Kaufmann (1988) 256-261
16. Rissland, E.L., Skalak, D.B. 1989. Combining Case-Based and Rule-Based Reasoning : A Heuristic Approach. In Proceedings of IJCAI-89, San Mateo, Calif.: Morgan Kaufmann (1989) 524-530
17. Shahar Y. A Framework for Knowledge-Based Temporal Abstraction. Artificial Intelligence 90(1-2) (1997) 79-133