G. Lanzola

Dipartimento di Informatica,
Universita' degli Studi di Pavia,
Pavia, Italy.

Synopsis

Collaborative Knowledge Processing

Introduction

Perhaps the first time researchers realized that solving complex problems requires collaborative knowledge processing was around the late 1980s. By that time, it became apparent that KBSs were bound to remain small in size and only able to address very simple problems if their whole knowledge bodies were to be built every time from scratch by a small pool of experts, if not just by a single expert. Thus, in order to improve the performance of KBSs and solve problems with increasing levels of complexity, researchers started pursuing new ways of preserving existing knowledge bases and building on them, thereby bringing forward knowledge sharing and reusability issues.

The network infrastructure was still in its early stages and, at least in the beginning, sharability and reusability could only take place in terms of component modules that could be easily exchanged between different applications. Nevertheless, what happened is that almost every research group implemented those concepts on top of a specialized framework, thereby giving rise to a large number of still isolated islands and ending up with exactly the same kind of problems the new methodology was supposed to solve.

Fortunately, in less than one decade the situation concerning the network infrastructure has dramatically changed, and an experiment such as the Internet has definitely proven its excellent potential as a cheap and widely available means of sharing information. Time is now ripe for a new paradigm shift which entails looking at it as the underlying framework for building and distributing a whole set of knowledge-level services required to proficiently support the collaboration among health-care personnel.

All the papers in this section illustrate interesting approaches which could have an enormous impact on collaborative knowledge processing if bridged with network technologies, although only one of them [1] explicitly addresses the topic. In order to support it we have to develop frameworks for *Knowledge and Information Representation and Interchange* [1-3]. There is a strong emphasis on this in the medical domain since it is known that it takes years for important changes in medical knowledge to make their way into the daily practice. Then we need suitable *Automatic Knowledge Acquisition* techniques to help in representing and formalizing the increasing amounts of information available [4,5]. Finally, we should provide frameworks for *Knowledge Based Applications* in order to proficiently and effectively exploit knowledge for increasing the quality of health-care services while reducing the associated costs [6].

Knowledge and Information Representation and Interchange

One of the main bottlenecks which hinders a proficient collaboration among health-care workers is represented by the existing impediments in accessing relevant information and knowledge. While this is an important issue for the daily practice it becomes even more serious in research environments. Analyzing biological systems calls for an ever-increasing specialization because of their inherent complexity. However, research strictly depends on the integration of contributions from multiple heterogeneous sources as part of the process of generating new knowledge.

Several paradigms have been developed over the years with the aim of facilitating formalization and exchange of knowledge. Conceptual Graphs were first introduced by Sowa [7] as a mean of formalizing natural language analysis and understanding, and given their similarity with semantic networks they exhibit very high expressiveness and representation power. Thus, they are frequently used in complex problems addressing knowledge representation issues since they provide a unifying model which facilitates the mapping of a representation onto another one while preserving the underlying semantic content. The paper by Graves et al. [2] tries to leverage this feature proposing a conceptual model aimed at simplifying the process of designing and implementing databases for a genome center, and considers collaboration from two different perspectives. At the database design level, conceptual graphs are introduced as a means of achieving cooperation between the domain expert and the software engineer since they provide a middle ground for exchanging knowledge and bridging the gap that exists between their different expertise. The database is constructed through a sequence of modeling and refinement steps centered on the interaction between the two experts, through which the most appropriate structures for representing data are identified. Nevertheless, the semantic model underlying the database is also exploited for improving cooperation among biologists and overcoming one of the main problems experienced in this area. While the research could certainly benefit from a more abstract analysis aimed at pointing out similarities and differences among DNA sequences of genes belonging to different organisms, the lack of a conceptual model has prevented, so far, the integration of that information across species boundaries.

Exploiting the information technology for implementing large-scale clinical repositories has always been a challenging task, despite the availability of good methodologies for developing models which closely capture the application domains. What happens is that rendering the inherent complexity of a medical domain into a database automatically translates into poorly performing applications. Often, ad-hoc solutions are adopted that in trying to reduce the complexity of the application also cause the loss of some of the underlying semantic constraints. The paper by Johnson [3] illustrates the "generic data modeling" technology with the aim of overcoming the problem and reconciling the two opposite requirements concerning data modeling and database design. While conventional approaches are still appropriate for describing the conceptual schema, generic data modeling techniques may be adopted for making explicit a set of transformations over that schema in order to achieve an efficient patient database design. Even in this paper, conceptual graphs are adopted for formally defining the generic transformations thereby preserving the underlying semantics of the model. The proposed approach also increases the flexibility and the adaptability of the database. This is an interesting feature since formalized knowledge of a health-care process is always evolving and incomplete, and the rapid emergence of new insights may continuously require changes in the domain structure.

Any discussion addressing collaborative environments can not refrain from quoting the Internet and the Web which has turned it into an easy to use technology having the potential of revolutionizing the way in which medical information is accessed. Web Browsers were initially conceived only as a means of distributing information by supporting hypertextual navigation combined with some capabilities for rendering images. Nevertheless, their widespread availability along with the adoption of a uniform way for accessing and displaying information has rapidly and forcibly turned them into generic and platform-independent interface building tools. All of this happened in spite of the obvious limitations shown by the adopted protocols, whose exploitation according to this new paradigm required some clumsy expedients going to the detriment of the performance. The use of network services for providing a collaborative environment for information exchange among domain experts is also the central theme in the paper by Peterson et al. [1] which illustrates how database technologies may be successfully merged with Web ones in order to easily make available high-level services worldwide. More specifically, the paper describes a system acting as a front end for both a modeling environment and a cluster of databases. The system is able to facilitate the collaborative development of models among experts while enforcing the referential integrity among the model, the parameters adopted for running the simulations and the relevant results which are all stored in the databases. Web support is exploited for making available through the Internet a sophisticated graphical environment. This simplifies the process of building models and adjusting simulation parameters even for domain experts in the area of neuroscience, who do not usually have the expertise in any particular simulation package, nor a deep knowledge in mathematics.

Automatic Knowledge Acquisition

Building frameworks for disseminating information is only part of the problem of developing collaborative care environments. There is an urgent need also for methodologies and tools helping human experts to deal with the growing amount of data in order to easily and rapidly extract useful knowledge. Statistical and probabilistic techniques based on the Bayes' theorem assuming a conditional independency of the data were probably among the first ones to be used for performing automatic classification within medical contexts. They perform adequately well in narrow domains, but exhibit problems in others. Furthermore, those approaches are inherently affected by a lack of transparency which is a serious drawback for medical applications. The same black-box nature is also typical of other recently emerged techniques which are increasingly gaining popularity within the AI community. Artificial neural networks, for example, have proved to be adequate for formulating many different kinds of predictions while being competitive in terms of performance. However, since neural networks do not formalize knowledge, it is impossible to modify them by any means other than by training on specific sets of data. This also means that it is not possible to analyze what they have learnt and, even worse, they cannot be explicitly taught any medical expertise which is unavailable in the initial training data set.

On those grounds more principled approaches have always been preferred for AI applications in medicine, such as those performing classification tasks by creating hierarchies of concept descriptions. The paper by Ohmann et al. [5] actually addresses this topic by analyzing the performance of six different techniques of automatic rule induction applied in a prospective study, and comparing them with the standard model based on Bayes' theorem. The different algorithms presented no striking differences between each other and with the standard Bayes' reference model. Unfortunately they only showed an average diagnostic accuracy comparable to that of a junior physician. The reason for this probably lies in the dimensionality of the problem, suggesting that a closer integration with distributed databases may result in a possible performance improvement.

A different perspective emerges from the paper by Braaten [4] which uses the ID3 algorithm to build a binary decision tree for identifying important signs in the diagnosis of newborn syndromes. A decision support system is generated after a set of artificial cases built from the information available in the literature and subsequently run on a different set of test cases. The performance is rather poor, probably because of the implicit unrealistic assumption of a closed world. Nevertheless, the algorithm classified as important all the well-established clinical signs. This may be of potential help for rapidly focusing a human expert while addressing new situations.

Knowledge-Based Applications

Perhaps the most renowned results of applying AI methodologies in health-care are represented by computer-based clinical decision support systems. The open challenge driving the most ambitious research efforts in this area has been that of emulating within a computer system the reasoning capabilities of the human expert for solving highly structured tasks, such as diagnosis or therapy planning, in a complex and less formalized domain like medicine.

Several studies point out that additional recommendations, such as those provided by computerized medical systems, could definitely improve the effectiveness and efficiency of patient care and even achieve a better compliance of the clinicians in following treatment protocols and guidelines. Nevertheless building those systems is still a daunting task given that huge bodies of knowledge must be acquired and represented even for solving small classes of closely related problems, and further methodological and technological advances are required. This is witnessed by the large number of papers illustrating the construction of prototypical systems accomplishing knowledge-level tasks compared to the very few ones addressing real evaluations of the effects arising from adopting the systems in the daily practice, or simply trying to assess their reliability and performance through sound controlled trials [8].

Furthermore, there are still many open issues concerning ethical, legal and technical aspects preventing the exploitation of those tools in the real clinical practice. The more cognitive scientists understand the complex and ever-changing nature of medical knowledge the clearer it is that human experts will always be involved as key characters in every decision-making process. In fact there is no evidence so far, and probably there will not be any in the near future, that computer systems, no matter how complex they are, will measure up to the capabilities of the human mind in dealing with new and unexpected situations, and in integrating different and apparently unrelated pieces of information which are instead determinant for properly undertaking clinical decisions.

Thus, instead of envisioning tools supporting the whole diagnostic or therapeutic planning processes, a challenging area where AI methodologies could be proficiently exploited even on a short-term perspective is that of intelligent real-time patient monitoring systems. These systems may definitely help health-care personnel in improving the quality of assistance by being operational all-day long with low running costs and performing smoothly over time. Achieving the same level of assistance only with humans would be almost impossible, as witnessed by surveys pointing out that errors are very likely to occur on protracted shifts and especially across their changes. Moreover, whenever those systems are about to undertake important decisions, they may easily request an acknowledgment by the human expert who in the meanwhile may be available for accomplishing different tasks.

This issue is well addressed in the paper by Dojat et al. [6] which evaluates a KBS connected to a ventilator able to predict the ability of patients to tolerate total withdrawal from ventilatory support. The system runs on a microcomputer located at the patient's bedside, is connected both to the ventilator and a gas analyzer and accomplishes two different tasks. On one hand it monitors the patient's clinical situation and continuously adapts the ventilatory assistance to his/her needs, possibly decreasing it gradually in order to prepare the patient for the weaning. On the other hand the system selects an optimal strategy for weaning and indicates the most appropriate moment for performing tracheal extubation. The system operates in closed-loop and successfully combines AI techniques concerning knowledge representation and temporal reasoning with real-time issues in order to guarantee a timely response even in complex situations. The paper reports on the evaluation of the system performed on 38 real patients and aimed at comparing the advice provided by the KBS with the decision usually undertaken by a physician and based on a set of conventional tests. Although the number of cases is small, the results are encouraging and show even a better performance than that achieved by the physicians adopting conventional methods.

Conclusions

It is clear that despite several interesting achievements during the past decade in areas concerning knowledge processing, a real application of those methodologies in clinical practice is still to come, and the reasons for that failure are manifold.

Certainly solving medical problems is a challenge on its own, since diverse and complex forms of knowledge representation are involved, requiring multiple reasoning strategies. Nevertheless, it must be said that the whole research in AI in the past ten years has been mostly shaped according to a *technology-push* paradigm rather than after an *application-pull* one. More specifically, the research has focused too much on basic issues concerning knowledge modeling and representation, and too little interest has been put on pursuing suitable methodologies and techniques for exchanging and processing that knowledge in a shared environment.

However, managing patients is definitely a collaborative and knowledge-intensive process involving several persons, including physicians, nurses, therapists, technicians, administrators and clerks, all with a slightly different view about what the process is. Furthermore, medical care delivery has distinguishing features that make such cooperation difficult to support since it is inherently connected with the distribution of resources, skills, control and data along two different dimensions such as time and space.

Supporting the requirement of medical applications requires further research efforts in all the traditional AI areas *concerning knowledge modeling and representation, machine learning and adaptation, cognitive modeling and epistemology*. However, it is clear that no application is ever bound to succeed if it does not also tackle those organizational issues which are part of the medical practice. This is also the goal of new emerging projects such as InterMed which tries to exploit the network for providing knowledge-level services within a medical context.

References

1. Peterson BE, Healy MD, Nadkarni PM, Miller PL, Sheperd GM. ModelDB: An environment for running and storing computational models and their results applied to neuroscience. J Am Med Inform Assoc 1996;3:389-98.
2. Graves M. Bergeman ER, Lawrence CB. A graph conceptual model for developing human genome center databases. Comput Biol Med 1996;26,3:183-97.
3. Johnson SB. Generic data modeling for clinical repositories. J Am Med Inform Assoc 1996;3:328-39.
4. Braaten O. Artificial intelligence in pediatrics: important clinical signs in newborn syndromes. Comput Biomed Res 1996;29,153-61.
5. Ohmann C, Moustakis V, Yang Q, Lang K. Evaluation of automatic knowledge acquisition techniques in the diagnosis of acute abdominal pain. Artif Intell Med 1996;8:23-6.
6. Dojat M, Harf A, Touchard D, Laforest M, Lemaire F, Brochard L. Evaluation of a knowledge-based system providing ventilatory management and decision for extubation. Am J Respir Crit Care Med 1996;153:997-1004.
7. Sowa JF. *Conceptual Structures: Information Processing in Mind and Machine*. Reading, MA: Addison Wesley, 1984.
8. Wyatt J, Spiegelhalter D. Evaluating medical expert systems: what to test and how? Med Inform 1990;15:205-17.

Address of the author:
Giordano Lanzola,
Dipartimento di Informatica,
Universita' degli Studi di Pavia,
Via Ferrata 1,
27100, Pavia, Italy.
E-mail: giordano@aim.unipv.it