Guided Interactive Diagnostic Assistance

Giovanni Maria Sacco
Università di Torino, Italy

INTRODUCTION

One of the significant challenges and opportunities in computer-assisted healthcare is how to improve diagnosis and treatment of patients. With respect to the past, diagnosis and treatment have become more and more difficult because of two major reasons: a global society where people, goods, and, consequently, diseases travel very easily; and a growing awareness and emergence of very rare "orphan" pathologies that affect an extremely small percentage of the population. Physicians are very often required to diagnose and correctly treat diseases that they are not familiar with. Especially in this context, but also with more frequent pathologies, computer assistance can become an essential tool in healthcare, especially if it supports all relevant information such as protocols, exceptions, best practices, condition-specific guidelines, and so on, in addition to diagnosis.

BACKGROUND

Medical diagnostic systems have a relatively long history with prototype systems developed as early as the '70s. The emphasis was on Artificial Intelligence (AI) techniques with an evolution from systems in which diagnostic knowledge from experts was captured in the form of empirical classification rules (Buchanan & Shortliffe, 1984) to theoretical and model-based approaches, such as set-covering (Reggia, Nau, & Wang, 1983), abductive diagnosis (Console, Theseider Dupré, & Torasso, 1989), and consistency-based diagnosis (de Kleer, Mackworth, & Reiter, 1992).

Regardless of the theory of diagnosis used, a dramatic limitation of most diagnostic systems is their system-centric, rather than user-centric approach. This implies a master-slave relationship in which the system is in charge of diagnosis, and the user is used to supply the system with observations. This architecture limits the application of diagnostic systems in areas such as healthcare, where highly skilled professional users do not readily accept a subordinate role. In addition, not taking users into account is one of the known causes of failure in knowledge-based systems (Brézillon, 1999). Finally, the cost of building and maintaining a diagnostic system based on AI techniques may very well be so large as to be impractical.

Alternatively, the diagnostic problem can be recast in terms of information access, rather than in terms of artificial intelligence. Complex AI architectures can be replaced by a collection of electronic texts that describe pathologies, which is searched by traditional techniques. Although this approach greatly simplifies knowledge base creation and maintenance, the limitations of information retrieval have been known for some time (Blair & Maron, 1985), and they indicate that locating the right information may be quite hard, and the result is usually not exhaustive.

More recently, systems based on hypertext technology have been proposed. An example is OncoDoc, a system for the assisted selection of clinical practice guidelines for cancer treatment (Bouaud et al., 1998; Séroussi et al., 2001). OncoDoc is an interactive system that encodes domain knowledge in the form of a decision tree implemented through a number of pages linked by hypertext links: the physician is presented with a sequence of choices that lead her to the guideline to be applied. Systems based on decision trees are less system-centric than conventional AI diagnostic systems, but user interaction is still quite rigid and follows predefined paths. Creation and maintenance are expensive, and the addition of a single new pathology may well disrupt the entire structure, and consequently user familiarity with the system.

DYNAMIC TAXONOMIES

In order to overcome the problems in current diagnostic systems, we propose a user-centric architecture in which the diagnostic system guides the user in exploring and systematically reduce the number of candidate pathologies until their number is sufficiently small for manual inspection. This architecture is based on a new information access paradigm represented by dynamic taxonomies (Sacco, 1987, 2000; also called *faceted search systems*), that focuses on exploratory, guided searches, rather than on traditional retrieval-based on precise specifications. Emphasis is placed on the user, on interactivity and on transparent operations, so that the main causes of failure in practical applications of diagnostic systems are removed. In addition, dynamic taxonomies easily support dynamic knowledge "encyclopedias," which include not only pathologies, but also emergency alerts, guidelines, and, in general, all relevant information. Since information items need not be textual, photos, x-rays, and so on can be easily accommodated.

Dynamic taxonomies are a general knowledge management model for complex, heterogeneous information bases. The intension is a taxonomy that does not require any other relationships in addition to subsumptions (e.g., IS-A and PART-OF relationships). Dynamic taxonomies require a multidimensional classification (i.e., items are classified under several topics at any level of abstraction, as appropriate). A concept C is just a label that identifies a set of items (items(C)), the set of the items classified directly under C, or under any of C's descendants (i.e., the *deep extension* of C).

This set-oriented approach has two important consequences. First, logical operations (and, or, not) on concepts can be performed by the corresponding set operations on their extension. Second, dynamic taxonomies can infer all the concepts related to a given concept C, which represent the conceptual summary of C. Concept relationships other than subsumptions are inferred by empirical evidence, through the extension only, according to the following *extensional inference rule*: two concepts A and B are related iff there is at least one item D in the infobase which is classified at the same time under A (or under one of A's descendants) and under B (or under one of B's descendants). For example, we can infer a (unnamed) relationship between Raphael and Rome, if an item that is classified under Michelangelo and Rome exists in the infobase. At the same time, since Rome is a descendant of Italy, also a relationship between Raphael and Italy can be inferred.

Dynamic taxonomies can be used to browse and explore the infobase in the following way. The user is initially presented with a tree representation of the initial taxonomy for the entire infobase. Each concept label has also a count of all the items classified under it (i.e., the cardinality of items(C) for all Cs). The initial user focus F is the universe (i.e., all the items in the infobase). In the simplest case, the user can then select a concept C in the taxonomy and zoom over it. This operation changes the current state as follows. First, concept C is used to refine the current focus F, which becomes $F \cap items(C)$; items not in the focus are discarded. Second, the tree representation of the taxonomy is modified in order to summarize the new focus. This means that all and only the concepts related to F are retained, and the count for each retained concept C' is updated to reflect the number of items in the focus F that are classified under C'. The reduced taxonomy is a conceptual summary of the set of documents identified by F, exactly as the original taxonomy is a conceptual summary of the universe.

The retrieval process can be seen as an iterative thinning of the information base: the user selects a focus, which restricts (thins out) the information base by discarding all the items not in the current focus. Only the concepts used to classify the items in the focus (and, because of subsumptions, their ancestors) are retained. These concepts, which summarize the current focus, are those and only those concepts that can be used for further refinements. From the human computer interaction point of view, the user is effectively guided to reach his goal, by a clear and consistent listing of all possible alternatives. Though guided, exploration is unconstrained because the user can freely focus on any concept, with the exception of those concepts not in the reduced taxonomy. These concepts are not related to the current focus, and focusing on any of them would produce a (useless) empty result.

The advantages of dynamic taxonomies over traditional access methods are dramatic in terms of convergence of exploratory patterns (Sacco, 2006), and in terms of human factors (Yee, 2003). Dynamic taxonomies have been successfully applied to a number of very diverse information access problems (Sacco, 2005a) and efficient implementations exist (Sacco, 1998).

DIAGNOSTIC ASSISTANCE THROUGH DYNAMIC TAXONOMIES

A dynamic taxonomy is used to recast the diagnostic problem in terms of guided information access to items that represent pathological situations, rather than in terms of logic manipulation. Each item is classified

Related Content