Interactive Image Cases for Medical Problem Solving

Erika Rogers and Barb Ericson

ABSTRACT

The development of an intelligent assistance agent for diagnostic radiology faces a number of challenges: e.g., how to gain access to the knowledge used in such visual problem solving tasks, how to operationalize it for further processing, and how to present it in the most useful manner for diagnosis. Our previous work in this domain placed an emphasis on the first two issues, and has shown that both perception and problem solving need to be supported in these types of tasks. However, the interface to the underlying intelligence in our first prototypes did not reflect how medical practitioners physically interact with images in the context of their work. This paper describes current work in progress on an image-centered approach to the agent's interface design. This approach will allow users to interact dynamically with the image, and to store these interactions for future retrieval.

1. INTRODUCTION

In many types of medical diagnosis and treatment, a patient’s radiological image or set of images forms a key knowledge component which is utilized by a variety of practitioners working on the patient’s case. Each practitioner will look at the same image in a somewhat different manner, and for a different purpose. For example, a radiologist will examine the image(s) and prepare a report which describes any abnormalities seen, and which may establish, confirm or deny diagnostic hypotheses about the patient’s condition. Surgeons will look at the image(s) to determine whether to perform an operation, and if so, where the path of the surgery should go (in this case the image will often be examined during the actual course of the surgery itself). Surgeons, radiologists and primary care physicians will use images for post-surgery assessment (e.g., was the bleeding successfully stopped, the aneurysm clipped, or the tumor removed). Nurses also examine the images to determine patient treatment based on functional concerns (e.g., if the patient has a tumor in a particular area of the brain, what physical and/or mental functions will be affected, and what kinds of behavior and capabilities should be expected pre- and post-surgery). Physicians frequently attend weekly in-house “conferences” where patient cases and treatments are presented to and discussed with colleagues. In these types of group discussions, the patient’s medical images are typically used as the major focus of attention.

The domain knowledge associated with these different usages is prodigious, and usually implicit – the radiology report is often the only formalized summary of what is observed in the image, and this information is usually a textual description which may or may not be electronically captured. Typically patient information, including one or more sets of film images, is collected in a physical file, which then travels as needed through the hospital system.

2. PREVIOUS WORK

There are numerous approaches to providing decision support for diagnostic radiology (e.g., [4, 5]), as well as an increasing number of world wide web resources containing sample cases and other support materials (e.g., [1, 2]). In contrast with these, our work to date on the development of an intelligent assistance agent for diagnostic radiology has focused primarily on studying underlying cognitive links between perception and problem solving, which enable humans to perform complex tasks such as interpretation and diagnosis from an image representation [6, 8]. Aspects of this cognitive model have been embedded in a blackboard-style architecture, which supports the opportunistic type of problem solving characteristic of this domain, and a prototype system called VIA-RAD (Visual Interaction Assistant for Radiology) was implemented and tested in a small study [7]. The original program was designed to run on a two-monitor system as shown in Fig. 1. This consisted of a high-resolution Barco monitor hooked up to a Pixar computer for image display, together with a Sun workstation which was used for knowledge processing and diagnostic menu display. This configuration was chosen for several rea-
reasons: i) it was expected that the very high display resolution would minimize the impact of digitization on the diagnostic process; ii) it was the same system on which the original cognitive data had been collected, and therefore it was felt that results would not be inadvertently affected by a change in display; iii) the underlying blackboard model was organized into logical panels of information (compatible with the cognitive model), and this allowed perceptual and problem solving components of the system to be mapped onto individual displays; iv) the location of the system in the hospital ensured convenient access for the subjects of our study, who were primarily radiology residents.

While this original system suffered from numerous limitations, including restricted memory, black and white monitor, slow mouse interaction, and proprietary image display routines, the comments of the study participants showed an interest and enthusiasm for the knowledge-based enhancements as well as the diagnostic problem solving assistance that encouraged further development. Our experience showed medical practitioners to be very open to technological innovation, and willing to consider its usage if it would help them perform more effectively.

VIA-RAD Blackboard Panels and Interface

In the original VIA-RAD system, the knowledge processing information was contained in four main blackboard panels, described below, and illustrated with an example screen shot in Fig. 2.

**Context Panel.** This area contains information that is known about the overall problem context. In the domain of radiology, this is where knowledge about the type of image and the particular anatomical area under consideration is kept (e.g., chest x-ray, brain mri, etc.). Specific objects or landmarks in particular configurations are considered standard or normal, and the presence or absence of such objects and their normal or abnormal classification constitute contextual events posted to this panel.

Patient-specific information is also posted in the current context panel. However, in the original system, since patient information was limited to age, sex, and occasionally, a very abbreviated form of symptoms, this part of the context was de-emphasized. With recent developments in electronic patient record technology, it is expected that this part of the current context will become much more important, allowing the intelligent assistant to retrieve and display previous patient data and other relevant information directly from the hospital record.

**Perceptual Panel.** The intelligent system obtains perceptual input about features in the image in a dynamic manner, and posts this information to the perceptual panel. Particular values for features are used by the problem solving knowledge sources to support and/or to rule out diagnostic hypotheses, adjusting belief levels accordingly. While the user is the primary source of perceptual input in the original system, extensions to incorporate automated image analysis techniques can be easily incorporated.

**Hypothesis Panel.** This panel contains the current hypotheses that constitute the partial (or complete) solutions that are evolving as a result of the problem solving activity. It is divided into two subpanels, corresponding to the two types of hypotheses that must be considered: Visual Hypotheses, which reflect what is currently known about abnormal or unexpected objects in the image (these correspond to Findings), and Reasoning Hypotheses, which constitute explanations of those objects, or collections of objects (these correspond to Diagnoses).

The main emphasis of the VIA-RAD system has always been on the visual hypotheses or findings, since in this domain, the characterization of what is in the image is considered to be the highest priority. Since the image to be interpreted is a representation of the real patient’s body, there is some uncertainty involved. Thus the findings consist of labeled objects or configurations with associated confidence factors. In the absence of plausible explanations, these findings constitute acceptable partial solutions.

The visual hypotheses (often combined with domain knowledge) serve the reasoning hypotheses (i.e., diagnoses) in either triggering or supporting capacities. The latter represent a deeper level of problem solving activity and progress toward a solution. These hypotheses have belief ratings associated with them that are a combination of the user’s assessment of their validity and the system’s own uncertainty management policies. The top-down relationship between reasoning hypotheses and visual hypotheses takes the form of expectations, which, if matched by observations, assist the program in converging upon a solution.

**Attention Panel.** This panel is the locus of the visual
focus-of-attention mechanism. It is also partitioned into two subpanels. The first contains Attention Directives aimed at the perceptual process of the user, while the second displays the image data. The image is considered to be part of the blackboard since it, too, is a dynamic part of the problem solving activity. The system directs the user’s attention not only by textual suggestions, but by automatically executing appropriate image enhancements designed to facilitate the user’s perception of the feature(s) in question. The user, in this initial system design, did not interact directly with the image itself, but indicated areas of interest through the menu-based selections.

The VIA-RAD system relies on the user for information about what is seen in the image, and also for an assessment of its proposed hypotheses. The assistance it offers in return is meant to aid both the perceptual and the problem-solving efforts of the user. The original user interface, considered to be a part of the control architecture, was divided into two parts: the Logical User View, which controlled how much of the blackboard was visible to the user, and the Presentation Manager, which controlled the form of the interface presented to the user (e.g., menus, graphics, direct manipulation windows, etc.). The user, as a key participant in the problem solving process, was expected to read from and write to the blackboard, similar to other knowledge sources which were, however, purely under system control. This was designed to follow the principles of cooperative decision aiding as described by [9].

The logical user view was expected to handle the amount and nature of blackboard information revealed to the user based on two different criteria: i) experience level of the user (e.g., a novice would see more details than an expert), and ii) the purpose of the system (e.g., a research tool, a training tool, or a clinical tool. In this way, the logical user view component of the user interface would allow the system to be adapted for various purposes without compromising its basic problem-solving approach.

On the other hand, the presentation manager was designed to provide the actual human-machine interface of the system through a displayed representation of the logical user view. This could take a number of forms including menus, icons, graphics, and/or direct manipulation windows, and was expected to extend to audio as well as other visual mechanisms. The presentation manager was considered to be responsible for how the user would be informed about changes to the blackboard made by the system, and also for enabling the user, in turn, to post information to the blackboard in the capacity of a knowledge source and decision-maker.
The User Interface component, therefore, had a dual role in the initial VIA-RAD design. On the one hand, the system was able to accommodate and direct the user's opportunistic behavior within the framework of the problem-solving session. On the other hand, changes to the blackboard that were effected by the user were monitored by the User Interface, and knowledge sources whose preconditions were met were then activated. Control was then passed to the VIA system so that it could make some contribution to the solution, if possible.

VIA-RAD II

A re-implementation of the VIA-RAD system using a commercial blackboard software package (GBB), began the process of reviewing the entire design of the intelligent agent. In the GBB system, events serve as an interface between blackboard database operations and the control shell, and are either predefined events that occur when the state of the blackboard database changes (e.g., a frame is instantiated on a panel), or events that occur for user-defined purposes. Events are used to trigger user interface functions as well as the knowledge sources which perform the cooperative problem solving. The control shell in this version of the prototype is based on a precondition/action model in which the activation of a triggered knowledge source is determined by a precondition function, and the execution order of the resulting action function is controlled by an execution rating returned by the precondition function [3].

Although these changes gave more power to the underlying intelligence, the only adjustment to the interface was a consolidation of the menu interaction and the image display to a single-screen format, as shown in Fig. 3.

3. INTERACTIVE IMAGE CASES

In the current phase of the project, we are addressing the interface and knowledge acquisition aspects of the problem. Although our underlying approach to provide both perceptual and problem solving assistance appears to be sound, and has been supported by preliminary test results, the original interface designs did not take into account many of the results we obtained from our earlier and subsequent cognitive studies. A consideration was whether the separation of the perceptual component from the rest of the system's problem solving behavior through the two-monitor display might interfere with the diagnostic process since it forced the physician to shift attention between the image and the selection panels. Putting everything onto one screen did not really address the fact that our external interface does not correspond closely to the way radiology is conducted, and is also not conducive to adapting the interface for other medical personnel to interact with the image.

Typically, a patient's image film is placed on a light box and becomes the center of attention for the discussion of the patient's case. Parts of the image are highlighted either by pointing, or by marking; sometimes a magnifying glass and/or a ruler is used, and generally the discussion is fairly concise and to the point. If a radiologist is examining the image for diagnostic purposes, the examination may be more comprehensive, whereas if the image is being used for peer presentation during a weekly conference, the main points of the radiological findings may be presented as a punctuation to the discussion of treatment and outcomes. With advances in multimedia interfaces, graphical windowing environments and image processing techniques, it is clear that an image-centered approach to radiological knowledge navigation and knowledge acquisition will make the intelligent assistance agent a more powerful decision support tool.

Current work on the new interface design of this agent incorporates interactive image annotation to link the current domain knowledge and the patient context to a set of visual overlays for the radiological image. Each overlay will represent an assessment by a particular type of practitioner (e.g., radiologist, surgeon, nurse, etc.) and display of multiple overlays should be possible if desired. The annotations will be adaptively constructed from a collection of relevant concepts (e.g., landmarks, findings, features and diagnoses), relevant tools (e.g., ruler, region-of-interest, etc.), and additional comments areas. These annotations may also be, in turn, hyperlinked to more extensive medical background knowledge or to other similar cases. However, to build a large enough collection of exemplars for case-based reasoning techniques to be exploited, we need to make it easy for the practitioners themselves to provide the knowledge about what they see, and what they think about what is seen. Therefore the following goals have been identified for VIA-RAD III:

- to capture the medical expert's assessment of the image in a manner that is intuitive to the expert, and is directly linked to the image itself;
- to utilize this input as knowledge acquisition for further intelligent processing - both for visual interaction assistance and for future case-based retrieval and processing;
- to adapt the interface and the domain knowledge presented to the particular type of practitioner using the system (e.g., primary care physician, radiologist, surgeon, nurse, etc.);
- to store the information as visual cases, which other users can retrieve as overlays to the image for either treatment purposes or training purposes; these cases can also be further indexed for more extensive case-based retrieval and processing;
- to make a tool which can be of immediate stand-alone use to the medical experts themselves, in addition to its value as a knowledge acquisition tool for our project.

The new design puts the image at the center of the screen, with knowledge navigation surrounding it. Landmark classification is still the initiating diagnostic activity, but in this version, an “abnormal” classification triggers several different behaviors: a) the agent presents a region of interest tool which allows the radiologist to mark the location of the abnormal finding; b) depending on the landmark selected, more specific localization is supported through a selection of illustrated parts; and c) the next step of labeling the finding itself is presented as a concept hierarchy which allows the radiologist to choose the appropriate label for the finding and drag it to annotate the selected region of interest on the image. The latter capability is designed to adapt more closely to cognitive results obtained in the initial study of radiological diagnosis. In this study, it was found that findings (i.e., abnormalities in the image) were characterized at different levels of specificity, and that the particular words used gave not only an indication of where the physician was in the problem solving process, but also suggested what type of further supporting information might be needed. For example, the statement “there is a density” really only means “I have detected a light area”, without conveying any further semantic information regarding what this object might represent. On the other hand, a term such as “there is a mass” suggests that certain associated features such as size, shape and edges should be investigated, whereas a statement about a “malignant tumor” indicates default values for specific features, (e.g., large, round mass with lobulated edges) as well as diagnostic hypotheses. By allowing the user to choose from the hierarchy, the agent obtains immediate knowledge about what additional knowledge and support might be needed. General image enhancements as well as those specifically related to the current level of problem solving in the region of interest are provided as selectable thumbnail images – the user can view enlargements of these when desired, while still maintaining the context of the current image. A new region of interest would automatically trigger a different set of enhancements to become available. An example of the interface layout is shown in Fig. 4.

4. CONCLUSIONS

With the introduction of high-speed networks, high resolution computers, internet connectivity and advanced imaging techniques, the changing face of healthcare is creating a demand for more effective decision support
tools to help medical practitioners cope with the information overload while providing high quality and cost-effective patient care. These tools must not only demonstrate knowledge-based behavior to assist in medical problem solving, but they must also provide intuitive, interactive and intelligent interfaces that allow the users to perform their tasks efficiently and effectively. It is expected that the new directions taken in VIA-RAD’s interface design, which tie the underlying cognitively-based intelligence more closely to the image-based work context of the users will provide a more flexible and extendible approach to radiological diagnostic support.

5. ACKNOWLEDGMENTS

This research is currently supported by NSF Grant IIS-9896178. Special thanks go to Eric Sparling and Michael Brown for their assistance.

6. REFERENCES


