Cognitive Engineering for Intelligent Man-Machine Systems

Erika Rogers

Abstract—This paper discusses a methodology for cognitive engineering which provides a framework for the *cognitive studies* → *intelligent system development* → *cognitive studies* cycle. The methodology has been developed to address the problem of providing knowledge-based cooperative assistance in visual reasoning tasks, and is discussed in terms of its impact on three different application domains: diagnostic radiology, geographical information systems, and tele-assisted robot supervision.

I. INTRODUCTION

In many complex problem-solving applications, knowledge-based computer systems are unable to effectively handle degraded information or novel situations with ambiguous conditions [8]. This means that designers of such intelligent systems must try to keep the human "in the loop", and there is a growing interest in using automation to amplify human skills rather than to replace them (e.g., [3, 4, 8, 10]). Hoc has pointed out the need for "expert systems which share common representations with the user", leading to what are known as cooperative problem-solving systems [2]. In order to achieve this goal, cognitive science research must be coupled with complex system design so that interaction mechanisms may be effectively integrated with underlying domain knowledge [1].

However, methodologies for attaining this integration frequently employ a one-way relationship between the cognitive studies and the computer system development, with an emphasis on the latter. One approach is to extensively study human performance in solving certain kinds of problems, then to apply this theory to the development of a system which can replace some of the "faulty" aspects of human reasoning. Further testing and evaluation emphasizes improvement of computer system performance capabilities. A different approach begins with the goal of developing a computer system which addresses certain kinds of problems, and cognitively-oriented studies are conducted after its development to test and improve usability of the system, and user satisfaction with it. This, again, leads primarily to computer system refinement, rather than improved human performance in the problem-solving task.

Figure 1: Development Cycles.

Relatively few paradigms have emerged to date which not only result in functioning application programs, but which also allow feedback from these programs in order to further develop the underlying cognitive theories. This paper discusses a methodology for cognitive engineering which provides a framework for such a development cycle, as shown in Fig. 1. This methodology has been developed to address the problem of providing knowledge-based cooperative assistance in visual reasoning tasks, where understanding and interpretation of visual images are key elements of the decision-making process. The goal is to facilitate and stimulate human reasoning capabilities in such tasks. In this case, the "intelligence" consists of knowing what type of assistance is needed and when it may be cognitively useful to afford it. The system can then automatically invoke relevant visualization tools at the appropriate stages in the problem-solving task in order to improve performance by a) reducing some cognitive load (e.g., little or no tool selection needed), and b) selecting performance-enhancing rather than simply appearance-enhancing algorithms.

II. VIA OVERVIEW

In Fig. 2, are shown the basic stages of the methodology used to develop the Visual Interaction Assistant (VIA). There were three primary products of this effort: i) a three-stage data collection and analysis methodology which includes an encoding scheme for protocol analysis, ii) a cognitive model of the interaction between perception and problem-solving, and iii) a blackboard-based system design.

Figure 2: Stages of the Methodology.

Figure 3: Environment Analysis.

A. Data Collection and Analysis

Understanding how humans process information, and what type of assistance may be needed, requires a deeper study of how people do work than has been typical to date. In the VIA project, this has been accomplished by identifying three levels of data collection and analysis, called Environment Analysis, Protocol Analysis and Contextual Analysis, respectively.

The first stage of data collection, also called Environment Analysis, is where information is gathered about how the task is accomplished. A guiding principle is that how people *think* they perform a task does not always correspond with how they *actually* do it. Therefore, to reveal as many dimensions of the task as possible, it is important to collect data from a number of sources such as external observation, performance studies, and verbal protocol studies. This leads to a telescoping series of activities, each one providing results which form a basis for subsequent experiments, as illustrated in Fig. 3.

The purpose of the Protocol Analysis stage is to examine in detail the data obtained in the experiments, and to identify and label relevant categories and concepts in the verbal statements of the subjects. This is done through

Figure 4: Protocol Analysis.

Figure 5: Contextual Analysis.

the vehicle of an Encoding Scheme, as shown in Fig. 4. Application domains with an extensive specialized vocabulary pose a particular challenge to the development of a general encoding scheme. In the VIA methodology, this is handled by a two-step encoding process, where protocol statements are pre-processed using domain-specific labels, followed by a more general descriptive and cognitive concept analysis.

The final stage of Contextual Analysis is then applied to determine whether any clusters or patterns of concepts can be discerned. Generally, the encoded statements are examined from the perspective of important overall task-related, time-related and/or experience-related patterns. In the analysis leading to VIA, only the time and task dimensions were explored as shown in Fig. 5, although, in future, comparisons between expert and novice performance will also be examined.

This type of extensive data collection and analysis provides the cognitive foundation for the development of a theory of the interaction between perception and problem-solving in a visual reasoning task.

B. Model of Visual Interaction

In these kinds of applications, humans are much better than computers at perceiving information in the images, and using this information to achieve solutions. This is especially the case in real-world dynamic situations, where the nature of the challenges cannot easily be predicted. Therefore the goal of the VIA system development is to facilitate and enhance these capabilities, rather than to reproduce them. That is, the computer should be used to do image, data, and knowledge processing in a way that is aligned with an understanding of the user's visual information processing in the task. This means

that, for effective cooperation, the system should contain knowledge about descriptive features, levels of abstraction, context, focus of attention and expectation. In particular, these issues need to be expressed in the context of their contribution to the avoidance and/or recovery from errors and oversights. All of this information must be presented in a coherent and consistent framework that allows the system to "understand" how to provide cognitively effective assistance. This is achieved through the model developed in [6]. It is based on the information-processing paradigm of cognition, and proposes a *mediating process* between perception and problem-solving, called the visual interaction process (VIP).

The VIP does not usurp any of the functions of either problem-solving or perception. Rather it acts as a transformer at those points in the visual reasoning task where the problem-solving process requests perceptual input, and where the perceptual process is delivering such relevant information. The primary functions of this process are identified as hypothesis management and attention direction, and the working memory structures which support these activities are described in terms of two conceptual buffers and a visual context store. These symbolic structures represent how working memory may be organized to accommodate the information flow during the visual reasoning task. The purpose of this model is to serve as a bridge between understanding the cognitive activities of the humans, and building an intelligent system which supports these activities.

This model is the beginning of an iterative, evolutionary process, aimed at understanding how perceptual information is utilized by higher cognitive processes in order to solve problems. Furthermore, it provides a framework for task-related knowledge about the user, especially with respect to where assistance might be needed in the visual interaction cycle. In this sense, the model represents a *practical* theory which not only tries to account for experimental data results (both our own, and that of other researchers), but which also lends itself to incorporation in the design of a cooperative computerized assistant.

C. System Design

The general VIA system design closely resembles aspects of the model of visual interaction, and consists of a blackboard-style architecture which allows incorporation of the user, the image(s), the knowledge base, and the functional aspects of the program into an integrated modular system. The working memory of the model suggests a logical partitioning of the blackboard that facilitates the corresponding types of information transfer. These partitions or panels are divided into four main categories: the Context Panel contains information that is known about the overall problem context, including expected objects or landmarks in particular configurations that are considered standard or normal; the Hypothesis Panel contains the current hypotheses that constitute the

partial (or complete) solutions that are evolving as a result of the problem-solving activity; the Perceptual Panel allows the system to dynamically obtain perceptual input about what is in the image; and the Attention Panel is the locus of the visual focus-of-attention mechanism.

The Hypothesis Panel is further divided into two subpanels: one contains *Visual Hypotheses*, which reflect what is currently known about abnormal or unexpected objects in the image, and the other contains *Reasoning Hypotheses*, which correspond to explanations of those objects, or collections of objects. The Attention Panel is also partitioned into two subpanels. The first contains attention directives aimed at the perceptual process of the user, while the second consists of the image itself, which is considered to be a dynamic part of the problem-solving activity. The system directs the user's attention not only by textual suggestions, but by executing appropriate image enhancements selected to facilitate the user's perception of the features in question. This general design provides a relatively domain-independent framework which can be used as a template for building systems in a variety of applications where visual interaction is a component of the problem-solving task.

The overall control structure of the VIA system reflects the functionality of the visual interaction process of the model in terms of hypothesis management and attention direction. The former impacts the blackboard through the activities of hypothesis-related knowledge sources, while the latter is concerned with focussing attention by altering the image and/or presenting suggestions to the user of what to look at next. In order to pass control from the Hypothesis Manager to the Attention Director, a third control module is needed to select strategies. Since there may be several ways in which attention could be focussed in order to obtain perceptual information, the Strategy Selector makes this decision, based on current blackboard information. The user is conceptualized as an additional independent knowledge source, who reads from and writes to the blackboard in cooperation with the system's knowledge sources.

III. APPLICATION DOMAINS

Fig. 6 presents an overview of the VIA cognitive engineering methodology applied to three different application domains. The diagnostic radiology branch was used to establish the theoretical and architectural foundations for the initial VIA system described in the previous section. A preliminary prototype system, called VIA-RAD, was implemented and tested with radiology subjects, and feedback obtained from this study indicated a positive trend in performance as well as favorable impressions by the subjects [6]. However, the limitations of the first prototype were such that improvements in implementation must be carried out before more detailed evaluations can be conducted.

The second project, called VIA-GIS (Visual Interaction Assistance for Geographical Information Systems), involves decision-making tasks based on visual display of

Figure 6: VIA Methodology in Three Application Domains. (Solid filled boxes represent completed stages, solid unfilled boxes represent work-in-progress, and dashed boxes and lines represent future work.)

geographical database information. The cognitive data gathered in this domain will provide a basis for evaluation of the cognitive model of perception and problem-solving, and the refined model will then be used to design and develop an intelligent GIS prototype system.

The third project, called teleVIA, is concerned with providing intelligent assistance to a human operator in the supervision of robotic systems in hazardous or otherwise remote environments [7]. The role of the teleVIA system is primarily to filter and enhance data from the remote robot when it requires assistance from the local operator. In this work, there is not only cooperation between the human and the intelligent VIA system, but an additional computational agent must be included in the cooperative group. Therefore, the initial emphasis of teleVIA is on the extension of the general VIA system design to accommodate this configuration, and to handle the particular type of perceptual data generated by the robotic sensors.

In each of these three applications, there is a similar relationship between the human user and the environment which affords the perceptual information involved in the visual reasoning task. That is, the user does not deal directly with the actual external world environment; rather, the problem-solving is accomplished through interaction with a technologically-mediated task environment as shown in Fig. 7.

In the case of diagnostic radiology, this consists of the

Figure 7: Environmental Interaction

medical image(s) provided by the scanning devices of the hospital, which may be supplemented by some image manipulation capabilities. (Although a number of different modalities are available through current technology, the images are still typically static, black and white "snapshots" of the affected parts of the patient's body.) On the other hand, in geographical information systems, the user interacts with a software program which allows the selection and display of various attributes of the underlying geographical database(s). The display usually takes the form of visual maps in addition to text. Finally, the tele-assistance domain requires interaction of the human supervisor with a computational agent in the form of a semi-autonomous robot, which is engaged in exploration and/or remediation of a remote environment.

In all three applications, perceptual characteristics consist of image display and enhancement capabilities. Issues which must be addressed include dimensionality of the features and direction of attention. The problem-solving characteristics include contextual issues, the role of expectation, and levels of abstraction. In each application, the user is subject to errors and oversights, which must be identified and addressed if effective assistance is to be provided. All of these relationships and issues influence how the cognitive engineering methodology is applied, and may in turn, be themselves altered by the introduction of an intelligent cooperative computer system.

A. Diagnostic Radiology

In diagnostic radiology, the task consists of two parts: i) gathering information by looking at the image, and ii) deciding what's in the image, i.e., determining abnormalities and proposing diagnoses. The user relates to the task environment through these two activities, but does not cause any direct changes to that environment. This means that the outcome of the problem-solving depends solely on the quality of the image, and the capabilities of the user, requiring intensive training together with natural perceptual and problem-solving skills or talents.

In the task of examining computer-displayed chest x-ray images, it has been shown [6] that three kinds of oversights can occur: detection errors, where a relevant abnormality is not seen or noticed; identification errors, where an abnormality is seen, but incorrectly labelled; and/or integration errors, where an abnormality is noticed and correctly labelled, but not used in the genera-

tion of diagnostic hypotheses. The first problem can be addressed if the environment (i.e., the image) is allowed to change while the diagnostic process is underway. This can be accomplished by the provision of image processing and enhancement tools, now available on most commercial medical imaging systems. The integration type of oversight may be remedied to some extent by the addition of an expert decision support system, which proposes hypotheses and solutions based on current evidence. In both of these cases, the cognitive load of the user is increased since there is the added burden of learning how to effectively use these computational tools, as well as solving the original problem of diagnosis.

The VIA-RAD system, on the other hand, attempts to address the range of oversights by providing an integrated problem-solving environment. The system tracks the user's diagnostic reasoning and tries to assist at appropriate times by suggesting hypotheses and/or automatically enhancing the image.

B. Geographical Information Systems

The application of radiology and GIS are similar in terms of having a computational tool which affords the perceptual information to the user through a representation of the real world environment. On the other hand, the task itself is quite different in the two domains. In radiology, the user carries out a diagnostic task, where abnormalities in an image must be detected, labelled, and incorporated into diagnostic hypotheses. In GIS, the task is to explore the information encoded into a spatially-indexed database. The user is not necessarily seeking anomalies to some "normal" state of the environmental image, but rather seeks to build the image or map itself, supplemented by additional details contained in the fields of the database. Once the relevant visual and textual representation is obtained, the user may perform further decision-making tasks based on the displayed information.

One challenging aspect of this task is that the user may not have a mental or contextual model of the map that is going to be generated. He/she may have some concept of the approximate shape and size of a region to be considered, but to a great extent, the user must trust the reliability of the software system and the accuracy of the underlying database to produce an adequate representation for extracting the desired information.

Preliminary observations of GIS users suggest that difficulties are often encountered with trying to master the software to produce the appropriate representation needed for the task. Therefore, although there are perceptual and problem-solving aspects to this task as well, the nature of the cooperative assistance required is different from that of diagnosis. Rather than generate hypotheses about what's in the image, it is more important to assist the generation of the images or maps themselves. This means that the knowledge encoded into the intelligent system must include not only system commands

and database characteristics, but this information must be indexed according to the various types of user tasks, and the visualization capabilities needed to perform those tasks.

C. Tele-assisted Robotics

The domain of tele-assisted robotics differs substantially from the previous two in that the user interacts with a computational *agent* rather than a *tool*. The semi-autonomous robot has the task of exploring and/or working in a remote real world environment. However, because of its limited sensing capabilities, it may require the assistance of a local human operator when sensing failures cannot be resolved. In this case, the task of the human user falls under the general category of diagnosis: i.e., looking at the data provided by the robot, and attempting to determine what has caused the failure and how to resolve the problem.

In an unaided system, the local task environment of the user presents numerous cognitive challenges: direct querying of the remote robot may be too slow; transmission of all related data may include unnecessary information; the sensor data itself may be in formats that are difficult for humans to understand and interpret. The display may contain different types of images obtained from various sensors involved in the failure, as well as some textual information on the hypotheses generated and tested through the robot's autonomous exception-handling mechanism. Any of this information could be faulty or misleading, and the user must quickly determine what is relevant, what it means, and what to tell the robot. Multiple robots operating on the site would typically require multiple local supervisors to monitor their activities.

The purpose of the teleVIA system is, therefore, to act as an intermediary between the local supervisor and the robot, requesting relevant data in case of failure, automatically enhancing that data where appropriate and providing additional domain knowledge in order to diagnose and repair the failure. The reduction in cognitive load may make it more feasible for one local operator to supervise several robots, potentially recruiting them to provide additional data to help the troubled robot.

V. DISCUSSION

The VIA system can be thought of as a computational assistant, (more than a tool, but less than an agent), in a relationship with the human user as well as the task environment. This leads, on the one hand, to a joint cognitive system viewpoint, where the human is coupled with the computational system to act on the task environment [10]. A different approach, however, leads to a joint task environment perspective, where the human is situated in an environment which includes the computational system (e.g., [9]). In the VIA approach, both of these views are accommodated and integrated into a larger, dynamic problem-solving perspective, as

Figure 8: Dynamic Problem-solving Perspective

shown in Fig. 8. The management of hypotheses and selection of strategies are founded on an information-processing type of model. On the other hand, by directing attention, VIA reaches beyond the user to change the environment itself through image enhancements. In this way, the viewpoint can change according to the current needs of the visual reasoning activity. This type of cognitive flexibility is felt to be more characteristic of a truly cooperative system—sometimes the user and the VIA system are coupled in considering the environment, the current problem-solving context, and the application of additional long-term knowledge. At other times, the VIA system is coupled with the environment, changing it through appropriate image and other visual enhancements in order to allow the user to fully exploit his/her perceptual capabilities.

In order to couple the computational system with the human in a joint cognitive system, it is important to have an understanding of the cognitive processes of the user in the execution of the task. On the other hand, the joint task environment viewpoint requires a thorough task analysis. In the cognitive engineering methodology described above, both of these requirements are met through the initial data collection, analysis and cognitive modelling stages. The flexibility to dynamically change perspective is introduced in the system design stage. At this time, a blackboard-style architecture serves this purpose. However, as further implementation is carried out, a different design paradigm may emerge as a viable alternative. Once a prototype system is implemented, further data collection leads not only to testing and evaluation of the computational system itself, but also of the two joint perspectives. This feedback is then used to modify and refine the earlier cognitive and domain theories.

VI. CONCLUSION

In each of the three applications presented: diagnostic radiology, geographical information systems and tele-assisted robotics, it has been shown that an intelligent cooperative assistant would be helpful to overcome some of the difficulties faced by users in these complex problem-solving situations. Similarities such as the user's relationship to the task environment as well as basic perceptual and problem-solving issues which arise from the visual reasoning task have been identified, suggesting that a

common cognitive engineering methodology may be applied to conduct appropriate cognitive studies as well as to design the system itself.

The VIA methodology is one approach to this challenge, providing a platform for data collection and analysis, cognitive modelling, system development, and testing and evaluation. It is expected that the application of this methodology to the different domains will lead to insights regarding not only design and implementation of intelligent cooperative assistance systems, but also an evolving cognitive theory of how people utilize perceptual information in visual problem-solving tasks.

References

- [1] Fischer, G. and Reeves, "Beyond Intelligent Interfaces: Exploring, Analyzing, and Creating Success Models of Cooperative Problem Solving". *Journal of Applied Intelligence* 1, Boston: Kluwer, 1992, pp. 311-332.
- [2] Hoc, J.M., "Cognitive approaches to process control". In G. Tiberghien (Ed.), *Advances in Cognitive Science, Vol. 2: Theory and Applications*. Chichester: Horwood, 1989, pp. 178-202.
- [3] Hollnagel, E., Mancini, G. and Woods, D.D. (Eds.) *Intelligent Decision Support in Process Environments*. Berlin: Springer-Verlag, 1985.
- [4] Holtzman, S. *Intelligent Decision Systems*. Reading, MA: Addison-Wesley, 1989.
- [5] Rogers, E., "A Cognitive Foundation for Intelligent Computer Display Systems". In *Advances in Artificial Intelligence - Theory and Application*, Proc. of 6th Int. Conference on Systems Research, Informatics and Cybernetics. Windsor, Canada: Int. Inst. for Advanced Studies in Systems Research and Cybernetics, 1992, pp. 13-18.
- [6] Rogers, E., *Visual Interaction: A Link Between Perception and Problem-Solving*. Ph.D. Dissertation, Tech Report No. GIT-CC-92/59, Georgia Institute of Technology, Atlanta, GA, 1992.
- [7] Rogers, E. and Murphy, R.R., "Tele-Assistance for Semi-Autonomous Robots". *Proc. of AIAA Conference on Intelligent Robots in Field, Factory, Service and Space*. NASA Conference Publication 3251, 1994, pp. 500-508.
- [8] Rubin, K.S., Jones, P.M. and Mitchell, C.M., "OFMspert: Inference of operator intentions in supervisory control using a blackboard architecture", *IEEE Transactions on Systems, Man, and Cybernetics* 18, No. 4, 1988, pp. 618-636.
- [9] Stolze, M., "The Workshop Perspective: Beyond Optimization of the Joint Man-Machine Cognitive System". Working Notes for AAAI 93 Fall Symposium on Human-Computer Collaboration: Reconciling Theory, Synthesizing Practice, 1993, pp. 113-118.
- [10] Woods, D.D., Cognitive Technologies: The Design of Joint Human-Machine Cognitive Systems. *The AI Magazine* 6, No. 4, 1986, pp. 86-92.