Cognitive cooperation through visual interaction

Erika Rogers

The purpose of the paper is to describe a cooperative system which was designed to provide intelligent assistance in the domain of diagnostic radiology. The approach taken includes the collection and analysis of human performance data, cognitive modelling of the interaction between perception and problem solving, incorporation of this knowledge into a blackboard-based architecture, implementation of a prototype system and testing and evaluation of the prototype. These stages are described, and then discussed in the context of issues particularly relevant to the study of human-machine cooperation.

Keywords: cooperative systems, intelligent assistance, man-machine cooperation

The enhancement of human performance in complex tasks is an issue which has long concerned researchers, particularly with respect to the role of automation. There is a growing emphasis on assisting decision makers in gaining insight into problems, rather than on merely making or recommending right answers\(^1\). Woods has discussed the advantages of a joint human-machine cognitive system architecture where the 'system' is defined as the combination of human and machine, and he poses a challenge to cognitive scientists to provide 'models, data and techniques to help designers build an effective configuration of human and machine elements\(^2\).

The research described in this paper takes up that challenge in the particular arena of visual problem solving. More specifically, there are many applications in which understanding and interpretation of visual images are inherent parts of the problem-solving process. Examples range from diagnostic radiology to geographical information systems to operator-assisted telerobotics. In order to provide intelligent cooperative assistance in such domains, we need to know something about how humans perform these tasks, when and where assistance would be most useful, and what type of assistance is needed. This knowledge must then be incorporated into a system design which allows both the user and the computer to contribute to the eventual attainment of a solution or partial solution.

To build effective human-machine cognitive systems, techniques and concepts are needed to identify the decision-making/problem-solving requirements in some domain\(^3\). Traditional expert systems have approached this problem by extracting relatively shallow knowledge from a small number of domain experts, and representing this knowledge in a verbal rule-based format. To build more cooperative systems, however, these techniques must be extended to obtain a deeper level of knowledge covering a broader range of experience in the domain. Furthermore, perceptual requirements must also be understood, in order to accommodate the image or spatial reasoning needed for domains where the decision-making process relies on visual interpretation. Thus, our research has three main goals:

- to study and understand the interaction between perception and problem solving in particular visual reasoning tasks,
- to develop a proficiency in the domain knowledge of the particular application, and to extend that knowledge,
- to incorporate these results into a computer system designed to assist and enhance user performance in the task under consideration.

Initial work has focused on the domain of diagnostic radiology, and the main emphasis of this paper is on the methodology used to develop a visual interaction assistant for radiology (VIA-RAD). At each stage, we have tried to balance design and implementation decisions in
accordance with the three goals stated above. The experience of producing VIA-RAD has led to some insights about designing cooperative systems based on visual interaction, and some of these issues are discussed in the last part of the paper.

DATA COLLECTION AND ANALYSIS

Our first objective was to study how radiologists perceive computer-displayed chest X-ray images, and how this information is used to solve diagnostic problems. Initially, observational data was gathered which served to acquaint the investigator with the methods and behaviors of radiologists in the normal routine of the hospital chest X-ray reading room. Challenging aspects of the task, important features of the environment, and an initial domain vocabulary were identified in the context of diagnosing film images.

In the subsequent experiments, however, the environment would be changed to that of a laboratory setting, and diagnosis was to be done from computer-displayed X-ray images rather than film. Therefore, it was felt that some transitional data should be collected to determine whether these changes would have any major adverse effects on how the recruited subjects performed their diagnostic task. A limited study was conducted which compared certain timing, image characteristics, and performance parameters between the diagnosis of film and computer-displayed images. These results are described in more detail in Reference 5, but the overall outcome was that no major problems were discovered which would impact our subjects' ability to work with the CRT images.

The final experiment involved the collection of concurrent think-aloud protocols while the subjects were examining and diagnosing a number of chest X-rays on CRT. This type of articulation data has been described by Ericsson and Simon as the closest reflection of the cognitive processes utilized during the execution of a task. The images chosen represented a variety of abnormalities, and the subjects ranged in expertise from second year radiology residency to thirty years' experience. For each case, the articulation data included verbal reports, a formal dictation and a questionnaire that provided some quantitative data as well as introspective comments.

Results

Extensive analyses of these data revealed a number of issues relevant to the interaction between perception and problem solving in the context of the diagnostic process. These issues have implications not only for the modelling process, but for the system design as well.

- **Levels of abstraction for findings:** The same abnormality in an image could be referred to in a variety of ways, ranging from simple perceptual terms (e.g. density) to very diagnostically specific terms (e.g. malignant tumor). The expression of such levels of abstraction can be considered an indicator of where the user is in the current problem-solving activity, and it is important that the knowledge base of the cooperative system reflect these layers of findings.
- **Context:** Evans and Gadd point out that declarative context concerns the organization of knowledge in the domain, while procedural context describes the structure of plans associated with the use of domain knowledge. In this application, the former includes knowledge of the particular anatomical region under consideration, together with the visual characteristics of the imaging modality. This information provides further guidance for concepts such as landmarks which are important to include in the knowledge base of the system. Indications of plan-like activities such as deliberate landmark search, secondary abnormality search and finding characterization provide support for procedural context.
- **Evidence used in reasoning:** The direction of reasoning can affect the way in which perceptual evidence is used to support the diagnostic activity. For example, bottom-up or data-driven reasoning was supported by
  - the use of secondary findings to elicit (generate) diagnostic hypotheses,
  - the use of features of primary findings to specialize labelling of primary findings,
  - the use of features of primary findings to trigger diagnostic hypotheses.

On the other hand, top-down or expectation-driven reasoning involved
  - the confirmation of expectation of secondary findings to support diagnostic hypotheses,
  - the use of features of primary findings to rule out competing findings and diagnostic hypotheses,
  - the use of features of primary findings to trigger new (alternative) diagnostic hypotheses.

From these results, it is clear that intelligent assistance must involve not only knowing what kind of evidence is needed, but why it is needed at that time.

- **Attention:** In order to assist the visual interaction process, it is important to determine where in the image the user is focusing attention, and for what purpose. In this particular task, two types of attention were emphasized: a relatively fast noticing and labelling of an abnormality, which was called immediate visual capture, together with a more deliberate serial search of the landmarks.
- **Expectation:** Expectations may be used by the problem solver to optimize plans for the gathering of information that will converge on a solution. In the task of diagnosing chest X-rays, the expectations are largely perceptual in nature. That is, most expectations have to do with anticipations about what
can be seen in the image. In this sense, they appear to be dual in nature, originating with some abstract statement of intent, but resulting in an act of looking. This suggests a structure which provides a connecting link between perception and problem solving.

- **Oversights and errors:** A further link between perception and problem solving was suggested by the different levels at which relevant information could be overlooked:
  
  o **Perceptual level:** This type of detection oversight occurred when the subject did not notice or see the abnormal object or feature at all.
  
  o **Identification level:** In this case, a labelling error occurred when the subject saw the abnormality in question, but labelled it incorrectly.
  
  o **Problem-solving level:** Difficulty with integration occurred when the subject saw and labelled the abnormality correctly, but failed to use this information in the generation of diagnostic hypotheses.

These results are also consistent with Blesser's three-level categorization of radiological errors, and provide further justification for the approach of developing a cooperative assistant that addresses aspects of the entire visual interaction spectrum.

These issues are all closely coupled in the visual reasoning task. Context sets the scene for a particular collection of declarative and procedural knowledge components to be retrieved from memory and brought to bear on the problem. This knowledge creates expectations of what the practitioner is likely to see, and plans to explore these expectations emerge, which then guide the attention process in deliberate search. However, there are often unexpected phenomena in the image, which seem to capture attention immediately, and cause currently active plans to be interrupted or abandoned in favour of new exploratory activity. Descriptive features are used to characterize abnormalities, which, in turn, are labelled at different levels of abstraction. In the interplay between these different issues, a pattern of interaction between perception and problem solving begins to emerge. Descriptive features can be said to lie closer to the perceptual side, while context seems to originate with more abstract thought related to problem solving. Expectations lie between these two poles, originating with problem solving, but resulting in acts of looking.

**MODELLING VISUAL INTERACTION**

Unfortunately, to date, there are still few practical theories about how different aspects of cognition interact when humans attempt complex tasks. In the particular case of visual reasoning, one needs to know about (visual) perception, about problem solving, and, most importantly, about how perception and problem solving exchange information. Perception needs to deliver information to the problem-solving process, and conversely, the problem-solving process has to communicate directions to the perceptual process (e.g. I need this type of information rather than that type). Therefore both bottom-up and top-down activities must be incorporated, and the internal representation in working memory should be able to accommodate knowledge from both sides: visual information delivered by the perceptual process (e.g. percepts which describe findings in the image), and decision-related knowledge based on the current state of the problem-solving process (e.g. which hypotheses are active and the types of information they need for evidence). The model must also coordinate different levels of plans. For example, a plan to pursue hypothesis-directed search versus data-driven search is at a different level of abstraction than the detailed plan for gathering the specific perceptual evidence required by a particular hypothesis.

The model developed in Reference 7 and shown in Figure 1 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 at which 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 involved in the visual reasoning task, and building an intelligent system which assists and enhances these activities.

**VIA-RAD SYSTEM DESIGN**

The representation of the system's problem-solving activities, and how it retrieves the relevant portion appropriately in response to user queries, are important
issues in the design of cooperative systems. Both in our own experimental results and those of Lesgold et al., there was evidence of opportunistic planning or problem solving. This implies that the user changes the direction of reasoning (e.g. forward or backward chaining) according to the current state of the solution space. Partial solutions are constructed at various levels of abstraction, and alternatives are generated or abandoned as various pieces of evidence are incorporated. This naturally led to consideration of the blackboard paradigm, used extensively in artificial intelligence research on problem solving. Applications of blackboard systems are often used to solve complex and ill-structured problems, where the solutions involve diverse sources of knowledge.

In the development of the VIA-RAD system, the main domain concepts are landmarks, findings, features and diagnoses, and they are related as shown in Figure 2. These concepts are also reflected in the panels of the VIA-RAD blackboard, shown in Figure 3.

Landmarks refer to anatomical objects in the body, such as hearts and lungs, which are commonly visible in the X-ray image. This panel provides the context for what the radiologists expect to see in the image. The second panel holds two kinds of hypotheses: (a) findings, which are the potential or actual abnormalities observed in the X-ray image, and (b) diagnoses, which are the diseases causing the findings and other symptoms. The features panel holds descriptive attributes of the findings which are noted by the radiologist (e.g. smooth edges of a mass, large size). Finally, the attention panel contains not only directives to the user about what to look at next, but also displays appropriate enhancements to the image, according to the needs of the problem-solving process.

The overall control structure of the VIA-RAD 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 focusing 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 focused in order to obtain perceptual information, the strategy selector makes this decision, on the basis of 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.

**VIA-RAD interaction example**

In this example, a simple logical user view is adopted, which hides any ruled-out hypotheses.

In Figure 4 initial information is posted to the blackboard which indicates that the radiologist has assessed the lungs to be abnormal, and that the initial finding is hypothesized to be a density. Knowledge source KS-01 contributes to the problem-solving process by assisting the user to move up a level of abstraction in thinking about the finding hypotheses. In order to do this, it needs feature information, and, since this is not currently available on the blackboard, a strategy of Not-Enough-Information is invoked, which constructs the high level plan to examine the primary finding. This, in turn, leads to the generation of a detailed attention plan which sets the perceptual goals and posts directions to the attention panel of the blackboard for the user to see. In this hypothetical example, the user is given perceptual assistance in the form of image enhancements, which are executed at the appropriate stages.

In Figure 5, values for the features have been obtained and posted to the blackboard, thus fulfilling the first action step of knowledge source KS-01. With this information, it can now generate a list of candidate general finding hypotheses, which the user must examine and assess (e.g. mass, consolidation, infiltration). The user's attention is thus directed to that part of the blackboard in order to select the subset of general findings which is appropriate to the current case.

![Figure 2: Relationships between knowledge concepts](image)

![Figure 3: VIA-RAD system design](image)

![Figure 4: VIA-RAD interaction example](image)
Figure 4 VIA-RAD example part 1

Figure 5 VIA-RAD example part 2

Figure 6 VIA-RAD example part 3

KS-04 becomes active. This knowledge source needs evidence to fulfill its actions, and since there is only one current diagnostic hypothesis, the support diagnosis strategy is invoked. This strategy formulates a high level plan to examine the expectations associated with the diagnostic hypothesis with respect to secondary findings. Since bronchogenic carcinoma is presented in the example, the user is directed to focus attention on the bones, and to seek lesions. Again, a bone enhancement algorithm may be executed to also provide perceptual assistance to the user.

TESTING AND EVALUATION

An initial prototype VIA-RAD system was implemented in order to test the feasibility of some of the design ideas. The scope of the program was limited to a small number of cases which contained only mass abnormalities, both benign and malignant. The knowledge base consisted of nine landmarks, eight general findings, 15 specialized findings, six features and 12 diagnoses. Three types of strategies and 12 knowledge sources were also identified. The display of the VIA-RAD prototype consists of two computer screens, with one monitor displaying the X-ray image and its enhancements, while the other presents a window-based interface which corresponds to the panels of the blackboard. An example of the interface screen is shown in Figure 7.

An observational study was then conducted to obtain user feedback on the approach and the prototype system itself. Five different radiology residents were recruited, and were tested with four different cases, comparing diagnoses with the VIA-RAD system versus
those with a noninteractive computer-displayed image. The results of this study were quite interesting. In general, subjects gave more positive ratings when they were not using the VIA-RAD system: the image quality was rated better, confidence was rated higher, and cases were rated somewhat less difficult without VIA-RAD. On the other hand, an analysis of qualitative performance comparing the final dictation of the project expert with the subjects' dictations revealed a considerable improvement in accuracy when VIA-RAD was used. These differences could be due to a number of factors, including the increased time spent looking at cases when interacting with the VIA-RAD system. In addition, the presentation of the VIA-RAD selections may have influenced the terminology used by the subjects, although they were briefly exposed to the specific vocabulary of the system ahead of time.

From the postsession questionnaires, we also obtained subjective reactions to the program, which were, in general, quite positive. The limited knowledge base and slow response of the system were most often cited as hindrances. On the other hand, the subjects mentioned both image enhancements and the presentation of candidate hypotheses as helpful features. It was felt that the program often presented findings and/or diagnoses which should have been considered, but might otherwise have been overlooked.

The purpose of this study was to observe the interaction of radiologists with the VIA-RAD prototype, and to try to demonstrate the feasibility of further developing this program into a more fully functional system. In general, the choice of a blackboard design for this system has proven to have a number of advantages. First of all, it allows both user and system to respond opportunistically to the emerging problem solution. This leads to a close cooperation between man and machine, which was one of our initial objectives. The modularity of this design greatly facilitates both the extension of the knowledge base, and the addition of more knowledge sources, which communicate only via the blackboard. The structure of the blackboard itself allows the communication of both perceptual and problem solving information, through image enhancement, attention direction and hypothesis management. More active participation of the system in the perceptual analysis of the image through image processing and automatic feature extraction can easily be incorporated and the results presented to the user for further diagnostic assessment. This opens the door to more extensive use of computational capabilities without usurping the decision making of the human practitioner. Another type of computational extension that is facilitated by the blackboard design is the posting of case history information from a computerized patient database. This could then be used, not only by the radiologist, but also by the system knowledge sources, which would take this information into account in the management of the hypotheses. Current attempts to include images in...
patient databases will eventually allow computerized access to multiple image modalities as part of patient management. The flexibility and extensibility of the blackboard system, it is felt, will prove to be a great advantage in accommodating this future information load.

DISCUSSION

In developing the VIA-RAD system, we learned a great deal about diagnostic radiology, about perception and problem solving, and, more generally, about cooperative system design. The approach taken in this work includes the collection and analysis of human performance data, cognitive modelling of the interaction between perception and problem solving, and incorporation of this knowledge into a general system design. These steps are now described in the context of issues particularly relevant to the study of human-machine cooperation: analyses of how people work, knowledge needed for effective cooperation, and appropriate architectures for deploying this knowledge.

Analyses of how people work

One of the challenging problems in building computer systems is how to maximize the effectiveness of humans using them. This is compounded when the purpose of the system itself is to aid the human in performing a complex task. Understanding how humans process information, and what type of assistance may be needed, requires a deeper study of how people work than has been typical to date. A three-stage cognitive analysis methodology developed in the course of this work presents an approach to obtaining such in-depth knowledge.

The first stage of 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. In the VIA project, the data collection was formulated as a telescoping series of activities, each one providing results which formed a basis for subsequent experiments, as illustrated in Figure 8.

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 the vehicle of an encoding scheme, as shown in Figure 9.

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 preprocessed 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 Figure 10. (Comparisons between expert and novice performance were not considered at this time.) It was found that similar clusters of concepts did occur at corresponding times for particular cases, and that these could be interpreted as evidence for plan-like activities. Moreover, the grouping within the clusters or patterns showed tendencies towards perceptual or problem solving compositions, and suggested an ordering of activities in the movement towards a diagnostic solution.

This methodology represents a framework for acquiring domain knowledge that has provided a number of useful results in the area of diagnostic radiology, and current work is under way to examine its utility in other applications.
Knowledge needed for effective cooperation

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 of visual interaction described previously. 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 also lends itself to incorporation in the design of a cooperative computerized assistant.

Appropriate architectures

The primary requirement of the system is to facilitate cooperative man–machine problem solving, and to assist and enhance the human visual interaction process. To do this ‘intelligently’, the system must incorporate both domain knowledge and process knowledge, and can be conceptualized as an overlay to the user’s own visual interaction process. At critical stages, the system affords directions to the user to focus attention on particular aspects of the problem, either in the hypothesis space or in the image itself, and provides appropriately enhanced information. In order to do this, there must be a certain amount of internal simulation, where the computer system retrieves, combines and posts knowledge in a manner that is consistent with the model’s predictions. Therefore, the VIA system design 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.

A more general VIA blackboard has been designed, which represents an abstraction of those concepts shown in the VIA-RAD application. This design more closely resembles aspects of the model of visual interaction, especially with respect to working memory. The buffers of the model suggest 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. 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.

SUMMARY AND CONCLUSIONS

With the advent of powerful technologies for displaying multidimensional scientific data, the development of new strategies for efficient use of these capabilities is of foremost concern. It is becoming more and more evident that, in order to perform complex tasks, man and machine can no longer be treated as separate entities, but must be considered together as a unified decision-making system. The current trend is to try to augment human potential and productivity by tightly integrating interaction mechanisms with underlying domain knowledge in what are known as cooperative problem-solving systems.

The research described in this paper represents an approach to developing a cooperative computerized assistant for a particular type of visual reasoning task. In this work, the purpose of the computational agent is to assist and enhance the performance of the human agent in a process which links perceptual and problem-solving capabilities. An analysis of how people do this type of work provides information about the human visual interaction process, suggesting what type of assistance is needed. These results are then formalized in an information-processing model, which provides task-related knowledge about the user in a form which can be embedded in a computerized system. The model is used to determine when it would be cognitively effective to afford assistance. Finally, the computational agent itself is organized as a blackboard-style system which allows opportunistic interaction between the user and the system, and whose components are designed to
supplement the user's knowledge and visual interaction activities.

The joint man–machine cognitive system perspective stresses the use of computational technology to aid the user in the process of solving a problem\(^4\). While this approach is very appealing in many ways, it has also been criticized as being too oriented towards an information-processing viewpoint, which may bias the gathering of user data, and require the joint system to be optimized according to external requirements of the environment\(^14\). A different approach is to look at how the environment itself may be changed by the computer support system, thus affecting the user’s perception of the setting in which the activity is situated\(^15\). The VIA system represents a hybrid approach to cooperative assistance. The management of hypotheses and selection of strategies are founded on an information-processing type of model. On the other hand, the attention director 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.

The next immediate challenge is to pursue these ideas in new domains, and several projects are currently under way which will test the assumptions and validity of the VIA approach. Preliminary work has begun on the collection and analysis of visual reasoning data in the domain of geographical information systems (VIA-GIS). This newly emerging technology involves a variety of users making decisions on the basis of the visual display of geographical data. As such, it is an excellent candidate for the further study of human perception and problem solving, as well as the application of the cooperative assistance paradigm. Another area of application is in the teleassistance of robots (teleVIA). In this project, an additional computational agent (i.e. a semiautonomous robot) must be incorporated as part of the cooperative group, and the role of the teleVIA system is to filter and enhance data from the remote robot when it requires assistance from the local (human) operator. In all of these applications, there is an emphasis on understanding the underlying cognitive activities in the visual reasoning task. This enables the intelligent computer system to invoke appropriate visualization and knowledge tools in order to effectively cooperate with the human in complex problem solving.

**REFERENCES**