A Cognitive Model of Analytical Reasoning Using GRE Problems

Final Project: Cognitive Modeling and Intelligent Tutoring Systems

Scott B Kaufman
skaufman@cmu.edu

Helen J Ross
hjr@andrew.cmu.edu

Eli M Silk
silk@cmu.edu

Friday, December 14, 2001

A cognitive task analysis was performed to aid in the development of a cognitive model and associated tutor for the task of correctly solving a subset of the analytic ability questions from the Graduate Record Examination (GRE). This task combines a number of important issues relevant to cognitive theory including working memory load, building appropriate external representations, symbol use and manipulation, problem solving and goal structure, as well as abstract versus context-sensitive learning and transfer. All of these issues are considered here in an initial step towards the development of a more complete model of analytical reasoning using GRE problems.

1 Background

The subset of problems analyzed here includes only the GRE problems classified as analytical reasoning (AR) problems. AR problems are distinguished from logical reasoning (LR) problems which are also found in the analytical ability section of the GRE. The LR subset consists of problems requiring argument analysis in contrast to the AR subset which consists of constraint-satisfaction problems. This class of problems requires one to understand a given structure of arbitrary relationships among fictitious persons, places, things, or events, and to deduce new information from the given relationships.

Preliminary analyses centered around a variety of example problems taken directly from a past GRE exam. These analyses resulted in the identification of a number of enigmatic skills that were effective in determining correct solutions, ranging from the general skill of efficient symbol use and manipulation to the more task-specific ability to distinguish relevant constraints from both the passage and the given bulleted conditions. These skills were condensed into three possible sources of difficulty that, on the surface, appeared to offer the most opportunity for appropriate intervention strategies. Each of these will be discussed in turn.

Intermediate Inferences The GRE AR class of problems has at least two levels of difficulty. The easier level consists of those problems whose answers can be determined by directly applying the given constraints to the answer choices to eliminate those that are not consistent. The more difficult problems may consist of those that require intermediate inferences to be made from the given constraints. A strategy that was presented very early on in the analysis was to apply each given constraint, one at a time, to the answer choices, and successively eliminate those that were not consistent with each constraint. This strategy clearly breaks down when an inference must be made that is not given in the problem. Nevertheless, the strategy remains as an effective method for narrowing the range of possible answer choices.

Often it is the case that more difficult problems can be reduced to a key inference or insight which is the most challenging to form. Subsequent to the time that key inference is made, the remainder of the problem tends to unfold with relative ease. Reaching the point at which the key insight is useful while at the same time not yet having that inference either immediately available nor having it be easily constructed may account for longer response times on a particular problem.

The question that remains is what characteristic of the problems that require intermediate inferences make them harder to solve? One possible explanation is based on the theory of working memory capacity

[Baddeley, 1990]. When participants are required to solve a problem that can only be solved correctly by making an intermediate inference, they must first seek out the relevant rules, and only those relevant rules. Then they must combine these rules in order to induce a new rule which is the inference. The reason why this is not always so easily done is because it may be difficult to keep more than one rule in working memory at a given time. In cases where the participant must seek out more than two relevant rules to induce a new rule, this may overload the participants working memory, making it hard for that participant to make the connections necessary between all the rules in order to make that intermediate inference.

Therefore, it seems reasonable to provide hints that would specifically address the key intermediate inferences needed to solve more difficult problems. These hints would essentially be helping the participant to seek relevant rules, and allow them to constantly refer back to the hint when the working memory system becomes overloaded.

A number of research efforts have focused heavily on the development of methods to scaffold intermediate inferences. In particular, a major design principle of reasoning-congruent learning environments is to visibly render and openly provide comments on the normally invisible intermediate states between the given statements and the end solution [Merrill and Reiser, 1993]. The design of cognitive tutors has also stressed the importance of intermediate results in the form of explicit communication concerning the goal structure involved in successful problem solving [Anderson et al., 1995]. In every complex problem-solving context, the process of determining a final solution is greatly assisted by the utilization of effective skills for combining known results and applying those results on a path leading toward the final goal.

Organizational Representation Another way to classify this class of problems into two types, easier and harder problems, is by the type of representation required to solve them efficiently and effectively. Previous research has distinguished between problems with determinate solutions (*i.e.* state sufficient constraints to determine a unique satisfying model) versus those with indeterminate solutions and the types of organizational representations that are most appropriate for each of them [Cox et al., 1995]. In this research, graphical representations (*e.g.* tables, matrices, and graphs) were contrasted with sentential forms (*e.g.* using sentences comprised of words or symbols which adhere to a set of syntactical rules). Graphical representations have the advantage over sentential forms because they take advantage of the parallel searching mechanisms of the human visual perception system. On the other hand, graphical representations are considered to be weakly expressive since they are forced to leave no information unspecified, and as a result, there are many abstractions which they cannot express without the use of an indefinitely large number of diagrams.

Understanding in what problems each of these forms is appropriate is a valuable skill. Not only is it likely that appropriate choices for a graphical representation may be one of the most important factors for successful problem-solving in this domain, it is also likely that the skills involved in doing so are very complex and hard to learn. Previous work has been done which shows that beginning math students often choose graphical representations based on surface characteristics and similarities rather than their appropriateness to the specific sorts of analysis required [Baker et al., 2001]. Beginning students in all complex domains often spend little time weighing alternative representation choices, since its result is often not tangible and usually does not contribute any immediate, measurable progress toward an answer. Yet, the hidden skills involved in choosing the most effective representation may be the primary, underlying source for efficient and effective solution strategies. Therefore, it is important to consider appropriate interventions for scaffolding these skills.

From a cognitive standpoint, diagrams are useful because they lessen the working memory load on the user by dividing the load between the visual and textual systems. This freeing up of resources in the textual working memory system may be just what is needed to turn an unsuccessful solution into a successful one.

Decontextualization Although the GRE AR problems are presented in story format, they are often very complex, involving a range of objects, qualities, and relationships. Although a considerable amount of research has been undertaken which supports the claim that context is valuable in forming effective solution strategies, the complexity of this class of problems, as well as their fictitious cover stories lead to questioning the appropriateness of that claim for this particular task. The cover stories seems quite arbitrary in the sense that they rarely if ever give the participant more information about how to solve the problem. If anything, they cause the intrusion of irrelevant information [Passolunghi and Siegel, 2001].

Since there is so much information that the participant must organize and retain in working memory in order to solve these types of problems, decontextualization may significantly reduce the cognitive working memory capacity of the participant by essentially giving the participant “less to work with” that is more relevant to the problem at hand. It is quite clear in psychology that humans have a very limited working memory capacity. The more numbers and letters that are in working memory at one particular point in time, the less total numbers and letters that participant will be able to work with. Therefore, it would make sense that decontextualizing the irrelevant letters and numbers would give the participant more capacity to work with at a given time, thereby increasing the chances that the participant will make connections between necessary rules and infer appropriate immediate inferences.

2 Methods

The cognitive task analysis consisted primarily of two empirical descriptive approaches, an informal think-aloud study and a more formal difficulty factors assessment (DFA). The think-aloud study served to inform the design and implementation of the subsequent DFA. Easy access to large numbers of authentic problems from past exams and the simplified scoring system of the standardized tests supported the decision to concentrate on the DFA as the primary source for data.

Think-Aloud The first form of the empirical descriptive analysis was a short, informal think-aloud study. The purpose of this aspect of the study was to obtain a clearer initial idea of what strategies were common in solving GRE AR problems, and for exposure to some of the factors that were possibly common sources of difficulty. The think-alouds were conducted using a group of problems, containing both easier and more difficult problems, selected from a past GRE exam. The more difficult problems were appropriate in order to effectively challenge the think-aloud participants.

Each participant was given the questions in their original paper format, along with extra blank sheets to be used for scratchwork. Two subjects participated in the informal study and each were given as much time as they needed to solve the problems. Their spoken words were recorded and their papers collected for later inspection. The participants were both college undergraduates in their final year of school, with minimal or no experience solving GRE AR problems.

DFA The second form of the empirical descriptive analysis was a one factor study designed to determine the most appropriate and effective intervention strategy for success in the target task. The intervention factor was varied according to the three possible intervention strategies discussed above. Each of these strategies were confirmed to be worthy of further consideration based on results from the informal think-aloud study. Those strategies were:

- decontextualizing problem statements and answer choices;
- scaffolding intermediate inferences and key insights; and
- providing an organizational and/or graphical representation.

We chose four difficult problems whose difficulty was determined by the percent of examinees who answered the questions correctly on this particular edition of the test when it was used as an official GRE exam. The DFA consisted of four conditions for each problem. The four conditions were crossed with the four problems.

This resulted in four distinct test forms in which each difficulty factor and each problem were presented once. There was no control for problem ordering since that would require significantly more test forms than was appropriate for this initial cognitive task analysis study.

Twenty-four subjects were asked to participate in the study, each completing one of the test forms. Their final answers and solution times for each problem were recorded. The participants were taken from a random sample of college students or recent graduates. The participants varied in terms of experience with the GRE AR class of problems.

3 Results

Think-Aloud The results of the think-aloud study were the difficulty factors tested in the DFA. In Figure 1, the first participant made an inference regarding the relative positions of objects and landscapes, which were two types of drawings to be ordered in one group of problems. This inference lead the participant to recognize which type (object or landscape) must be in an even position, and which ones in an odd. This key inference aided the participant in several of the problems.

> No two drawings next to each other. Ok that means you must go object, landscape, object, landscape... Since you have 4 objects, and 3 landscapes, you have to have the lines begin with an object. Which also means you have to end with a landscape, no, object.

Figure 1: Subject 1 from the think-aloud study formed a key intermediate inference.

The second participant experienced difficulty relating to issues of context. Even though the subject decontextualized the problem in the beginning, he abandoned its use for the remainder. The subject also spent significant time attempting to understand the names. The participant provided the following comments shown in Figure 2. This seems to be a direct hint to the fact that contextualization and working memory are big issues in the difficulty of solving these problems.

> Fiona and Gabriela, Judith—I'm going to name these by their letters. F, G and J, K, M, S, T, and Y.
> ...
> I will say that for all of these the hardest thing is keeping track of all of these eight names!

Figure 2: Subject 2 from the think-aloud study recognized the need to translate the context to symbols, which reduces working memory load.

DFA The quantitative results from the DFA indicated no significant difference in regard to the cover stories, but there were significant differences between two of the difficulty factors and the original problem (Intermediate Inference and Decontextualization). The results are summarized in Table 1.

Condition	Original Problem	Representation	Interm. Inference	Decontextualized
% Correct	62.5	66.7	79.2	83.3
% Increase	-	4.2	16.7	20.8
Effect Size	-	0.3	1.0	1.3

Table 1: Results from the DFA, percent correct, percent increase, and effect size for each condition.

4 Discussion

The results of our DFA will influence the design of the cognitive tutor. Both decontextualization and intermediate inference scaffolding improved performance on the GRE problems, however organizational representation did not prove to be significant. It may be possible that organizational representation did not have an effect in our DFA because particicipants may have to create their own representation in order to interpret and use it correctly. Participants in this experiment noted that they either misunderstood the representations given or used it incorrectly. If this the case, then visual representations may be an effective problem solving strategy, but our approach to scaffolding them needs to be different such as teaching them how to make the decision what kind of representation to use depending on type of problem. This should a worthy next step in our research process.

Based on the factors in the DFA that did have a significant effect, our tutor will scaffold these actions by referring to the objects to be manipulated in a more symbolic form, and if necessary, showing the user how to decontextualize the problem. The hint messages will be largely based upon helping the user form intermediate inferences. As shown by the DFA it is important to aid the user in forming a deeper understanding of a problem instead of relying on a more surface understanding that may lead to incorrect assumptions.

One participant in the DFA, who had more experience taking GRE exams than most of the other participants, offered a retrospective analysis of strategies and answer choices. In particular, he related how on one problem he identified it as an “easy” problem from the context of the questions and therefore chose the corresponding “easy” answer without considering his decision more carefully. He cited his use of test-taking strategies specific to the GRE and the need to finish the problems in a short amount of time as justification for his strategy choice. The answer he chose was not the correct one. This supports the claim made here that relying on intuition is often a fragile and misleading approach to adopt when the goal is to answer these GRE AR problems correctly. The particular goal of this research is not to develop a model of desired performance that includes test-taking strategies, but rather, one that includes sound skills for consistently accurate solutions of this domain of constraint-satisfaction problems. In the same manner, it is likely that introducing context in problem statements, and more specifically those examinees who rely on that context, is more a source of difficulty to be overcome than it is a bridge to a variety of informal modes of correct reasoning. With the complexity present in the structures of these GRE AR problems, it may be a wise decision to develop a set of general, abstract modes of thinking that provides the validity and assurance needed for consistent success in this domain.

References

[Anderson et al., 1995] Anderson, J. R., Corbett, A. T., Koedinger, K. R., and Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2):167–207.

[Baddeley, 1990] Baddeley, D. D. (1990). *Human memory*. Erlbaum, Hove, England.

[Baker et al., 2001] Baker, R. S., Corbett, A. T., and Koedinger, K. R. (2001). Toward a model of learning data representations. In *Proceedings of the Cognitive Science Society Conference*, pages 45–50.

[Cox et al., 1995] Cox, R., Stenning, K., and Oberlander, J. (1995). The effect of graphical and sentential logic teaching on spontaneous external representation. *Japanese Journal of Cognitive Science*.

[Merrill and Reiser, 1993] Merrill, D. C. and Reiser, B. J. (1993). Scaffolding the acquisition of complex skills with reasoning-congruent learning environments. In *Proceedings of the Workshop in Graphical Representations, Reasoning and Communication from the World Conference on Artificial Intelligence in Education*, pages 167–207, Edinburgh, Scotland. The University of Edinburgh.

[Passolunghi and Siegel, 2001] Passolunghi, M. C. and Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, 80:44–57.