A DIALOG-BASED INTELLIGENT TUTORING SYSTEM FOR PRACTICING COMMAND REASONING SKILLS

Joan M. Ryder
CHI Systems, Inc.
Lower Gwynedd, PA

Arthur C. Graesser
University of Memphis
Memphis, TN

Jennifer McNamara, Ashish Karnavat, Ed Popp
CHI Systems, Inc.
Lower Gwynedd, PA

ABSTRACT

The Army Research Institute (ARI) has developed training/instructional materials, *Think Like a Commander (TLAC*), for coaching command reasoning through adaptive thinking exercises using battlefield situations. Development of an accompanying intelligent tutoring system (ITS) would allow additional practice without the need for additional human mentor involvement. However, developing such a system is a challenging problem that does not fit the mold of previously successful ITSs because of the open-ended, non-procedural nature of thinking exercises and the need for language-based interactions. Our approach is to couple two technologies that have been used successfully for different aspects of the problem. The system, called Automated Tutoring Environment for Command (ATEC), is adapting the dialog management capability from AutoTutor and linking it with a cognitive model-based instructional agent that replicates the knowledge and role of the human TLAC mentor.

The ATEC system presents a battlefield situation and then initiates a dialog between a virtual mentor (instructional agent) and a student as they collaboratively discuss the situation. The virtual mentor poses questions, evaluates student responses, determines the sequence of questions, and ultimately assesses performance on the basis of the specificity of questioning and the depth of probing and hinting that is needed to adequately answer the questions. The system includes natural language extraction, speech act classification, and dialog management. This paper discusses the development of the ATEC system as part of a Phase II SBIR, including the system architecture and functional components, the methods for accomplishing each function, the analyses conducted, and the issues remaining to be addressed. Ongoing complementary efforts include a related Phase II SBIR with a different ITS approach for battle command conceptual skills, and a computer-based program developed by ARI that human instructors are using in the Armor Captains Course at Fort Knox's University of Mounted Warfare.

AUTHORS BIOGRAPHIES

Dr. Joan Ryder is a Managing Cognitive Scientist at CHI Systems. She has led and contributed to numerous research and development efforts over the last twenty years involving cognitive modeling, training system analysis and development, user interface design, and human factors analysis. She is currently the Principal Investigator for the ATEC system development effort reported in this paper. Her interests are in the areas of cognitive modeling, cognitive task analysis, instructional agents, and advanced instructional systems. She holds a Ph.D. in experimental psychology from Brandeis University and an A.B. in psychology from Brown University.

Dr. Arthur Graesser is a full professor in the Department of Psychology and an adjunct professor in Mathematical Sciences at The University of Memphis. He is co-director of the Institute for Intelligent Systems and director of the Center for Applied Psychological Research. Dr. Graesser's primary research interests are in cognitive science and discourse processing. More specific interests include knowledge representation, question asking and answering, tutoring, text comprehension, inference generation, conversation, reading, education, memory, expert systems, artificial intelligence, and human-computer interaction. In recent years, he has led a team that has developed AutoTutor, a tutoring system that teaches physics, introductory computer literacy, and other topics by engaging the student in complex human-like tutoring dialogs using natural language. Dr. Graesser received his Ph.D. in psychology from the University of California at San Diego.

Ms. Jennifer McNamara is a Cognitive Engineer II at CHI Systems. Currently Ms. McNamara is involved in cognitive engineering and training systems research and design. She has a B.S. in psychology from Drexel University, and is pursuing an M.Ed. in Instructional Systems Training Design and Development at The Pennsylvania State University.

Mr. Ashish Karnavat is a Software Engineer at CHI Systems. He has a BS in Computer Engineering from the University of Pune, India, and will receive an M.S. in Computer Science from the University of Memphis. He has experience in computational linguistics and has worked on a number of natural language dialog systems.

Mr. Ed Popp is a Senior Human-Systems Engineer at CHI Systems. He has a B.S. in Aerospace Engineering from West Virginia University and is pursuing an M.S. in Information Systems at Drexel University. Mr. Popp has over sixteen years of experience in military aviation operations and technology development and assessment for R&D programs. He was a Captain in the US Army Aviation Branch.

A DIALOG-BASED INTELLIGENT TUTORING SYSTEM FOR PRACTICING COMMAND REASONING SKILLS

Joan M. Ryder
CHI Systems, Inc.
Lower Gwynedd, PA

Arthur C. Graesser
University of Memphis
Memphis, TN

Jennifer McNamara, Ashish Karnavat, Ed Popp
CHI Systems, Inc.
Lower Gwynedd, PA

INTRODUCTION

The Army Research Institute (ARI) has developed training/instructional materials, *Think Like a Commander (TLAC*), for coaching command reasoning through adaptive thinking exercises using battlefield situations (Ross & Lussier, 1999; Lussier, Ross & Mayes, 2000). The TLAC program is currently being used with Brigade Command designees attending the School for Command Preparation of the Command and General Staff College (CGSC) at Fort Leavenworth, KS (U.S. Army Research Institute, 2001) and in the Armor Captain s Career Course at Fort Knox, KY (Shadrick & Lussier, in preparation). In its current form, TLAC presents tactical situations (called vignettes) as PowerPoint presentations in a classroom setting. Following the presentation of a vignette, there is a classroom discussion of the vignette led by an instructor acting as tutor or mentor. The instructor begins by asking general questions to stimulate thinking, and then asks increasingly more directed questions to probe for themes that have not been addressed. The discussion is organized around eight themes that underlie common patterns of expert tactical thinking: (1) keep focus on mission and higher commander s intent, (2) model a thinking enemy, (3) consider effects of terrain, (4) use all assets available, (5) consider timing, (6) see the bigger picture, (7) visualize the battlefield, and (8) consider the contingencies and remain flexible. There is a set of questions distinctly tailored to each of these themes. Approximately a dozen vignettes have been created, all within a specific fictitious situation context.

An accompanying intelligent tutoring system (ITS) would provide additional practice without the need for additional human mentor involvement. However, the development of an ITS for interactive self-training of thinking skills, such as battlefield command reasoning, is a challenging problem. It does not fit the mold of previously successful ITSs because of the open-ended, non-procedural nature of the interactions. Furthermore, tutorial dialog systems are more complex than other ITSs, because they need to have a dialog management capability in addition to standard ITS components, and because natural language-based tutoring systems are in their infancy. Our approach to this challenge is to couple two technologies that have been used successfully for different aspects of the problem. Our system, called Automated Tutoring Environment for Command (ATEC) is integrating the dialog management capability from AutoTutor (Graesser, Person, Harter, & TRG, in press) with a cognitive model-based instructional agent that incorporates the knowledge and strategies of a human TLAC instructor.

DIALOG-BASED INTELLIGENT TUTORING SYSTEMS

The vision of having a computer communicate with users in natural language was entertained shortly after the computer was invented, but it was not until Weisenbaum s (1966) ELIZA program that a reasonably successful conversation system could be explored. Two decades of exploring human-computer dialog systems had a disappointing outcome. By the mid-1980 s, most researchers in artificial intelligence were convinced that the prospects of building good conversation systems was well beyond the horizon because of inherent complexities of natural language processing, because of the unconstrained, open-ended nature of world knowledge, and because of the lack of research on lengthy threads of connected discourse. However, the early pessimism was arguably premature. There have been a sufficient number of technical advances in the last seven years that researchers are revisiting the vision of building such dialog systems. The field of computational linguistics has recently produced an impressive array of lexicons, syntactic parsers, and semantic processing modules that are capable of rapidly extracting information from naturalistic text and discourse (Jurafsky & Martin, 2000). The world knowledge contained in an encyclopedia can be represented statistically in high dimensional spaces, such as latent semantic analyses (LSA) (Landauer, Foltz, & Laham, 1998); an LSA space can be created overnight and produces semantic judgments that are surprisingly close to human judgments. The representation and processing of connected discourse is much less mysterious after two decades of research in discourse processing (Graesser, Gernsbacher, & Goldman, in press). There are now generic computational modules for building dialog facilities that track and manage the beliefs, knowledge, intentions, goals, and attentional states of agents in two party dialogs (such as COLLAGEN, Rich & Sidner, 1998).

Computer-based natural language dialog is particularly feasible in tutoring environments when the knowledge domain is more qualitative (e.g., verbal reasoning) than precise (e.g., mathematics), and when the shared knowledge (common ground) between the tutor and student is low to moderate rather than high. Stated differently, it is entirely reasonable to build a natural language dialog system when the computer and tutor do not track what each other knows at a fine-grained level and when the computer produces dialog moves (e.g., questions, hints, assertions, short responses) that advance the dialog and help achieve the dialog goals. Indeed, it is well documented that human tutors are not able to monitor the knowledge of students at a fine-grained level, but they nevertheless provide productive dialog moves that lead to significant learning gains in the student. These considerations motivated the design of AutoTutor (Graesser, Person, Harter, & TRG, in press; Graesser, VanLehn, Rose, Jordan, & Harter, 2001; Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, & TRG, 1999) as well as the current ATEC system.

AutoTutor is a dialog-based tutor developed by Graesser and colleagues at the University of Memphis. AutoTutor tries to comprehend student contributions that are typed in natural language and to simulate dialog moves of human tutors or ideal tutors. AutoTutor has been implemented and tested on approximately 400 college students on the topics of introductory computer literacy and conceptual physics. It provides linguistic analysis capabilities, dialog turn generation (e.g., a clarification question needs an answer, an assertion needs an evaluative response), and tutorial strategy (e.g., when and how to provide hints, prompts, change of topic, ordering of topics). The goal of AutoTutor is to get the student to do the talking rather than to simply deliver correct information. AutoTutor first presents a major question or problem that requires a lengthy answer, which can be segregated into a set of *expectations*. As the learner types in information, distributed across several turns, AutoTutor constantly compares the learner s input to the expectations. LSA is used to compare the words in the student contributions to the words in the expectations. Some expectations are not articulated by the student, so AutoTutor asks questions or gives hints that attempt to get the learner to fill in the expectations; when all else fails, AutoTutor simply asserts the missing expectations. AutoTutor engages in mixed initiative dialog by answering student questions and responding appropriately to a broad array of student contributions. The management of the conversation is reasonably successful because bystanders cannot distinguish whether particular dialog moves are generated by AutoTutor or by a human tutor (Person, Graesser, & TRG, 2002).

INSTRUCTIONAL AGENTS IN INTELLIGENT TUTORING SYSTEMS

Although there has been considerable success in intelligent tutoring systems (ITSs) with constrained interaction and in well-structured learning domains such as geometry, programming, and physics (see Anderson, Corbett, Koedinger, & Pelletier, 1995; Shute & Psotka, 1995), many of the techniques used (such as model tracing) are not applicable to less-structured real-world training problems. The canonical ITS architecture includes, at a minimum, the following three components: (1) an *expert module* that contains a representation of the knowledge to be presented and a standard for evaluating student performance, (2) a *student module* that represents the student s current understanding of the domain, and (3) an *instructional module* that contains pedagogical strategies and guides the presentation of instructional material (e.g., Polson & Richardson, 1988; Wenger, 1987). These three aspects of intelligence need not be separate components. Current thinking is that the key to intelligent training is designing the system to behave intelligently in terms of providing adaptive instruction that is sensitive to a diagnosis of the student s knowledge structures or skills (Shute & Psotka, 1995). The indeterminacy and complexity of many domains, including battlefield thinking, preclude the use of model tracing approaches to student modeling, which are only applicable to procedural learning and reasoning in well-structured domains. Furthermore, recent pedagogical theories have focused on collaborative learning, situated learning, deliberate practice, constructive learning, and distributed interactive simulation, all of which call for modifications of the traditional ITS paradigm and the creation of alternative interactive learning environments.

Another approach to the problem has been to use cognitive modeling technology to create a model of an instructor that can be embedded in an interactive learning environment for these more complex, indeterminate domains. These models, called *instructional agents*, embody the reasoning of a human instructor and include all three aspects of tutoring intelligence in one model: domain knowledge, diagnostic reasoning, and pedagogical reasoning. The difficulty of diagnosing deficiencies in knowledge and skill or of selecting appropriate pedagogical strategies is not diminished using instructional agents.

However, the problem becomes more tractable when we analyze the expertise of an instructor using cognitive task analysis methods, and we create an executable model of the tutorial knowledge that is applicable in the instructional domain. Cognitive modeling provides a more natural language for representing human expertise than other AI formalisms. The associated cognitive task analysis provides a richer method for acquiring that knowledge than other knowledge engineering techniques. This approach has been used successfully in other complex domains that preclude the use of model tracing approaches to student modeling (e.g., Ryder, Santarelli, Scolaro, Hicinbothom, & Zachary, 2000; Zachary, Santarelli, Lyons, Bergondy, & Johnston, 2001).

An instructional agent, created using CHI Systems iGEN cognitive agent technology (Zachary, Ryder, Ross & Weiland, 1992; Zachary, Le Mentec, & Ryder, 1996), serves as the reasoning engine for ATEC, as described below. However, the instructional agent approach is being modified for this application to incorporate the pedagogical approach of AutoTutor and to integrate its language processing mechanisms.

ATEC RESEARCH AND DEVELOPMENT ISSUES

Pedagogical Approach

Command reasoning skills deal with battlefield thinking habits that are characteristic of expert tactical thinkers. Although well understood conceptually, the skills are often absent during realistic tactical problem solving of less experienced commanders. Schoolhouse learning involves primarily declarative knowledge about command principles and tactics, with training in task-specific procedures (in declarative form) that are based on these principles and tactics. Repeated real-time practice allows for proceduralization and chunking of the skill (deriving domain-specific problem-solving strategies and then integrating separate pieces of declarative knowledge). TLAC addresses the transition from schoolhouse learning to adaptive expertise by providing deliberate practice experience (Ericsson, Krampe, & Tesch-Romer, 1993) in a form most likely to facilitate the development of command expertise more rapidly and consistently than through experience alone. ATEC incorporates the pedagogical approach of deliberate practice through a set of vignettes (constructed to focus on difficult aspects of command reasoning) that ARI used to develop the TLAC program (Lussier, Ross, & Mayes, 2000). ATEC also attempts to replicate the coaching and scaffolding that human instructors provide in the TLAC program, which is entirely compatible with the constructivist style of instruction (Graesser, VanLehn, et al., 2001), in which students actively generate ideas and problem solutions. However, improving the dialog management logic and incorporating instructional intelligence to allow the virtual mentor to respond adaptively and appropriately is the major challenge of this effort (see Dialog Management and Meaning Representation discussions below).

The macrostructure of the dialog is controlled by a curriculum script that includes a set of questions organized around the eight themes in TLAC. For each theme, there is a general question meant to start discussion of that aspect of the problem. Responses are not considered correct or incorrect, but rather starting points for a dialog. Associated with each general question, there are anticipated good answers (called expectations) based on reasonable approaches to the problem posed. The virtual mentor assesses the student s response in relation to the possible good answers (see Meaning Representation below on the methodology for this). There is also a set of progressively more specific questions for the virtual mentor to ask to prompt the student into thinking about any aspect of the theme not discussed in response to the initial question. This approach is based on the AutoTutor curriculum script approach, but modified to provide a mentoring style of dialog rather than the tutoring style previously used in AutoTutor for teaching computer literacy or physics (subjects in which the material is conceptual and there is one correct answer to the questions posed by the tutor).

The classroom use of TLAC involves an instructor and a class of 20 or more students. Attendance at the class is required and the social aspect of the classroom situation provides motivation. In developing an instructorless system for individual use, the issue of motivation must be addressed. How can an automated system keep a commander in training engaged and bring him back to use the system? Our belief is that the intrinsic motivation of getting some value from system use is the most workable approach in this situation. A system for individual use will engage each student and allow him to actively participate in conversation with the virtual mentor without worrying about the reaction of classmates or the instructor. Primarily, however, the system must keep the student engaged by intelligent selection of questions and dialog that is conversationally appropriate and pedagogically effective (Person, Graesser, Kreuz, Pomeroy, & TRG, 2001).

Functional Description and System Architecture

ATEC is a web-based application that users can log onto from any computer with an Internet connection and a browser (with Flash and Java). A session with ATEC begins in the same way as the TLAC program in the classroom setting, with viewing of a vignette. After the vignette has been viewed, the tutorial dialog begins with the virtual mentor asking a question and waiting for an answer. The ATEC system is entirely dialog based with constant exchange of virtual mentor questions/comments (instructional agent) and student responses. Figure 1 shows the functional architecture of the ATEC system, indicating which components are handled by AutoTutor components, iGEN components, or Flash/Java components. The user interface components are implemented as Flash and Java, and the dialog manager and language processing components are derivatives of the AutoTutor system, whereas iGEN handles the domain knowledge and reasoning facilities associated with each vignette, the student model, and performance assessment components.

The Virtual Mentor controls and maintains the list of questions that should be asked (selected from the Curriculum Script), evaluates what knowledge the student has demonstrated (Student Input Evaluation), and maintains a representation of the student s discussion of vignette aspects (Student Model). The Language Processor parses the student input, classifies it into speech act categories, and passes a parsed and tagged representation of the student input to the Student Evaluation process within the Virtual Mentor. Frozen expressions (e.g., I don t understand , Could you repeat that?) are handled directly by the Dialog Manager, while assertions and questions are handled by the Virtual Mentor. The system incorporates a hybrid approach for evaluating student inputs. One type of evaluation uses statistical techniques (Statistical Comparison) for comparing the student input to the expected good answer(s). The statistical evaluation will be supplemented to the extent possible with some Deep Reasoning logic based on tractable aspects of Domain Knowledge (See Meaning Representation section below). Questions will be answered using Domain Knowledge to the extent possible. The Dialog Manager handles the microstructure of the tutorial dialog by specifying discourse markers, dialog moves, and frozen expressions, as well as the logic for constructing dialog moves that are responsive to the student input and its evaluation by the Virtual Mentor (see Dialog Management section below).

Dialog Management

Dialog management in ATEC consists of four

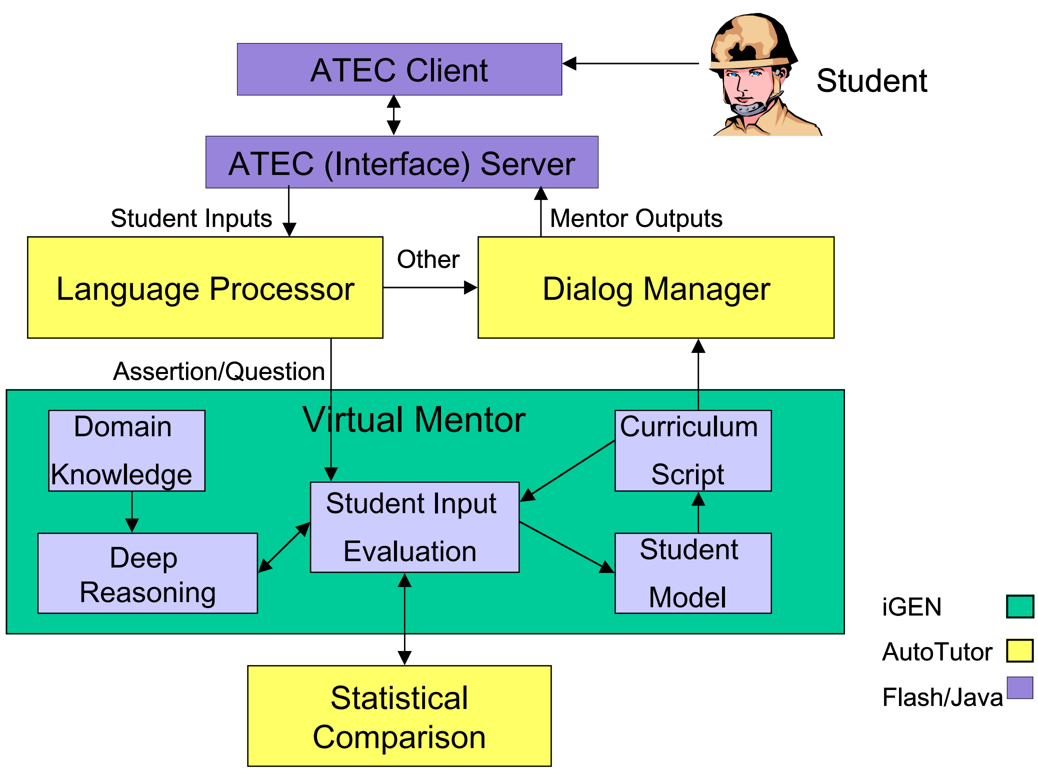

Figure 1. ATEC Functional Architecture

processing components that together should accommodate a smooth conversational interaction and at the same time provide effective training in battlefield tactical reasoning. First there is a hierarchical structure of goals that correspond to the different levels in the curriculum script: Vignettes Themes Expectation Question. That is, a particular Question (Q) is selected from the curriculum script that is designed to elicit the content of a particular Expectation (E) that addresses Theme (T) in Vignette (V). This nested hierarchy provides the core structure for guiding the dialog. Second, there is a Question selector that specifies which question from the curriculum script should be asked next. The question selector starts general and then gets progressively more leading and specific until all of the constituents of an Expectation (i.e., nouns, main verbs, adjectives) are covered. Third, there is a dialog advancer network (DAN) that appropriately reacts to a large set of speech act categories that the learners might type in. For example, learners sometimes generate metacommunicative acts (e.g., Could you say that again?), metacognitive acts (I m lost.), vague expressions (okay) and questions. Each of these categories of student speech acts spins a different thread of dialog from AutoTutor, as specified by the DAN. The DAN is formally an augmented state transition network. And fourth, there is a set of fuzzy production rules that generate particular AutoTutor moves when particular states are recognized by the system. For example, if the learner types in a lengthy answer that has a high match to many of the Expectations in the curriculum script, then AutoTutor generates a positive short response (e.g., Very good.) and then continues in the DAN.

There were several differences between the dialog management of ATEC and the dialog management of AutoTutor when students learn computer literacy or physics. Unlike AutoTutor, ATEC does not attempt to correct learner misconceptions and give a large number of feedback responses (positive and negative). The conversational ground rules of ATEC are not to correct learners and to converge on an exact answer, but rather to get the learner to reason about battle tactics. Therefore, feedback is sparse and either positive or neutral. Unlike AutoTutor, there are no prompts to get the learner to fill in precise words; instead, the only dialog moves associated with an Expectation are questions (of varying specificity) and assertions. The DAN is very similar between AutoTutor and ATEC but some differences were needed. For example, when the learner asked most questions, ATEC would not directly answer them but put the onus on the learner to answer it (e.g., Well you are the commander, so how would you answer your question?). There are a number of generic dialog moves to shift the burden on the learner to take initiative. Once again, ATEC is more like a conversational prosthesis to get the learner to do the talking, as opposed to an information delivery system. In the future, we hope to expand iGEN by incorporating some of the components of COLLAGEN that tracks beliefs, knowledge, goals and attentional states of the learner.

Meaning Representation

The meaning representation in the current ATEC version has three components. First, there is a Curriculum Script with the hierarchical structure of Vignettes, Themes, Expectations, and Questions. This content consists of sentences in natural language, which is much easier for lesson planners to develop than the preparation of symbolic structured code (i.e., LISP or Prolog expressions). The learner s contributions are continuously compared with the content of the Expectations in order to assess what information has been covered and what information has not yet been addressed. Second, there is an LSA space that was prepared from a corpus of documents on battlefield tactics. The LSA space has 300 dimensions that are used when computing the extent to which one bag of words (such as the words typed in during a learner s turn) is conceptually related to another bag of words (such as the words in one Expectation (E)). Third, there is a glossary of terms and definitions that are used in the military. Whenever a learner asks a definitional question (What does X mean?), then the glossary is consulted and a definition is produced for an entry in the glossary.

Future versions of ATEC will phase out the LSA component (which has an expensive patent) and substitute a hybrid mechanism for computing pattern matches between learner contributions and Expectations. The hybrid mechanism will combine a statistical and a symbolic matcher. The statistical component will compare words in an Expectation (E) with words in the learner s contribution (C), but the words will be weighted by the inverse of word frequency. That is, rarer words are weighted higher than frequent words (such as *the* and *person)*. The symbolic component will represent each Expectation in a structured, propositional code that is native to iGEN. The proposition unit has a predicate (which includes main verbs, adjectives, connectives) and noun-phrase arguments that have thematic roles (such as agent, object, location). Syntactic parsers will be used to assign these structures to the verbal descriptions associated with E and C. The advantage of the propositional symbolic representation is that they support theorem proving and other systematic forms of inference generation. The disadvantage is that they are brittle when there is incomplete information. The statistical component is less brittle so a hybrid mechanism is desired. The exact pattern matching function of the hybrid mechanism has not yet been decided.

Later versions of ATEC will support deeper semantic interpretation and reasoning to the extent that it integrates iGEN. This will be accomplished by identifying approximately 20 categories of *canonical reasoning packages* that are used frequently in battlefield tactics. Examples of such packages are Transporting Troops from Location A to Location B and Taking Control of a Town. Each package has a distinctive set of Questions for the curriculum script, inference rules, and semantic composition. The hope is that these packages can be recognized on the fly and instantiated dynamically during the tutorial session. The same set of 20 packages should be reusable for any specific battlefield scenario that is being considered.

Performance Assessment

The ATEC system faces some interesting assessment and evaluation challenges. ATEC is entirely dialog based, with the goal of expanding the student s critical thinking skills. Responses are not correct and incorrect; rather, the system must assess and evaluate student responses to determine if the student is thinking the situation by fully using relevant aspects of the vignette. The evaluation goal is to provide a useful assessment to allow the student (and supervisors) to see progress. The user s beliefs regarding how well the system works and how natural the interaction feels will greatly affect his or her decision to use the system. We must be careful not to lose the user s confidence in the system by providing inappropriate or incorrect feedback. The issues in performance assessment are determining how to measure performance and how to present the information to the student.

The approach we have taken to assessment is based on the depth of questioning needed in each aspect of each of the eight patterns of expert thinking (themes). Performance is considered better if the student generates a response to a high-level question that incorporates all information relevant to that theme. Performance is deemed poorer to the extent that additional, more specific questions and prompts are needed to elicit responses incorporating the additional relevant information. Following this logic, a student who presents a thorough response to the vignette, a response that touches upon the relevant information with a few probe questions, would be assessed higher than a student who provides the same amount of information but requires more extensive probing or prompting. The Student Model contains a representation of the question hierarchy in the Curriculum Script and the Expectations contained in each. The topics and Expectations covered by the student are tagged. In addition, the depth of questioning and prompting needed to elicit reasoning discussion of the topic are calculated, to enable the Virtual Mentor to provide the assessment.

Presentation of the assessment information is the second issue. Whenever one is planning to conduct assessment or evaluation of a person with a career in the military, there needs to be a serious consideration of how the evaluative information will be used by the military, as well as how the recipient of the information will feel about the evaluation. Since the ultimate goal of ATEC is to encourage Army officers to expand their thinking skills, we decided that the assessment and evaluative feedback should be presented in a non-threatening way. Given the dialog-based nature of the system, we decided to build the assessment and evaluation feedback into the dialog with the student. Thus, at the end of the mentor-student discussion of the vignette, the virtual mentor will tell the student which themes he covered well and what information he didn t consider without detailed questioning.

Interaction Style

The user interface (see Figure 2) consists of a section of the screen reserved for displaying information including the vignette, and supporting documents. At any time during the tutorial interaction, the student can manipulate interface functions and display reference materials, re-play the vignette, zoom in or out on the vignette maps, or access supporting materials while working on the scenario. There are buttons allowing the vignette to be replayed and exited at any time. There are also vignette manipulation buttons that provide the ability to zoom in or out on the different maps of the vignette. There are buttons labeled eight themes , orders , view map , and download docs . The eight themes button brings up a brief description of the eight themes to consider while working on the scenario. The orders button brings up a list of available documents. This list is displayed in the information display section of the screen as is any selected document. The view map button simply sets the information display back to the map view. Finally, the download docs button brings up a list of field manuals that can be viewed on-line. There is a space for a talking head where the image of the narrator of the vignette appears during the scenario presentation, and the mentor appears during the tutoring interaction. Finally, there is a section of the screen reserved for dialog boxes for the mentor output and the student

Figure 2. ATEC User Interface

input. The mentor s questions are presented as text, and the student types his responses. In preliminary tests, officers at the level of potential users were willing to type reasonable amounts of text in response to sample questions, indicating the text input would be sufficient for ATEC.

STATUS AND PLANS

The development of the ATEC system is in progress. The basic structure and function of the system have been determined, and a prototype system has been developed including the user interface and AutoTutor components. An initial Student Model is being expanded. The research focus is currently on determining what aspects of deep reasoning can and should be included, and developing the Virtual Mentor processes to incorporate them. Additional effort is also underway to improve the dialog management logic to provide appropriate mentoring style responses and to incorporate the results of Virtual Mentor reasoning into the dialog management logic. Plans include iterative testing and refinement of the dialog content and style, expanding coverage of the system from one to multiple vignettes, and field user testing.

Ongoing complementary efforts include a related Phase II SBIR with a different ITS approach for battle command conceptual skills (Domeshek & Ross, 2002), and a computer-based program developed by ARI that human instructors are using in the Armor Captains Course at Fort Knox's University of Mounted Warfare (Shadrick & Lussier, 2002).

ACKNOWLEDGEMENTS

This research is supported by SBIR Contract DASW01-01-C-0040 from the Army Research Institute monitored by Dr. James Lussier and Dr. Ted Shlechter. We would like to acknowledge additional contributors to this effort: Vassil Iordanov, Dr. Jean-Christophe Le Mentec, and Dr. Wayne Zachary of CHI Systems; Dr. Natalie Person of Rhodes College; Drs. Max Louwerse and Xiangen Hu of the University of Memphis; and Michael Corrigan and Sita Machuga of Provant. Scott Shadrick and MAJ Bill Rademacher of ARI also provided valuable information and assistance.

REFERENCES

Anderson, J.R., Corbett, A.T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*(2), 167—207.

Domeshek, E. A., & Ross, K. (2002). Automated Socratic tutors for high-level command skills. *Proceedings of the 2002 Interservice/Industry Training*

Simulation, and Education Conference [CD-ROM]. .Orlando, FL: NTSA.

Ericsson, K.A., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance, *Psychological Review, 100(3)*, 366-406.

Graesser, A.C., Gernsbacher, M.A., & Goldman, S. (in press) (Eds.). *Handbook of discourse processes*. Mahwah, NJ: Erlbaum.

Graesser, A.C., Person, N., Harter, D., & TRG (in press). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education, 12*, xxx-xxx.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9*, 359.1-28.

Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine, 22*, 39-51.

Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & TRG (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research, 1*, 35-51.

Jurafsky, D., & Martin, J.H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Landauer, T.K., Foltz, P.W., Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.

Lussier, J.W., Ross, K.G., & Mayes, B. (2000). Coaching techniques for adaptive thinking. *Proceedings of the 2000 Interservice/Industry Training Simulation, and Education Conference* [CD-ROM]. .Orlando, FL: NTSA.

Person, N.K., Graesser, A.C., & the Tutoring Research Group (2002). Human or computer?: AutoTutor in a bystander Turing test. In S. A. Cerri, G. Gouarderes, & F. Paraguacu (Eds.), *Intelligent Tutoring Systems 2002* (pp. 821-830). Berlin, Germany: Springer.

Person, N.K., Graesser, A.C., Kreuz, R.J., Pomeroy, V., & TRG (2001). Simulating human tutor dialog moves in AutoTutor. *International Journal of Artificial Intelligence in Education, 12*, 23-39.

Polson, M., & Richardson, J. (Eds.) (1988). *Foundations of intelligent tutoring systems.* Hillsdale, NJ: Erlbaum.

Rich, C., & Sidner, C. (1998). COLLAGEN: A collaborative manager for software interface agents. *User Modeling and User-adapted Interaction, 8*, 315-350.

Ross, K.G., & Lussier, J.W. (1999). A training solution for adaptive battlefield performance. *Proceedings of the 1999 Interservice/Industry Training Simulation, and Education Conference* [CD-ROM]. Orlando, FL: NTSA.

Ryder, J, Santarelli, T, Scolaro, J., Hicinbothom, J., & Zachary, W. (2000). Comparison of cognitive model uses in intelligent training systems. In *Proceedings of IEA2000/HFES2000* (pp. 2-374 to 2-377). Santa Monica, CA: Human Factors Society.

Shadrick, Scott B. & Lussier, James W., (2002). Think Like a Commander at the Armor Captain s Career Course. *Proceedings of the 2002 Interservice/Industry Training Simulation, and Education Conference* [CD-ROM]. Orlando, FL: NTSA.

Shadrick, Scott B., & Lussier, James W. (In Preparation). *Think Like a Commander: Captain s Edition - Prototype 1.0.* (ARI Research Product). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Shute, V.J., & Psotka, J. (1995). Intelligent tutoring systems: Past, present, and future. In D. Janassen (Ed.), *Handbook of Research on Educational Communications and Technology*, Scholastic Publications.

U.S. Army Research Institute (2001). *Think Like a Commander* [Computer CD-ROM and Materials]. (Available from the U.S. Army Research Institute for the Behavioral and Social Sciences, 851 McClellan Ave, Fort Leavenworth, KS 66072).

Weizenbaum, J. (1966). ELIZA — A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*, 36-45.

Wenger, E. (1987). *Artificial intelligence and tutoring systems*, Los Altos, Morgan Kaufmann, 1987.

Zachary, W. Santarelli, T., Lyons, D., Bergondy, M. and Johnston, J. (2001). Using a Community of Intelligent Synthetic Entities to Support Operational Team Training. In *Proceedings of the Tenth Conference on Computer Generated Forces and Behavioral Representation*. Orlando: Institute for Simulation and Training.

Zachary, W., Le Mentec, J-C., & Ryder, J. (1996). Interface agents in complex systems. In C. Ntuen & E.H. Park (Eds.), *Human interaction with complex systems: Conceptual Principles and Design Practice*. Norwell, MA: Kluwer Academic Publishers.

Zachary, W.W., Ryder, J.M., Ross, L., & Weiland, M.Z. (1992). Intelligent computer-human interaction in real-time multi-tasking process control and monitoring systems. In M. Helander and M. Nagamachi (Eds.), *Design for Manufacturability*. New York: Taylor and Francis.