Understanding the Impact of Intelligent Tutoring Agents on Real-Time Training Simulations

Keith W. Brawner, Heather K. Holden, Benjamin S. Goldberg, Robert A. Sottilare
Army Research Laboratory (Human Research and Engineering Division)
Orlando, Florida
{Keith.W.Brawner, Heather.K.Holden, Benjamin.S.Goldberg, Robert.Sottilare}@us.army.mil

ABSTRACT

Abstract: Over the past two decades, the use of agent-based technology within simulated training environments has increased. Intelligent Tutoring Systems (ITS) technology may include reactive or proactive simulation agents that monitor and support computer-based training without human tutors. Reactive agents are able to provide hints and feedback on trainee performance within static scenarios. Based on the trainee's competency and their progress toward training objectives, proactive ITS use computational methods in real-time to decide when to change content, complexity and/or instructional methods within a training scenario (Niehaus & Riedl, 2009). This paper evaluates the advantages and disadvantages of reactive and proactive agents in computer-based tutoring systems; and discusses design considerations for the use of reactive and proactive agents in training simulations.

Historically, intelligent tutoring agents have been simple, passive observers within simulation environments. These reactive agents monitor the trainee's progress and provide hints or other feedback only when there is sufficient variance from expected norms. Reactive agent actions are often based on simple heuristics or scripted behaviors. This can be desirable if the goal of the training is repeatability. However, reactive agents often know little about the trainee and the training context beyond performance data.

Proactive agents have a higher computational cost in that they need to sense and understand more about the trainee, environment and training context, but are better able to predict trainee needs and adapt both feedback and scenario content. Complex military scenarios (e.g. ill-defined domains like bilateral negotiations) provide the opportunity to use more proactive agent techniques in assessing individual and team performance, and in adapting training scenarios to maintain challenge and flow.

Keywords: Intelligent Tutoring, Simulation, Scaffolding, Feedback, Dynamic Scenarios, Adaptive Training, Intelligent Agents, Automated Scenario Generation

Report Documentation Page			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE **Understanding the Impact of Intelligent Tutoring Agents on Real-Time Training Simulations**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Laboratory,Human Research and Engineering Division,Orlando,FL,32819**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

To be presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2011, 28 Nov ? 1 Dec, Orlando, FL

14. ABSTRACT

Over the past two decades, the use of agent-based technology within simulated training environments has increased. Intelligent Tutoring Systems (ITS) technology may include reactive or proactive simulation agents that monitor and support computer-based training without human tutors. Reactive agents are able to provide hints and feedback on trainee performance within static scenarios. Based on the trainee?s competency and their progress toward training objectives, proactive ITS use computational methods in real-time to decide when to change content, complexity and/or instructional methods within a training scenario (Niehaus & Riedl, 2009). This paper evaluates the advantages and disadvantages of reactive and proactive agents in computer-based tutoring systems; and discusses design considerations for the use of reactive and proactive agents in training simulations. Historically, intelligent tutoring agents have been simple, passive observers within simulation environments. These reactive agents monitor the trainee?s progress and provide hints or other feedback only when there is sufficient variance from expected norms. Reactive agent actions are often based on simple heuristics or scripted behaviors. This can be desirable if the goal of the training is repeatability. However, reactive agents often know little about the trainee and the training context beyond performance data. Proactive agents have a higher computational cost in that they need to sense and understand more about the trainee, environment and training context, but are better able to predict trainee needs and adapt both feedback and scenario content. Complex military scenarios (e.g. ill-defined domains like bilateral negotiations) provide the opportunity to use more proactive agent techniques in assessing individual and team performance, and in adapting training scenarios to maintain challenge and flow.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ABOUT THE AUTHORS

Mr. Keith Brawner is a researcher for the Learning in Intelligent Tutoring Environments (LITE) Lab within the U. S. Army Research Laboratory - Human Research & Engineering Directorate (ARL-HRED). He has 4 years of experience within U.S. Army and Navy acquisition, development, and research agencies. He holds a Masters degree in Computer Engineering with a focus on Intelligent Systems and Machine Learning from the University of Central Florida, and is currently in pursuit of a doctoral degree in the same field. The focus of his current research is in machine learning, adaptive training, student modeling, and knowledge transfer.

Dr. Heather Holden is a researcher in the Learning in Intelligent Tutoring Environments (LITE) Lab within the U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED). The focus of her research is in artificial intelligence and its application to education and training; technology acceptance and Human-Computer Interaction. Dr. Holden's doctoral research evaluated the relationship between teachers' technology acceptance and usage behaviors to better understand the perceived usability and utilization of job-related technologies. Her work has been published in the Journal of Research on Technology in Education, the International Journal of Mobile Learning and Organization, the Interactive Technology and Smart Education Journal, and several relevant conference proceedings. Her PhD and MS were earned in Information Systems from the University of Maryland, Baltimore County. Dr. Holden also possesses a BS in Computer Science from the University of Maryland, Eastern Shore.

Mr. Benjamin S. Goldberg is a member of the Learning in Intelligent Tutoring Environments (LITE) Lab at the U.S. Army Research Laboratory's (ARL) Simulation and Training Technology Center (STTC) in Orlando, FL. He has been conducting research in the Modeling and Simulation community for the past 3 years with a focus on adaptive learning and how to leverage Artificial Intelligence tools and methods for adaptive computer-based instruction. Mr. Goldberg is a Ph.D. student at the University of Central Florida and holds an M.S. in Modeling & Simulation. Prior to employment with ARL, he held a Graduate Research Assistant position for two years in the Applied Cognition and Training in Immersive Virtual Environments (ACTIVE) Lab at the Institute for Simulation and Training. Recent publications include a proceedings paper in the 2010 International Conference on Applied Human Factors and Ergonomics and a poster paper presented at the 2010 Spring Simulation Multiconference.

Robert A. Sottilare, Ph.D. is the Associate Director for Science & Technology within the U.S. Army Research Laboratory - Human Research & Engineering Directorate (ARL-HRED). Dr. Sottilare has over 25 years of experience as both a U.S. Army and Navy training & simulation researcher, engineer and program manager. He leads the international program for ARL's Simulation & Training Technology Center (STTC) and participates in training technology panels within both The Technical Cooperation Program (TTCP) and the North Atlantic Treaty Organization (NATO). He has a patent for a high resolution, head mounted projection display (U.S. Patent 7,525,735) and his recent publications have appeared in the Educational Technology Journal, the Journal for Defense Modeling and Simulation, the NATO Human Factors and Medicine Panel's workshop on Human Dimensions in Embedded Virtual Simulation and the Intelligent Tutoring Systems Conference 2010. Dr. Sottilare is a graduate of the Advanced Program Managers Course at the Defense Systems Management College at Ft. Belvoir, Virginia and his doctorate in modeling & simulation is from the University of Central Florida. The focus of his current research program is in machine learning, trainee modeling and the application of artificial intelligence tools and methods to adaptive training environments. Dr. Sottilare oversees the research conducted at the Learning in Intelligent Tutoring Environments (LITE) Laboratory at STTC in Orlando, Florida.

Understanding the Impact of Intelligent Tutoring Agents on Real-Time Training Simulations

Keith W. Brawner, Heather K. Holden
ARL (HRED)
Orlando, Florida
Keith.W.Brawner@us.army.mil, Heather.K.Holden@us.army.mil

Benjamin S. Goldberg, Robert A. Sottilare
ARL (HRED)
Orlando, FL
Benjamin.S.Goldberg@us.army.mil
Robert.Sottilare@us.army.mil

INTRODUCTION

As simulation technology has advanced, real-time and distributed simulation has become increasingly prevalent. Military interface standards and protocols such as Distributed Interactive Simulation (DIS) (IEEE 1278) and High Level Architecture (HLA) (IEEE 1516) have made it possible to connect a variety of diverse systems in real-time. The natural consequence is the increased development of real-time simulation and training (Adelantado, Siron, & Chaudron, 2010). These simulations address a wide assortment of issues, including flight simulators, undersea operations, ground forces control, and naval exercises. Many of these training systems incorporate tactical, ill-defined, or other non-procedural activities, with their own technical implementation difficulties (Levine, Topor, Troccola, & Pullen, 2009). Traditionally computer-based tutors have not been integrated as part of these systems.

Adaptive training, intelligent tutoring, and agent-based technologies have advanced in parallel with simulation technology. These systems typically obey turn based, reactive, strategies. The newest developments in this area, such as Dynamic storytelling (Niehaus & Riedl, 2009) and learner state assessment (D'Mello & Graesser, 2009), are the beginning of a move towards a more proactive approach.

The advancement of these technologies is an opportunity for intelligent tutors and agents to make use of real-time components in a proactive, rather than reactive, manner. The ability of the simulation to measure real-time performance combined with the capability of intelligent tutoring systems to provide feedback enables the development of intelligently aided training. Rather than discard the lessons learned in the creation of reactive and proactive tutors, the authors look at how they may aid in the development of future intelligent tutoring systems for military and training purposes.

The remainder of this paper will be spent in several ways. First we discuss the historical successes and limitations of computer-based reactive tutoring, followed by the indicated historical successes of human proactive tutoring, how they have been addressed by computer tutors, and the limitations of proactive computer-tutoring. We then discuss the high-level considerations that are seen for the design of a proactive computer tutor, the trends in student competency and cognition, application of team tutoring, possible domains of operation, and how these tutors can be fairly evaluated. We part with a word on the way forward for proactive tutoring and simulation.

REACTIVE TUTORING

Strategies and Successes

Even though it has been shown that one-on-one tutoring is significantly more effective than traditional classroom instruction (Bloom, 1984), constraints in budgets and personnel prevent achieving the learning gains accompanied with individual tutoring. Advances in computer science have allowed for computer-based tutors to instruct trainees with personal attention. The development of computer-based tutors and adaptive training has thus far been costly, with a trend towards single-purpose tutoring, but has ultimately proven to be more effective than traditional instruction (Verdú, Regueras, Verdu, De Castro, & Pérez 2008).

Well-defined domains are uniquely suited to intelligent tutoring and adaptive training. These domains have explicit correct and incorrect answers, steps or processes, and ability to adjust task complexity and demand. The domains of mathematics and physics are traditionally thought of as well defined because of the ease of including computer-generated problems and feedback. These areas have seen significantly increased amounts of learning due to the influence of computer-based training because of their natural extensibility. Reported improvements after the addition of an Intelligent Tutoring System (ITS) to a well defined domain are typically between one half and one effect size. After introduction of an intelligent tutoring system in the hydrodynamics domain, 80% of students scored a letter grade of B or higher compared to 51% previously, with a failure rate decreased by half (Nirmalakhandan, 2007). Other examples, in the domains of mathematics (Hwang, 2003), logic (Yacef, 2005), and computer programming (Guzmán, Conejo, & Pérez-de-la-Cruz, 2007) domains report effect size changes of 1.45, between .66 and 1.05, and .93, respectively.

Limitations and Liabilities

Human or computer-based tutoring has two aspects to its activities. The first aspect is that of managing the material presented to the user in order to present content on the edge of the trainee's current ability. The second aspect is that of managing the trainee's cognitive and emotional state in order to keep the trainee on task, interested, and ready to learn. While previous tutoring systems have been significantly useful for increasing the ability of trainees to learn, they have only addressed the issue of content management, rather than trainee management. These tutors do not manage emotional state, ask questions of the trainee, or other behaviors that have been observed in human tutoring. These tutors perform one of the two primary tutoring tasks, and have experienced corresponding increases of roughly one half of the effect size observed in human tutoring.

While the obvious limitation of reactive tutoring systems is the lack of proactive component, this absence has consequences that are further-reaching. Without proactive strategies, this tutor can only tutor the technical implementation details of the desired task. It reacts solely to planned events, and cannot respond to real-time simulation events which occur outside of the original designers' operating scope. It cannot effectively train more than one person to work closely with each other, as teamwork is a non-technical, non-procedural activity. It cannot train leadership or subjective judgment, other than as a strict procedural process, for the same reasoning. Additionally, each of these tutors is only operable in the original domain of design. These tutors are encoded with knowledge about what to teach, but not with how to teach it, and are difficult to expand to another domain or subject of interest.

PROACTIVE TUTORING STRATEGIES AND SUCCESSES

Human tutors are proactive

Although reactive computer-based tutors have been successful in the past, their actions, at best, are based on learners' cognition and understanding. Proactive computer-based tutors are more closely aligned to the strategies achieved in one-to-one human tutoring. In human tutoring, a central element attributing to the learning process is the tutor's ability to identify and adapt to the learner's affect (i.e. emotions, interest, and motivation). Approximately 50 percent of a human tutor's interactions with the learner are based on affective elements (D'Mello, Taylor, Davidson, & Graesser, 2008). Human tutors are devoted to learners' motivation goals as much as their cognitive and informational goals (Lepper & Hodell, 1989; Woolf, Burleson, Arroyo, Dragon, & Cooper, 2009). Since the connection between emotion and cognition is completely intertwined in regards to decision-making and memorization (Woolf, et al, 2009), human tutors rely on understanding both the learner's initial/real-time cognitive competency and the learner's affect to customize their instructional strategies.

With long-term interaction, the human tutor develops an ideal ―picture" of the learner (i.e. comprehensive trainee model) and is then able to (a) better initialize/adapt instruction, even if the task or domain changes, (b) predict the learner's actions/responses, and (c) provide feedback tailored to learner short-term cognition and affect. The tutor-learner relationship becomes more social and trusting, which influences learners' cognitive and affective development (Kim & Baylor, 2006). Interestingly, in traditional learning theories and environments tailoring to learner affect is often secondary or ignored compared to cognition (Picard, et al., 2004), which is most likely why one-to-one human tutoring is much more effective.

Additionally, a human tutor has the natural ability to maintain the learner's flow during instruction. Flow, or optimal experiences, is the feeling of being in control, concentrated and highly focused, or a match between the challenge at hand and one's skills (Csikszenmihayi, 1990; Woolf et al., 2009). A human tutor inherently discerns how and when to adjust challenge and/or provide intervention and feedback. Thus, the tutor continuously maximizes the learner's ―readiness" to receive instruction.

Current state of the art

While human tutors are inherently adept at judging emotional and cognitive states, computer tutors must rely on interpreted sensor data. If computer tutors can accurately assess state as well as a human tutor, it will be possible for computer-based tutors to emulate the same benefits as one-to-one human tutoring. Although the current state-of-the-art ITSs do not achieve this goal, related research is resulting in a better understanding of how to incorporate more proactive strategies. The overall research problem is divided into two categories: (a) the detection/identification of a learner's cognitive and affective states and (b) the appropriate ITS action (i.e., instructional strategy, feedback, etc.) based on a learner's states.

Detection of a learner's cognitive state is relatively straightforward. Initially, the ITS can use the learner's preliminary competency to decide how to begin instruction. During instruction, the ITS can use the learner's responses and actions to ascertain the learner's current comprehension of the material. The ITS can then make decisions on its next course of action.

Carnegie Learning's Cognitive Tutor, a computer-based tutor targeted to teach math, is able to understand trainee knowledge and problem-solving strategies by using a cognitive model. It monitors the learner's knowledge in real-time and tailors instruction based on continual cognition assessments (Corbett, Koedinger, & Anderson, 1997). ACT Programming tutor is another cognitive tutor developed by Carnegie Learning that is used to help trainees learn how to write short programming in Lisp, Pascal, or Prolog. It also uses a model and knowledge tracing technique to adapt instruction (Corbett, 2001). AutoTutor is another state-of-the-art ITS which reacts to a learner's understanding. AutoTutor is designed to hold a conversation with the learner in natural language. To further enhance its similarity to human tutors, AutoTutor uses an expectation and misconception tailored (EMT) dialogue, which contains a list of anticipated good answers and a list of misconceptions for each main question or problem. AutoTutor adapts its feedback based on the learners' responses using pumps, hints, assertions, corrections, and summarization strategies. It also answers trainee's questions (Graesser, Chipman, Haynes, & Olney, 2005). These tutors perform the task of cognitive assessment very well.

Detecting and integrating a learner's affective state can make ITSs much more proactive, but there is great complexity in the implementation of such strategies. In order for an intelligent tutor have natural interactions with humans, it must be able to recognize affect and possess social competencies (Picard et al., 2004). An affect-sensitive tutor could be used to combine cognitive, affective, and motivational states into an ITS' pedagogical strategies. Since motivational states are often based on affective and cognitive components of purpose-driven behavior (Dweck & Leggett, 1988; Woolf et. al, 2009), such incorporations could increase learner engagement, confidence, interest, and learning (D'Mello, Taylor, Davidson, & Graesser, 2008). While tutors like AutoTutor have defined strategies for teaching material, there is a decided lack of implemented automated strategies to manage emotional state.

Previous research has shown that learner affect can be detected within ITSs (Graesser et al, 2007; Woolf et al., 2009). Conati (2002) developed a system that can track multiple learner emotions during interactions within an educational game. The reliability of this system has been measured using learner self-reports; however, the system cannot support multiple affective states and additional complexities encompassed in a learner's reactions (D'Mello, Craig, Sulllins, & Graesser, 2006). The developers/researchers of Carnegie Learning tutors and AutoTutor are generating frameworks to incorporate affect into their ITSs decision making/pedagogical strategies. Current incorporations include: affect recognition (both short-term and long-term), the use of sensor data for real-time affect monitoring and classifications from machine learning strategies, etc. Related researchers are incorporating real-time flow and directional strategies such as scenario adaptation/story telling (Niehaus & Riedl, 2009). Such strategies provide multiple paths to success. In sum, taking more proactive measures during instruction can bring intelligent tutoring systems one step closer to being as beneficial as human tutors.

Limitations on Proactive Use

The most apparent limitation of a proactive tutor lies not in its' functional components, but in the difficulty in its creation. In order to construct a tutor with proactive strategies, one must first construct nearly all of the components of a reactive tutoring system. This process is time consuming, expensive, and explains why there has yet to be many systems which focus on proactive strategies.

Even assuming that a project has the resources for the construction of additional ITS features, or that there is a desire for additional improvement of an existing ITS, it will require the assistance of additional personnel. Psychologists and educational specialists are needed in order to assess what type of feedback should be provided, when and if it should be given (Razzaq & Heffernan, 2009). Tutors, whether computer-based or human, which interfere with the process of learning can easily produce negative learning gains.

Proactive ITS systems require extra computing resources on several different spectrums. A proactive ITS system will naturally require additional functional processing elements when compared with a reactive system. It requires components which process sensor input, which can include image processing of webcam data. It requires additional processing in making the decision to break or manage trainee activities. If a simulation or simulator is already pressing the computational limit of real-time performance, this additional processing overload has significant consequences.

FUNCTIONAL DESIGN CONSIDERATIONS FOR A PROACTIVE ITS

Implementation Details

In order to design an ITS with the capability to be proactive, there are several considerations that must be taken into account. A training session can be viewed of as a series of trainee interactions and state changes. After each of these events, the ITS must make a decision about the trainee, with regard to all of the previous events. This decision is based on events which can occur in series, such as the import of Learning Management System (LMS) data, or in parallel, such as the continuous feed of performance and emotional state data. Careful selection of hardware and software can allow these issues to be resolved in real-time.

The system must be able to monitor the user in order to effectively respond to their actions. While human tutors naturally assess trainee emotional state through interactions and observations, computer-based tutors may assess trainee state through the sensors available to them. These sensors capture mouse movements, measure physiological data, and make use of interpreted webcam images in order to assess the emotional and cognitive states. The exact composition of an ideal sensor setup is the subject of research (Dragon, Arroyo, Woolf, Burleson, Kaliouby & Eydgahi, 2008) (Arroyo, Cooper, Burleson, Woolf, Muldner & Christopherson, 2009), but will be necessary for successful adaption to the trainee.

An effective human or computer takes those actions and the trainee's knowledge into account in order to perform a state assessment. The trainee's knowledge is assessed in two ways: short- and long-term performance trends. These performance trends and sensor data may be combined to build a picture of the trainee's cognitive state. The trainee's emotional state can be assessed in a similar way via short-term sensory observations and long-term personality trait data.

Based on the trainee's combined cognitive and emotional state, it can be determined whether an intervention is required. If an intervention is required, it can be in the form of a domain-specific intervention, such as a hint, or domain independent intervention, such as a metacognitive prompt. While domain specific hints can aid in the learning of specific material, domain-independent feedback allows for the managing of trainee emotional state or meta-cognition (Rus, Lintean & Azevedo, 2010).

Each of these decisions has been taken into consideration into the architecture of Generalized Intelligent Framework for Tutoring (GIFT) (Sottilare, Holden, Goldberg & Brawner, In Review). GIFT consists of the sensor, trainee, pedagogical, LMS, and domain modules, which provide for each type of decision mentioned above. The sensor module adapts trainee actions and adjusts the trainee module, while the pedagogical module monitors the student situation for situations where action is required. If an action is required, it supplies this feedback directly to the trainee or asks the domain component for a recommendation. In this manner the system retains the ability to reactively adapt in response to individual performance, while still maintaining the ability to proactively manage trainee emotional state.

Short-Term Trends, Long-Term Trends, and Managing the Trainee

It has been shown that emotional state and mood have a significant effect on learning outcomes (D'Mello, Taylor, Davidson, & Graesser, 2008). A proactive ITS must analyze and evaluate the short-term and long-term trends of both the task and the trainee. These trends relate to trainee emotional/mood state and competency/cognitive state, and may be summed up as a few simple questions: ―What is the trainee's knowledge about the current task?", ―Are they adding to their knowledge?", ―What type of mood are they in?", and ―What is their overall mood shift?"

The question of long-term trainee knowledge can be answered via a LMS. This data is additionally stored as part of the LMS at the end of session as a result of the trainee's actions during the session. There is no required real-time design component for this type of information, and the design considerations can be addressed in the traditional way via standardized interfaces and formats.

The problem of assessing trainee knowledge and actions as a function of how they are adding to their knowledge does not inherently require a real-time component. While adaptive training research has done an admirable job adapting the trainee to potential problems when they are ready, there is no technical reason why the system should wait to adjust content. A real-time system component which can monitor trainee work has the ability to assess the trainees' general learning direction before the problem is complete. In this manner, content can be adapted or suggested to the trainee or system when it is clear that the trainee will make a mistake, rather than after the mistake is made. A trainee who is excellent at solving the last half of an hour long task should spend a minimal amount of time on this type of action. Focusing on problem areas and gaps in trainee knowledge allows for the optimization of learning activity.

Long-term trainee mood may be addressed in much the same way as long-term trainee knowledge. Long-term personality traits are unlikely to change from one session to another. These trends can be used in order to establish a baseline trainee emotional state, and can be stored in an LMS or other database system. Much in the same way LMS data is logged about the trainee at the end of a training session, overall training mood may be stored for future use or use in another domain. The transference of these traits may yield general models of trainee emotion-based performance, and is currently not studied in depth. There is not a need to make this data available to the rest of the system after communicating initially with any component requiring it.

Short-term trainee mood data is inherently changing frequently, and cannot be stored in an external system in the same way that content-based LMS data or long-term personality trait data may be. The current mood model must be either constructed for each trainee, or each trainee must be shoehorned into a predefined category of mood. These models are not likely to be transferrable from system to system, if only for the reason that different systems have different sensors and types of tasks. After the long term model is adjusted for the short term observations, a proactive ITS requires a provision to assess the value of the current state with regard to the ideal state. This data must be made available in real time in order to effectively make decisions regarding its use.

Team Training Considerations

Team training, in the technical and social aspects, is highly sought after in the military community where most activities are accomplished by groups (Salas, Cooke, & Gorman, 2010) (Heinrichs, Youngblood, Harter & Dev, 2008). Additionally, it is a task which is inherently linked to the affective state of each team member (Delise, Gorman, Brooks, Rentsch, & Johnson, 2010). However, if a tutor is aware of the emotional state of each trainee, this information can be communicated in order to enhance teamwork. The accomplishments in the individual tutoring domain can readily spread to the assistance of team tutoring.

However, team training poses different technical challenges. Team training via an ITS requires additional bandwidth for monitoring the performance of each trainee. Additionally, there must be a central monitoring station in order to scale content on a team, rather than individual, basis. Other complications rise on the issues of how to manage multiple students' emotional states, and the type of emotional state data which should be communicated to other teammates to aid in teamwork. Just as in individual tutoring, there are open research questions regarding how to communicate this information, when it should be done, how often it should be practiced, and the effect size of difference that it makes.

Applicable Domains

As conveyed above, current reactive intelligent tutoring technologies have shown significant learning gains over traditional training techniques within well-defined instructional domains. However, an issue with leveraging these tools for military wide training is that a large portion of tasks performed by Soldiers do not have structured steps for achieving critical objectives. An essential component of performing ill-defined procedures is knowing when to act and how to balance and resolve conflicting goals (Bratt, Domeshek, & Durlach, 2010). The target of military training is to instill a foundation of principles and values associated with problem-solving and decision making as to enable a trainee to make optimal choices under difficult situations when no response is apparent (Bratt, 2009). Army doctrine on training units places great importance on preparation. Soldiers must be able to anticipate change, recognize opportunities, and understand risks associated with potential actions in complex and constantly changing environments (Department of the Army, 2011).

Because of the nature of ill-defined tasks, there is no set procedure for achieving success. To this effect, monitoring performance alone is not enough to inform a training system that there is a need to intervene. To make this training approach most useful to practitioners, the experience needs to incorporate proactive interventions to ensure optimal learning conditions. Due to open-ended procedures for achieving success, feedback must be designed to facilitate metacognitive awareness among trainees that promotes monitoring one's own performance and assessing potential impact of decisions and actions. For both human and computer tutors, this entails monitoring trainee progress in real-time and initiating proactive feedback and content manipulations when a trainee is classified in a negative learning state. This promotes reflection within trainees and aims to build critical problem-solving strategies which can be used to determine the best course of action. Reflection, in the context of learning, is valuable in building new understanding of events based on intellectual and affective activities as the result of experience (Boud, Keogh, & Walker, 1985).

To facilitate effective training within ill-defined domains, practice environments which allow trainees to execute their acquired knowledge and skills among a variety of situations and conditions are required. Simulated virtual scenario-based training provides a means for performing a range of tasks without putting personnel and equipment directly in harm's way, as well as a means for reducing the load on instructor funding and other resources. An associated issue with this approach is that instructors are still required to monitor progress and provide feedback in real-time and through after action reviews (AAR) if no adaptive component is incorporated. By integrating intelligent tutor technology in scenario-based training systems, these platforms have the potential to reduce the role of the instructor. Feedback and content interventions are initiated based on trainee associated performance in accordance with expert models and by their current assessed cognitive and affective states. Understanding the role of affect in learning is vital for adaptive training to be realizable, and empirical testing of intervention techniques and strategies is required.

It is easier to extend an existing ITS domain to team training than it is to explicitly construct a team training domain from the beginning. Given that team training is an activity which can only take place within a framework of emotional context, a prudently chosen ITS domain should be able to be extended to team training on similar tasks. After proactive ITS techniques designed, implemented, and validated, they can quickly be applied to training several users to accomplish a task cooperatively. This can then be further expanded by allowing for the specialization of activities within the same task. The early choice of area of study dictates further expansion, with a poor choice limiting training groups and types and a considered choice allowing for increased variety of learning.

There remains the question of which domains would be best for early proactive tutoring strategies. A proactive tutor fits best in a domain where traditional training or traditional ITS means have not addressed critical training concepts. In order to maximize the potential impact of a proactive tutor, the implemented domain should contain the following features:

- The domain should not be readily addressable by typical ITS means, such as content scaling.
- The domain should have guidelines, but not strict rules, on the order or composition of activities
- The domain should have readily observable physical/emotional effects, such as high stress levels
- The domain should already be implemented in some training fashion, for reasons of practical evaluation
- The domain should have an implementation as an individual task, and expandable to a team task.

With these guidelines of domain choice, the authors have identified several available domains from which a proactive ITS designer may choose. These include room clearing activities, bilateral negotiations, first aid and medical training, and storylined training for tactical situations. The authors believe that each of these domains presents the opportunities outlined above, while still maintaining the interest of the military community.

Future Research and Evaluation of Proactive ITSs

An important aspect of ITS evaluation is to make sure that the evaluation of the system is fair. The implementation of proactive feedback has begun to be studied in other ITS systems (Hefferman & Razzaq, 2010). This research led to limited results, but looked to evaluate forced, proactive hinting against optional, reactive hinting. The conclusion that learning gain is increased when trainees are not needlessly interrupted is relatively unsurprising. These authors also noted that trainees considered the unsolicited feedback as distracting.

Other authors have realized that there is a fundamental tradeoff between studying via an ITS or another method (Hefferman & Razzaq, 2009). These authors look to see if the time spent studying with an ITS could better be spent in different types of learning activities, and present recommendations for different types of tutoring activities. This type of study is reflective of the correct way in which to view tutoring activities.

Fair evaluation of an ITS should hold time constant, and may consist of comparisons of different ITS systems, ITS strategies, or ITS comparison against the traditional means of instruction. Learning outcomes can either be measured through pre/post test, or retention testing. Standards of ITS comparison are beginning to emerge in the form of effect size calculation (Grubišić, Stankov, & Žitko, 2007). The findings of Hefferman and Razzaq indicate that proactive-only techniques may not live up to the learning gains produced by reactive-only techniques. However, it is the belief of the authors that a combined approach of proactive and reactive techniques may yield significant learning gains when compared with only a single technique type.

CONCLUSIONS

Reactive ITSs have been successfully implemented in practice and hold much promise for the future. This paper has reviewed the success of reactive ITS strategies, as well as their limitations. Research shows human tutors to be proactive by nature, and these types of activities are beginning to be included within computer-based tutoring. Their inclusion is trivial and requires significant research, presents developmental challenges, and incurs computational demand. Despite this high cost, this technology is of interest to a number of user communities, and fair evaluations of effect size are necessary to know the impact of system development on learning.

Proactive strategies are the next step for computer-based tutoring. Human tutoring has observed an improvement on performance of two standard deviations, while current reactive-only computer-based tutoring currently observes performance gains of roughly one standard deviation. While comparing historical human tutors to current computer tutors is not a fair comparison, the authors believe that it represents the untapped potential of computer tutoring. In order to realize the practical use of proactive tutoring methods, the predictive accuracy of real-time trainee state models (e.g., cognitive and affective) will need to improve. Proactive methods will need the ability to process behavioral and physiological sensor data in real-time (or near real-time) in order to make these types of judgments, while remaining at least as unobtrusive to the learning process as their human counterparts.

REFERENCES

Adelantado, M., Siron, P., & Chaudron, J.B. (2010). Towards an HLA runtime infrastructure with hard real-time capabilities. In *Proceedings of the International Simulation Multi-Conference*, Ottawa, Canada.

Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. In *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 17-24). Amsterdam, The Netherlands: IOS Press.

Bloom, B. S. (1984). The 2 sigma problem. The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4-16.

Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Journal of Applied Artificial Intelligence*, 16, 555-575.

Corbett, A. (2001). Cognitive Computer Tutors: Solving the Two-Sigma Problem. In M. Bauer, P. J.Gmytrasiewicz, & J. Vassileva. (Eds.), *Proceedings of the 8th International Conference on User Modeling 2001 (UM '01*) (pp. 137-147). Berlin: Springer-Verlag.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems . M. G. Helander, T. K. Landauer, & P. Prabhu, (Eds.) *Handbook of Human-Computer Interaction, 2nd edition* (Chapter 37). Amsterdam, The Netherlands: Elsevier Science.

Csikszenmihayi, M. (1990). *Flow: The Psychology of Optimal Experience.* NY: Harper-Row.

Delise, L. A., Gorman, C. A., Brooks, A. M., Rentsch, J. R., & Johnson, D. S. (2010). The effects of team training on team outcomes: A meta-analysis. *Performance Improvement Quarterly*, 22, 53-80.

D'Mello, S.K., Craig, S.D., Sullins, J., & Graesser, A.C. (2006). Predicting Affective States expressed through an Emote-Aloud Procedure from AutoTutor's Mixed-Initiative Dialogue. *International Journal of Artificial Intelligence in Education*, 16, 3-28.

D'Mello, S.K., & Graesser, A. (2009). Automatic detection of learner's affect from gross body language. *AAI*, 23: 123-150.

D'Mello, S., Taylor, R., Davidson, K., & Graesser, A. (2008). Self versus teacher judgments of learner emotions during a tutoring session with AutoTutor. In B. Woolf, E. Aimeur, R., Nkambou, & S. Lajoie (Eds.), *Proceedings of the Ninth International Conference on Intelligent Tutoring Systems (ITS 2008), LNCS* (Vol. 5091, pp. 9 – 18). Berlin: Springer-Verlag.

Dragon, T., Arroyo, I., Woolf, B.P., Burleson, W., El Kaliouby, R., & Eydgahi, H. (2008) Viewing Student Affect and Learning through Classroom Observation and Physical Sensors. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems (ITS 2008), LNCS* (Vol. 5091, pp. 29–39). Berlin: Springer-Verlag.

Dweck, C. & Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612-618.

Graesser, A., Chipman, P., King, B., McDaniel, B., & D'Mello, S. (2007). Emotions and learning with AutoTutor. In R. Luckin, K. Koedinger, & J. Greer (Eds.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED 2007)* (pp. 569-571*).* Amsterdam, The Netherlands: IOS Press.

Grubišić, A., Stankov, S., & Žitko, B. (2007). EVEDIN: A System for Automatic Evaluation of Educational Influence. In *International Journal WSEAS Transactions on Computers*, 6(1), 95-102.

Guzmán, E., Conejo, R., & Pérez-de-la-Cruz, J.L. (2007). Improving Student Performance using Self-Assessment Tests. *IEEE Intelligent Systems*, 22, 46-52.

Heinrichs, L.W., Youngblood, P., Harter, P.M., & Dev, P. (2008) Simulation for team training and assessment: case studies of online training with virtual worlds. *World J Surg,* 32(2),161–170.

Hwang, G-J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, 40 (3), 217-235.

IEEE 1278.1A-1998, Standard for Distributed Interactive Simulation - Application protocols Errata (May 1998).

IEEE 1516.4-2007, Recommended Practice for Verification, Validation, and Accreditation of a Federation an Overlay to the High Level Architecture Federation Development and Execution Process

Kim, Y., & Baylor, A. (2006). A Social-Cognitive Framework for Pedagogical Agents as Learning Companions. *Educational Technology Research and Development*, 54 (6), 569-596.

Lepper, M., & Hodell, M. (1989). Intrinsic motivation in the classroom. In C. Ames & R.E. Ames (Eds.), *Research on Motivation in Education Vol. 3*, (pp. 73-105). New York: Academic Press.

Levine, S., Topor, L., Troccola, T., & Pullen, J. (2009). Integration of Simulations, Battle Command, and Technology. In *Proceedings of IEEE European Simulation Interoperability Workshop.* Istanbul, Turkey.

Niehaus, J., & Riedl, M. (2009). Scenario Adaptation: An Approach to Customizing Computer-Based Training Games and Simulations. In *Proceedings on the 14th International Artificial Intelligence in Education (AIED) Conference, Workshop on Intelligent Educational Games.* Brighton, UK.

Nirmalakhandan, N. (2007). Computerized adaptive tutorials to improve and assess problem-solving skills. *Computers & Education*, 49, 1321-1329.

Picard, R., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D.,…Strohecker, C. (2004). Affective learning—a manifesto. *BT Technical Journal*, 2(4), 253-269.

Razzaq, L., & Hefferman, N. (2010). Hints: Is it better to wait or be asked? In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems (ITS '10), LNCS*, (Vol 6094, pp. 349-358). Berlin: Springer-Verlag.

Razzaq, L., & Heffernan, N. (2009). To Tutor or Not to Tutor: That is the Question. In V. Dimitrova, R. Mizoguchi, B. du Boulay & A. Graesser (Eds.) *Proceedings of the 2009 Artificial Intelligence in Education Conference* (pp. 457-464). Amsterdam, The Netherlands: IOS Press.

Rus, V., Lintean, M. C., & Azevedo, R. (2010). Computational Aspects of the Intelligent Tutoring System MetaTutor. In H.W. Guesgen & R.C. Murray (Eds.), *Proceedings of the 23rd International Florida Artificial Intelligence Research Society Conference* (pp. 522-527). Menlo Park, CA:AAAI Press.

Salas, E., Cooke, N., & Gorman, J. (2010). The Science of Team Performance: Progress and the Need for More . . ., *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 52(2), 344-346.

Sottilare, R., Holden, H., Goldberg, B., & Brawner, K., In Review. *Considerations in the design of modular, computer-based tutors for military training.* Submitted for publication in the *Journal of Military Psychology*.

Verdú, E., Regueras, L.M., Verdú, M. J., De Castro, J.P., & Pérez, M.A. (2008). Is Adaptive Learning Effective? A Review of the Research. In L. Qing, S. Y. Chen, A. Xu, & M. Li (Eds.), *Proceedings of the 7th WSEAS International Conference on Applied Computer & Applied Computational Science (ACACOS '08)* (pp. 710-715). Stevens Point, Wisconsin: WSEAS Press.

Yacef, K. (2005). The Logic-ITA in the classroom: a medium scale experiment. *International Journal on Artificial Intelligence in Education*, 15, 41-60.