Enhancing Performance through Pedagogy and Feedback: Domain Considerations for ITSs

Benjamin S. Goldberg, Heather K. Holden, Keith W. Brawner, and Robert A. Sottilare
U.S. Army Research Laboratory – Human Research and Engineering Directorate
Orlando, Florida
{benjamin.s.goldberg; heather.k.holden; keith.brawner; robert.sottilare }@us.army.mil

ABSTRACT

As computer-based instruction evolves to support more adaptive training, it is becoming increasingly more evident that such programs be designed around an individual trainee's characteristics, rather than focusing just on task performance. In other words, a trainee's state (e.g. how they learn, their affect and motivation) is an important factor in performance and retention. To optimize individual performance in computer-based training Intelligent Tutoring System (ITS) technologies (tools and methods) are combining artificial intelligence (AI) knowledge representations and programming techniques with the intent to deliver instructional content and support tailored to the individual (Conati & Manske, 2009). From a holistic perspective, such tools and methods personalize training by considering an individual's historical data, real-time behavior, and cognitive measures to predicting comprehension levels and affective states (i.e. frustration, boredom, excitement). This historical and real-time interpretation of the trainee is used for concurrent adaptation of pedagogical and feedback strategies within training content.

Several ITS studies within academic settings report significant learning gains among students receiving adaptive ITS support when compared to students in a traditional schoolhouse environment (Koedinger, Anderson, Hadley & Mark, 1997; Kulik & Kulik, 1991). However, the majority of those systems supported domains with well-defined problems that require well-defined solutions (i.e. physics, algebra). With recent trends in virtual scenario-based training in the defense and medical communities, there has been a major push to simulate more ill-defined tasks that require critical decision-making and swift problem-solving. Primary issues associated with ill-defined scenario training are the lack of a suitable design framework, and determining an appropriate level of support/direction through pedagogy and feedback. This paper will compare ITS pedagogical design considerations between well-defined and ill-defined tasks, identify the variables of interest that have the greatest impact on performance and skill acquisition, and present a high-level design architecture.

ABOUT THE AUTHORS

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.

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. 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. Robert Sottilare is the Associate Director for Science & Technology within the U.S. Army Research Laboratory - Human Research & Engineering Directorate (ARL-HRED). He has over 25 years of experience as both a U.S. Army and Navy training & simulation researcher, engineer and program manager. He leads STTC's international program and participates in training technology panels within both The Technical Cooperation Program (TTCP) and 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 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. 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.

Enhancing Performance through Pedagogy and Feedback: Domain Considerations for ITSs

Benjamin S. Goldberg, Heather K. Holden, Keith Brawner, and Robert A. Sottilare
U.S. Army Research Laboratory – Human Research and Engineering Directorate
Orlando, Florida
{benjamin.s.goldberg; robert.sottilare; heather.k.holden; keith.w.brawner}@us.army.mil

INTRODUCTION

Pedagogy and instructional design are fundamental components to successful training implementation. This is true more than ever with the incorporation of computer-based training platforms that promotes and mediates self-directed learning. Pedagogy focuses on adaptive instruction that individualizes training experiences to the needs of a particular learner. Based on knowledge, competency, and state, training is tailored to meet skill level and feedback/interventions are incorporated as tutoring mechanisms to aid in restoring or maintaining a positive learning state. Such methods are being pursued by the military and medical communities to instantiate alternative solutions to expensive and resource straining live exercises. To gain the full benefits of such techniques, design considerations need to incorporate qualities of instructor characteristics that provide real-time performance assessment and feedback. Intelligent Tutoring Systems (ITS) are one such approach that use Artificial Intelligence (AI) knowledge representations with machine learning techniques for the purpose of producing concurrent performance and state (cognition and affect) diagnoses on the individual and team level (Conati & Manske, 2009). However, an issue with this approach is knowing when and how to adapt training content when an individual is classified in a negative "readiness to learn" state.

Much of the research conducted to answer this question involves studying the tactics performed by human tutors. Tutors are found to be most effective because humans are able to adapt and individualize instruction based on the particular needs of a trainee (Lane & Johnson, 2008). This involves an active balance of participation by the trainee with guidance facilitated by the tutor. The driving factors for determining guidance are based on competency exhibited through performance and dynamic state variables (e.g., engagement, frustration, boredom, confusion, etc.) that fluctuate during interaction. The objective is to have a trainee perform as much of a task as possible while a tutor provides constructive feedback aimed to minimize frustration and confusion (Merrill, Reiser, Ranney, & Trafton, 1992).

Though there has been empirical evidence of improved performance among ITSs, the results still do not meet or exceed the effectiveness of human tutors. This is due in part to a human's ability to read and interpret cues linked with affective states associated to cognitive performance. For selection of the most advantageous pedagogical strategies, the system must know how the interacting trainee is feeling cognitively and emotionally to adapt content that matches their current state. To produce computer-based platforms that exhibit the same benefits seen in one-to-one human instruction (Bloom, 1984), the system must be able to make state determinations as well as or better than that of a person. A large amount of experimentation has been conducted incorporating sensor technology that monitors both behavioral and physiological markers believed to be correlated with cognitive and affective states to make this a realization (D'Mello, Taylor, & Graesser, 2007; Berka et al, 2007; McQuiggan, Lee, & Lester, 2007; Ahlstrom & Friedman-Bern, 2006). From a holistic perspective, such tools and methods personalize training by considering an individual's historical data, real-time behavior, and cognitive measures to predicting comprehension levels and affective states. This historical and real-time interpretation of the trainee is used for synchronized adaptation of pedagogical and feedback strategies within training content to maintain appropriate challenge and to curb boredom.

Yet, decisions on how to adapt content based on state assessment follows few standards. With hundreds of theories on instructional design (see Marzano, 1998; Marzono, 2003; Bransford, Brown, & Cocking, 1999), there are no deemed best practices for adaptive instruction. This paper will highlight design considerations based on domain characteristics and will present a high-level pedagogical architecture. The architecture is based on empirical evidence from past studies in the field and theoretical perspectives on instructional design geared for computer-based education. Appropriate instructional strategy selection requires analysis of a number of variables that drive task mechanics. Establishing a defined framework that assists in decomposing task components for instructional design can improve adaptive capabilities and reduce time in transitioning systems to training houses.

Trends of ITS Implementation

ITS research aims to make training environments adaptable to different learning needs and abilities on an individual level. Majority of systems control the user experience through interacting models that correspond with the elements utilized by a human tutor. This includes knowledge about the student (Student Model), instructional strategy selection rules (Pedagogical Model), knowledge about the domain being trained (Domain Model), and knowledge of how to successfully perform domain tasks (Expert Model) (Durlach & Ray, 2011). The type of data fed into these models is dependent on the domain and available sensing technology designed into the platform.

Current fielded applications in academia using ITS technologies have shown significant learning gains over long-established instructional methods, with the best platforms producing an average increase of 1.0 standard deviation over conventional practices (Anderson, Corbett, Koedinger, & Pelletier, 1995; VanLehn, Lynch, Schulze, Shapiro, Taylor, & Treacy, 2005; Koedinger, Anderson, Hadley & Mark, 1997; Kulik, 2003). However, these successful ITS applications are administered within well-defined domains (e.g., math, physics, chemistry), which involve specific procedures for satisfying task objectives. In this context, performance is easily assessable based on models of expert performance. When actions performed delineate from successful routines, feedback and/or content manipulations (change of pace and/or difficulty) are administered to reduce errors and enhance training transfer. Other approaches for well-formed domains that apply artificial intelligence and cognitive science for adaptation include production systems, case-based reasoning, Bayesian networks, theorem proving, and constraint satisfaction algorithms (Shaffer & Graesser, 2010).

Because performance alone cannot accurately gauge overall training effectiveness, new directions are being taken to enhance the capabilities of ITS components. Incorporating mechanisms that can track affective states among individual trainees will improve the diagnostic capacity of such systems to classify emotional responses that may hinder learning (i.e., boredom, frustration, fatigue). Research efforts are looking at better ways to collect data on the trainee for accurate real-time assessment that can be used to tailor training to match strengths and weaknesses. This includes indentifying techniques that can passively gather information while remaining unobtrusive to the individual. With new data streams being fed into the learner model, pedagogical decision functions can leverage this information to understand more about the current state of the trainee. If such a platform can diagnose competence/performance, motivation, and emotional response, intervention selection can facilitate multiple options outside of feedback driven by performance.

FRAMEWORK FOR DOMAIN DEFINITION

There are a number of variables to consider when adapting a training experience, with domain explicates being a major influencing factor. Dependent on the process and complexity, instructional and pedagogical design practitioners need to decompose task actions into basic functional components. This is led by defined objectives training aims to prepare. The process to achieve task objectives is based on the structure of actions required for successful performance, as well as how well the initial state and goal state are specified (Goel, 1995). If the structure of actions follow specific standards and involves the same process in each instantiation, the task components are considered well-defined. In the instance where procedures are ambiguously defined and there are no clear set actions for meeting goal objectives, the domain of interest is considered ill-defined. This is further described as problems that have less specific criteria for determining when an objective has been satisfied and all information required for a solution is not supplied (Simon, 1973). The domain definition must also take into account the complexity of the task being performed. Complexity is comprised of how difficult (easy/hard) the task is to conduct, the type of environment it’s performed within, and the extraneous factors (weather, opposition, visibility, etc.) that may influence its outcome.

Well-defined domains that score high in complexity often times require training of skills outside specific task execution. In the context of military readiness training, exhibiting performance standards is essential. However, independent of their military occupational specialties (MOS), Soldiers are required to demonstrate higher order thinking skills that exhibit the capability of adapting decision-making tactics in unstable environments where situations and conditions rapidly change (US Army Training and Doctrine Command, 2000). The United States Military Academy's (USMA) Center for Enhanced Performance identified the following elements as critical for performance improvement among warfighters: metacognitive awareness, attentional control, goal-setting, stress management, and visualization (Zinsser, Perkins, Gervais, & Burbelo, 2004). Zinsser et al (2004) further state that establishing these competencies empower individual Soldiers to create efficient thinking habits; improve attentional resources for enhanced situational awareness; and provides experience for coping with physical, emotional, and mental responses during high-demanding tasks. Though many MOS tasks may follow well-defined routines, extraneous factors can significantly impact the procedure or environment the task is being performed within. Because of this, personnel must be able to adapt in real-time to ensure objectives are reached. Training these competencies among all Soldiers is vital for an adaptive Force.

Instilling these elements in trainees requires effective and efficient training paradigms. Providing training that achieves efficient acquisition of these elements is not easily administered. Defining performance criteria for such skills is difficult and is based on specific scenario interactions. Because of this ill-defined classification, standards need to be developed that highlight key strategies and adaptation approaches for design practitioners to enhance system support features for maximizing training effectiveness. With a push by the military for a learner-centered approach to training, tools and methods that promote self-directed learning are required (TRADOC, 2011). TRADOC further identifies the need for pursuing adaptive training and tutoring technologies and to develop standards for implementing these capabilities in computer-based platforms. Yet, a number of issues must be addressed in ITS design to facilitate development of the critical competencies associated with desired training outcomes.

With a two dimensional approach to domain definition, instructional strategies can be specified based on the component characteristics identified within the domain designation. Through task analysis, procedure and complexity can be classified and used for formulation of training objectives. Based on competency and experience, specific training objectives are tailored on the individual level to promote efficient progression from novice to expert. As a trainee progresses through initial content, training objectives are adapted to introduce extraneous factors that will influence execution. Time spent on training, progression tactics from basic functional training to skill mastery, repetition and remediation techniques, and technology aids all play an important role in training implementation (Zipperer, Klein, Fitzgeral, Kinnison & Graham, 2003).

DESIGN CONSIDERATIONS DEPENDENT OF DOMAIN

Considerations for pedagogy and feedback are founded on empirical evidence from past studies using computer-based learning environments and classic learning theory literature. Instructional strategy selection and feedback implementation are the variables of interest, with a goal to discern those that have the highest impact on learning outcomes. Distinguishing a list of best practices based on domain definition is premature at this point and requires rigorous empirical evaluations of the following findings, testing their validity across multiple domains. This paper aims to identify the current state of pedagogy and feedback research in the ITS community and to identify the strategies that have shown significant improvements over traditional instruction or computer-based systems that lack adaptive faculties.

The remainder of the section will review considerations instructional designers must reflect on when developing an ITS application. This requires sound design procedures that takes into account all elements associated with a given training event and the interface components used. As mentioned above, learning objectives are strongly tied to domain definition. Based on the categorized domain a given training objective falls within, methods for pedagogical design need to be identified based on the type of actions being performed. Following domain definition based on task procedure and complexity, an instructional designer should follow a routine process to create the training experience. Five rudiments have been identified that must be addressed in adaptive training design:

- Curriculum: Specifies explicit content for training. This involves scenario design through task designation and decomposition.
- Instructional Strategy: Highlights how content will be presented, the pace of training, the types of actions and procedures performed by the trainee, and scenario difficulty/complexity.
- Performance Measures: A challenge associated with training is defining what is deemed as successful performance. Variables must be identified that measure performance

and are congruent with positive training transfer for the specific outcome of interest.

- Pedagogical Interventions: Determines feedback and support considerations associated to the learning objectives and 'readiness to learn' state. This includes manipulations of content and feedback interventions.
- Student Model Data: Highlights specific data that is needed for cueing interventions. Decision functions must also be defined that trigger adaptive interventions (performance data vs. cognitive/affective state determinants). Data fed into model will vary depending on system functionalities.

Each aforementioned element requires awareness during the design phase of an ITS application. The following subsections will review empirical studies of ITS applications within both well-defined and ill-defined domains. The review will highlight instructional and feedback implementation strategies and their impact on learning outcomes when compared to control settings. This effort aims to identify similarities among successful empirically tested systems and to categorize a domain definition with sound techniques for administering adaptive training.

Well-Defined Domains

Instructional strategies of existing ITSs are based on human teaching, informed by learning theories, and are facilitated by technology (Woolf, 2009). Four strategies based on human instruction commonly used in ITSs are apprenticeship training, problem solving, tutorial dialogue, and collaborative learning (Woolf, 2009). Tutors that cater to well-defined domains can incorporate any combination of these strategies. ITSs that use apprenticeship training contain an expert model to track student performance, provide advice on demand, and support multiple avenues to solutions. Such tutors will scaffold instruction as needed, but will often stay in the background having the student be more responsible for their performance. This strategy is ideal for tasks that are best learned by doing. Sherlock is an apprenticeship training-based environment that simulated the structure/function of a complex electronic diagnostic board (Lesgold et al., 1992). Although Sherlock only provided feedback when requested by the learner, the scores of those who used the environment increased approximately 35% when compared to learners who received no feedback (Corbett, Koedinger, & Anderson, 2007).

Problem-solving is another traditional ITS instructional strategy that uses error-handling techniques and production rules to navigate instruction. For example, the Andes physics tutor uses problem definition, physics rules, a solution graph, action interpreter, and a help system to provide procedural, conceptual, and example guidance (Gertner & VanLehn, 2000). This tutor has been shown to increase scores by one standard deviation (Schulze, Shelby, Treacy, Wintersgill, VanLehn, & Gertner, 2000) and one-third of a letter grade (Gertner & VanLehn, 2000). Like most problem-solving tutors for well-defined domains, the Andes tutor uses model-tracing techniques to monitor the students' progress through a problem solution. In model-tracing, the tutor tries to infer the process by which a student arrived to a solution and uses that inference as the basis for remediation. Model-tracing tutors typically contain expert (production) rules, buggy rules (for error handling), a model tracer, and a user interface (Kodagnallur, Weitz, & Rosenthal, 2005).

Although model-tracing tutors have shown to increase student performance, their adaptation and accountability for individual differences is significantly limited. Traditional model-tracing tutors do not allow for new questions or multi-step lines of questioning. However, second/third generation model-tracing tutors are created to better personalize instruction. For example, the Pump Algebra Tutor (PAT) is a problem-solving tutor that includes a cognitive (psychological) model to assess the process of cognition behind successful and near-successful student performance. PAT also uses knowledge tracing to monitor student's learning from problem to problem. This technique identifies students' strengths and weaknesses relative to the cognitive model's production rules (Koedinger, Anderson, Hadley, & Mark, 1997). PAT is used within thousands of schools and has found to improve student performance on standardized test by 15-25 % (Koedinger & Corbett, 2006). Ms. Lindquist, another model-tracing based tutor for algebra, added a 'tutorial model' that provides the capability asking questions and thinking about the knowledge behind the next problem solving step. Ms. Lindquist allows for the ability of engaging in dialog with the learner (Heffernan, Koedinger, & Razzaq, 2008).

Well-defined model-tracing tutors previously mentioned have improved their adaptability to learner cognition and have been shown to increase student performance; however, they still do not cater to natural interactions seen in one-to-one tutoring. Other ITSs, such as AutoTutor (Graesser, Chipman, Haynes, & Olney, 2005), incorporate natural language interfaces that allows spoken dialogue and adaptation. This promotes collaborative inquiry learning, which has been shown to increase student performance and other learner outcomes.

Ill-Defined Domains

To make ITSs effective across a number of domains, focused feedback and scenario adaptations are required that assist trainees in knowledge/skill acquisition. Empirical studies have been conducted looking at varying pedagogical approaches and to view their effect on learning outcomes. The issue with providing real-time feedback is identifying the mechanisms and decision functions that trigger an intervention, and designing environments that guide trainees to optimal interactions without limiting performance. This requires user models that account for the uncertainty, dynamicity, and multiple interpretations of how to execute ill-defined tasks (Lynch, Ashley, Mitrovic, Dimitrova, Pinkwart, & Aleven, 2010).

Pedagogy within ill-defined domains takes many forms, each of which applies theoretical underpinnings believed to maximize training effectiveness. The underlying issue is definitive feedback often given for well-structured domains are difficult to provide in an ill-defined setting (Walker, Ogan, Aleven, & Jones, 2008). Techniques such as model-tracing, expert systems, and constraint-based reasoning are not optimal in this context because they lack the specifications to support tutoring services outside of task performance (Bratt, 2009). Expert systems have the potential to be leveraged for such training, but comparing a trainee's solution against an ideal solution does not always provide an explanation or reasoning of how the result was constructed (Fournier-Viger, Nkambou, Nguifo, & Mayers, 2010). Depending on the nature of the task and the learning objectives training aims to prepare, specialized instructional delivery and feedback mechanisms must be designed to facilitate training in ill-specified problem spaces.

The challenge is defining a solution path that caters to learning critical elements associated with conducting ill-defined tasks. Lynch, Ashley, Pinkwart, and Aleven (2008) proposes solution paths for ill-defined domains are constructed through: (1) multiple characterizations of a problem to specify components and constraints that have been undefined for selection and discrimination among alternatives, (2) experience for adapting to second and third order effects given the context of a specific problem space, and (3) to justify scenario actions taken with concepts and principles linked to training curriculum. The intent of this approach is to develop a trainee's knowledge and reasoning skills so to avoid the worst outcomes when choosing an action response (Bratt, 2009). Development of such skills is through practice, and simulation-based training environments offer a low-cost alternative to running live exercises that are often resource straining.

Simulated training experiences promote the constructivist and experiential methodologies of learning by facilitating a trainee to solve multiple problems of varying complexity across a number of situations (Bratt, 2009; Raybourn, 2007). However, this approach requires significant instructional guidance at times to avoid negative transfer. To enhance the capabilities of simulation-based platforms used solely as practice environments, adaptive intelligent tutors have been added in systems to aid in instilling the critical aspects of performance, to ensure trainees avoid practicing mistakes or displaying misconceptions, and to reduce the role of the instructor (Thomas & Milligan, 2004). A workshop held at the 9th International Conference on Intelligent Tutoring Systems identified the following explicit domains as ill-defined, which require specialized feedback considerations: medical diagnosis and treatment, intercultural relations and negotiations, inquiry learning, ethical reasoning, robotics operation, and object-oriented design (Lynch et al, 2008). Each of these domains requires higher order thinking skills that enable an individual to perform decision-making, problem solving and goal-conflict resolution.

Five tutorial strategies commonly used for development of higher order thinking skills are question prompts, clarification, hints, examples, and redirection (Alvarez-Xochihua, Bettati, & Cifuentes, 2010). Each strategy is intended to support problem-solving situations and promote metacognitive awareness. An ITS designed for training problem solving with cybersecurity personnel incorporated a Mixed-Initiative framework for providing feedback (Alvarez-Xochihua et al, 2010). The framework was applied to a case-based instructional system that allowed execution of problem solving techniques across a number of distinct scenarios. The system was designed as a re-active platform that responds to student requests. Based on state-determinations generalized from interactions prior to the help request, the mixed-imitative component of the ITS decides the format of feedback to deliver from the five strategies listed above. The system is currently being empirically evaluated to assess if the mixed-initiative approach provides relevant feedback.

Another approach to feedback implementation in ill-defined domains is providing adaptive guidance during scenario interaction and during an after-action review

(AAR) for the purpose of promoting reflection. This approach was implemented in ELECT BiLat, a game-based trainer used for teaching and practicing cultural awareness and negotiating skills (Lane, Core, Gomboc, Karnavat, & Rosenberg, 2007). The environment is designed around interactive narrative between the trainee and artificial agents that respond to the exchanges taken by the system user. Feedback is determined by an expert model. Based on the current state of a scenario, the expert model runs a search algorithm that identifies all available action selections, filters out actions not appropriate for meeting objectives, filters out actions previously performed, and identifies the action selections congruent with expert performance (Lane et al, 2007). Based on the trainee dialogue selection, feedback is generated providing either a hint, a positive remark for a good action, or negative feedback with a short explanation (Lane, Hays, Core, Gomboc, Forbell, Auerbach, & Rosenberg, 2008). Study results comparing effectiveness of feedback implementation in comparison to no coaching showed 89% participants who received real-time guidance completed the scenario successfully while only 59% who received no feedback met training objectives (Lane et al, 2008).

Natural dialogue-based tutoring is an additional method for providing real-time feedback in an ITS environment. Systems designed for well-defined domains have shown success incorporating natural language dialogue that tasks a trainee with articulating the reasoning process during scenario execution. AutoTutor, an ITS that teaches introductory computer literacy, showed an increase in performance of 0.5 standard deviations in comparison to learners who received instruction from a text book (Graesser, Wiemer-Hastings, Wiemer-Hstings, Kreuz, & Tutoring Research Group, 1999). The system prompts trainees to provide explanations for 'How', 'Why', and 'What-if' questions. A concern with this approach is the accuracy of dialogue systems and their effect on training performance. A study was conducted with AutoTutor to gage the effect speech recognition errors had on learning, with the results conveying a subtle impact on performance as well as on a participant's emotions and attitudes (D'Mello, King, Stolarski, Chipman & Graesser, 2007).

Natural language dialogue is now being integrated in ITSs geared for ill-defined domains. The EER-Tutor was designed to instruct individuals on database design by utilizing a hands on practice environment that incorporates dialogue interaction with a computer-based tutor (Weerasinghe, Mitrovic & Martin, 2009). Tutor interventions are applied when an error is present and is facilitated by tutorial dialogues. An independent dialogue was designed for each identified error type and feedback was implemented by a rule-based reasoning system. Interventions were designed to facilitate remediation, aid in completing the session and assisting with technical problems, and for helping with the interface components (Weerasinghe et al, 2009). Feedback is also designed to be adaptive based on a learner's current domain knowledge and reasoning skills. The model proposed for this tutor was evaluated by five acting judges to rate the appropriateness of feedback selection and timing. Conclusions from this study supported the models ability for error-remediation in ill-defined tasks and indicated that trainees acquired domain concepts in the natural dialogue sessions.

Because of the uncertainty in performing ill-defined tasks, knowing the appropriate pedagogy to enact is difficult to determine. As can be seen in this paper, the ITS research community is taking multiple avenues to facilitate adaptive instruction across well- and ill-defined domains. Building standards and guidelines for pedagogical intervention selection based on defined training objectives can enhance systems to be interoperable across a number of domains and aid in reducing time on instructional design. This requires a modular architecture that incorporates all ITS components for the purpose of guiding feedback selection.

ITS PEDAGOGY ARCHITECTURE BASED ON DOMAIN CONSIDERATIONS

The Generalized Intelligent Framework for Tutoring (GIFT) architecture (see Figure 1) introduces the components of sensor, trainee, pedagogical, learning management system (LMS), and domain modules. A domain module in GIFT processes feedback requests and scenario changes by adjusting scenario elements or user interface components. It allows for the pass-through of domain independent interventions, and allows the domain to respond to specific hint requests. A domain module can assess feedback either through *a priori* knowledge of the correct answer, having the ability to calculate the correct answer, or through comparison of a built-in expert model, depending on how well-defined a domain is. The architecture supports a hybrid model approach to feedback selection and requires production rules that dictate activation. Through system use the production rules will be iteratively updated based on the data fed into the

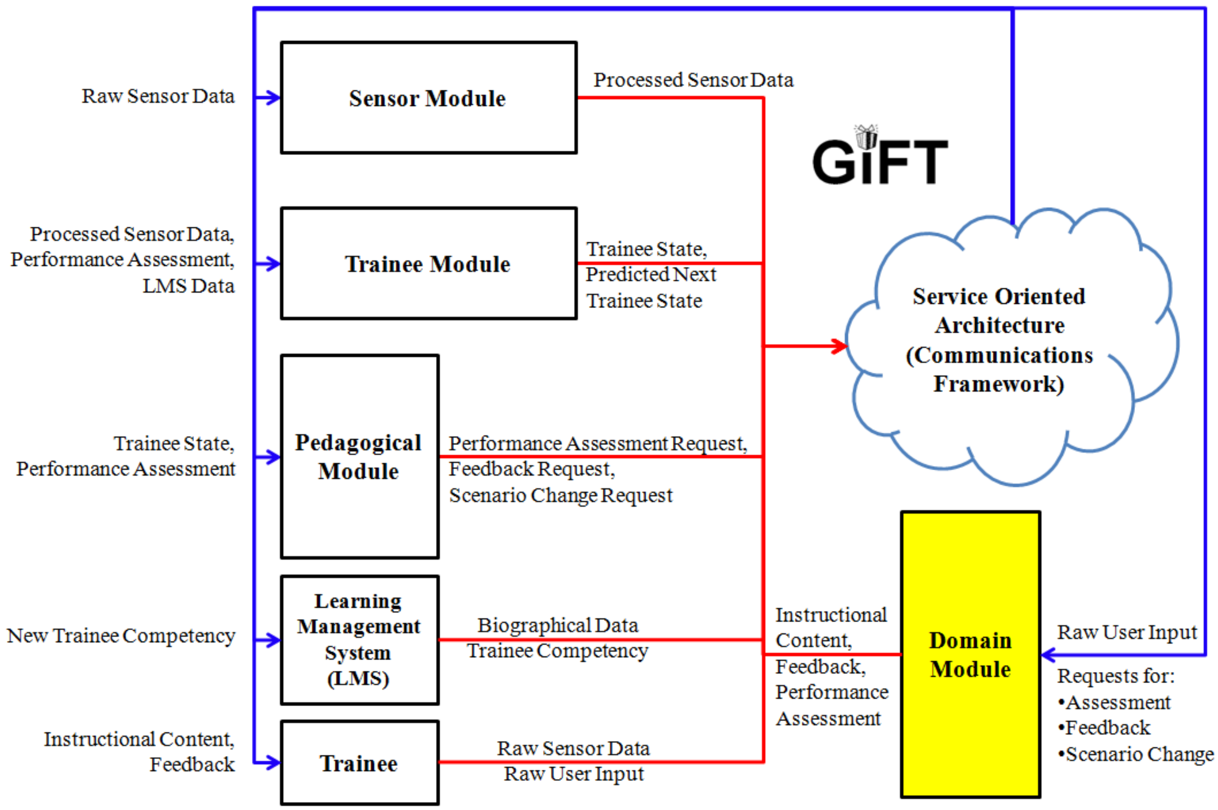

Figure 1: Generalized Intelligent Framework for Tutoring (GIFT)

student and domain models. Variables of performance, competency, affect, and cognition will all be determinants of implementing a training intervention. The architecture will allow for grouping of feedback/manipulation strategies with a triggering variable and supports empirical evaluations to test their effect on training outcomes.

The primary outputs of pedagogical strategy decisions are whether to make an intervention, and the next instructional content which must be presented. While the instructional content decisions are processed out of view of the user, interventions have a sizable effect. Interventions take one of two forms, either that of a domain-specific feedback such as “aim higher”, domain-independent feedback such as an emotional, or a metacognitive prompt. Decisions to change instructional content also come in different forms. Content decisions can modify task demand, modify task complexity, or change the types of content presented. A well-designed domain-specific component must address these things.

In order for pedagogy and feedback to be successful, the architecture requires a domain module that supports the types of feedback requested. While the vast majority of the components of an ITS may be made domain independent, there must always be a specific component of the architecture to deal with the problems that the instructor desires to teach. The fundamental problems of domain dependent components are how to assess student actions, how to respond to instructional changes, how to respond to requests for immediate feedback, and an interface which supports learning (Sottilare, Holden, Goldberg, & Brawner, In Review). The architecture designed must have built-in support for these types of instructional activities.

In a GIFT prototype system for the well-defined domain of addition, this module is being constructed in the following manner. Student action assessment is based on whether or not each digit in a multi-digit number is computed correctly. Feedback generation is handled either through a pass-through emotional prompt or buggy mistake prompt. Complexity is handled by adding or subtracting digits to the numbers to be added, while task demand is changed by allowing less time for the problem to be solved. The user interface is a simple screen with the ability to input added numbers. This is a proof-of-concept system that shows the ease of use of inputting a given domain into a more generalized architecture.

CONCLUSIONS AND FUTURE WORK

While we intuitively know that it is better to have more information when we are making decisions to tailor instructional feedback and content to individual trainee needs, the influence of specific trainee attributes on instructional decisions can be debated. Additional experimentation is needed to quantify the impact of trainee attributes. For example, the importance of personality attributes like openness to performance might differ by task type (e.g., ill or well-defined tasks; individual or collective tasks).

Additionally, implementing sound pedagogy in computer-based training will require accurate state classifications that will determine timing and type. Research is needed to evaluate the influence of macro and micro variables in classifying trainee cognition and affect. Macro variables are generally known at training exercise start and include trainee states like affect (e.g., personality), domain competence, learning preferences and demographic data (e.g., gender, training history).

Micro variables include real-time behavioral and physiological attributes. Behavioral data may be captured by recording interactions within the training simulation (e.g., mouse movement or control selection) or through sensor methods (e.g., motion capture). Physiological attributes (e.g., interbeat heartrates, brainwaves) are generally captured via sensor methods. It will be critical to build validated models of trainee cognitive and affective states using behavioral and physiological measures, but to make the use of these models practical for military training, it will be essential to develop sensor methods that are: low-cost, unobtrusive and portable. Empirical evaluations, validating classification models, and feedback approaches across domains is required for developing standards of feedback implementation in adaptive training. This will enable congruent approaches across domains in terms of pedagogical approaches based on domain definition and state assessment.

REFERENCES

Ahlstrom, U., & Friedman-Bern, F.J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics,* 36(7), 623-636.

Alvarez-Xochihua, O., Bettati, R. & Cifuentes, L. (2010). Mixed-initiative intelligent tutoring addressing case-based problem solving. *Technical Report TAMU-CS-TR-2010-7-2*, Texas A&M University.

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

Berka, C., Levendowski, D.J., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P.D., Craven, P.L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation Space and Environmental Medicine*; 78(5, Suppl.):B231-B244.

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.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bratt, E.O. (2009). Intelligent tutoring for ill-defined domains in military simulation-based training. *International Journal of Artificial Intelligence in Education*, 19, 337-356.

Conati, C., & Manske, M. (2009). Evaluating Adaptive Feedback in an Educational Computer Game. In Z. Ruttkay, M. Kipp, A. Nijholt, & H.H. Vilhjálmsson (Eds.), *Proceedings of 9th International Conference on Intelligent Virtual Agents (IVA '09), LNCS* (Vol. 5773, pp. 146-158). Berlin: Springer.

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

D'Mello, S. K., Taylor, R., & Graesser, A. C. (2007). Monitoring Affective Trajectories during Complex Learning. In D. S. McNamara & J. G. Trafton (Eds.), In *Proceedings of the 29th Annual Cognitive Science Society (*pp. 203-208). Austin, TX: Cognitive Science Society.

D'Mello, S.K., King, B., Stolarski, M., Chipman, P. & Graesser, A. (2007). The effects of speech recognition errors on learner's contributions, knowledge, emotions, and interaction experience. *In Proceedings of the SLaTE Workshop on Speech and Language Technology in Education*. Farmington, PA, pp. 49-52.

Fournier-Viger, P., Faghihi, U., Nkambou, R., & Nguifo, E. (2010). ITS in Ill-Defined Domains: Toward Hybrid Approaches. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring*

Systems (ITS '10), LNCS (Vol. 6095, pp. 318-320). Berlin: Springer.

Gertner, A., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson, & K. VanLen (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems (ITS '00), LNCS* (Vol. 1839, pp. 133-142). Berlin: Springer.

Goel, V. (1995). *Sketches of Thought*. Cambridge, MA: MIT Press.

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

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

Heffernan, N., Koedinger, K., & Razzaq, L. (2008). Expanding the model-tracing architecture: A 3rd generation intelligent tutoring tutor for Algebra symbolization. *The International Journal of Artificial Intelligence in Education*, 18(2), 153 – 178.

Kodaganallur, V., Weitz, R., & Rosenthal, D. (2005). A comparison of model-tracing and constraint-based intelligent tutoring paradigms. *International Journal of Artificial Intelligence in Education*, 6(2), 117-144.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8, 30-43.

Koedinger, K.R., & Corbett, A. T. (2006). Cognitive Tutors: Technology bringing learning science to the classroom. In Sayer, K. (Ed.). *The Cambridge Handbook of the Learning Sciences* (pp. 61-78). Cambridge, MA: Cambridge University Press.

Kulik, J.A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say.* Final Report (SRI Project No. P10446.001). Arlington, VA: SRI International

Lane, H. (2006). Intelligent tutoring systems: Prospects for guided practice and efficient learning. In Army's science of learning workshop. Hampton, VA.

Lane, H.C., Core, M.G., Gomboc, D., Karnavat, A. & Rosenberg, M. (2007). Intelligent tutoring for interpersonal and intercultural skills. *In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC 2007)*, Orlando, FL.

Lane, H.C., Hays, M., Core, M., Gomboc, D., Forbell, E., Auerbach, D., & Rosenberg, M. (2008). Coaching intercultural communication in a serious game. In T.W. Chan (Ed.) *Proceedings of the 16th International Conference on Computers in Education* (pp. 35-42). Jhongli City, Taiwan: APSCE.

Lane, H.C., & Johnson, W.L. (2008). Intelligent Tutoring and Pedagogical Experience Manipulation in Virtual Learning Environments. In J. Cohn, D. Nicholson, & D. Schmorrow (Eds.), *The PSI Handbook of Virtual Environments for Training and Education* (pp. 393-406). Westport, CT: Praeger Security International.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larken & R. Chabay (Eds.), *Computer-Assisted Instruction and Intelligent Tutoring System: Shared Issues and Complementary Approaches* (pp. 201-238). Hillsdale, NJ: Lawrence Erlbaum.

Lynch, C., Ashley, K.D., Pinkwart, N., & Aleven, V. (2009). Concepts, structures, and goals: redefining ill-definedness. *International Journal of Artificial Intelligence in Education*, 19, 253-266.

Lynch, C., Ashley, K., Mitrovic, T., Dimitrova, V., Pinkwart, N., & Aleven, V. (2010). Intelligent Tutoring Technologies for Ill-Defined Problems and Ill-Defined Domains. In *Proceedings of the 4th International Workshop on Intelligent Tutoring Systems and Ill-Defined Domains at the 10th International Conference on Intelligent Tutoring Systems (ITS2010)*, Pittsburgh, PA.

Marzano, R. J. (1998). *A theory-based meta-analysis of research on instruction*. Aurora, CO: Mid-continent Research for Education and Learning. (Eric Document Reproduction Service No. ED 427 087)

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

McQuiggan, S., Lee, S., & Lester, J. (2007). Early prediction of student frustration. In A.C. Paiva, R. Prada, & R.W. Picard (Eds.), *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction (ACII '07)* (pp. 698-709). Berlin, Heidelberg: Springer-Verlag.

Merrill, D.C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences*, 2(3), 277-305.

Raybourn, E.M. (2007). Applying simulation experience design methods to creating serious game –based adaptive training systems. *Interacting with Computers*, 19, 206-214.

Schulze, K.G., Shelby, R.N., Treacy, D.J., Wintersgill, M.C., VanLehn, K., & Gertner, A. (2000). Andes: An intelligent tutor for classical physics. *The Journal of Electronic Publishing*, 6(1).

Shaffer, D.W., & Graesser, A. (2010). Using a Quantitative Model of Participation in a Community of Practice to Direct Automated Mentoring in an Ill-Formed Domain. In C. Lynch, K. Ashley, T. Mitrovic, V. Dimitrova, N. Pinkwart, & V. Aleven (Eds.), *Proceedings of the 4th International Workshop on Intelligent Tutoring Systems and Ill-Defined Domains held at the 10th International Conference on Intelligent Tutoring Systems (ITS '10)* (pp. 61-68).

Simon, H.A. (1973). The structure of ill-structured problems. *Artificial Intelligence*, 4, 181-204.

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*.

Thomas, R. C., & Milligan, C. D. (2004). Putting teachers in the loop: tools for creating and customizing simulations. *Journal of Interactive Media in Education (Designing and Developing for the Disciplines Special Issue)*, 15.

U.S. Army Training and Doctrine Command (2000). *Objective Force capability*. TRADOC Pam 525-66 Draft) Fort Monroe, VA.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Taylor, L., Treacy, D.,…Wintersgill, M.C. (2005). The Andes physics tutoring system: Five years of evaluations. In G. McCalla & C. K. Looi (Eds.), *Artificial Intelligence in Education* (pp. 678-685). Amsterdam: IOS Press.

Walker, E., Ogan, A., Aleven, V., & Jones, C. (2008). Two Approaches for Providing Adaptive Support for Discussion in an Ill-Defined Domain. In *Proceedings of the Workshop on Intelligent Tutoring Systems and Ill-Defined Domains: Assessment and Feedback in Ill-Defined Domains held at the 9th International Conference on Intelligent Tutoring Systems (ITS '08)*, Montreal, Canada.

Weerasinghe, A., Mitrovic, A. & Martin, B. (2009). Towards individualized dialogue support for ill-defined domains. *International Journal of Artificial Intelligence in Education*, 19, 357-379.

Zinsser, N., Perkins, L.D., Gervais, P.D. & Burbelo, G.A. (2004). Military application of performance-enhancement psychology. *Military Review*, September-October, 62-65.

Zipperer, E., Klein, G., Fitzgerald, R., Kinnison, H., & Graham, S. (2003). *Training and Training Technology Issues for Objective Force Warrior. ARI Research Report 1809*. U.S. Army Research Institute for the Behavioral and Social Sciences: Alexandria, VA.