Intelligent Tutoring Methods for Optimizing Learning Outcomes with Embedded Training

Randy Jensen
Stottler Henke Associates, Inc.
951 Mariner's Island Blvd., Suite 360
San Mateo, CA 94404
650-931-2700 (phone)
650-931-2701 (fax)
jensen@shai.com

Dr. Jeffery Mosley
Optimetrics, Inc.
3115 Professional Dr.
Ann Arbor, MI 48104-5131
734-973-1199 (phone)
734-973-1199 (fax)
jmosley@omi.com

Col Mike Sanders
Modeling and Simulation Coordination Office
1901 N. Beauregard St., Suite 500
Alexandria, VA 22311-1705
703-681-7000 (phone)
mike.sanders@us.army.mil

Maj Jason Sims
Army Command and General Staff College
1 Reynolds Ave.
Fort Leavenworth, KS 66027-1352
937-361-5207
jason.sims@us.army.mil

Abstract

The advancing state of the art in virtual and constructive computer based tools provides an opportunity to construct increasingly engaging embedded training. However, even with the benefits of training conducted while immersed in the operating environment and coupled with advanced simulation technology, if the key ingredient of tailored performance feedback is absent then learning outcomes may still be limited. This amounts to one of the central challenges for embedded training, where the goal is to deliver instructional benefits in a setting where the availability of human instructors is limited or none. Within such a framework, structured training methods provide a means to achieve learning objectives and concept retention, with minimal instructor involvement. Intelligent structured training applies real-time automated evaluation and feedback methods based on Intelligent Tutoring System (ITS) techniques. This paper reviews two case studies of embedded training prototypes developed for the U.S. Army which employ structured training methods to optimize learning without direct instructor involvement. These prototypes include a man-wearable trainer for dismounted operations, and a robotic vehicle control station trainer. This paper also summarizes results from selected preliminary experiments which give indicators for the promise of the structured training approach in the embedded setting, particularly with respect to acquisition of procedural knowledge. Experimental results also highlight areas of friction where technology obstacles in the embedded setting make some task areas more “trainable” than others. Finally, this paper gives an overview of potential directions for enhancing ITS approaches with future embedded training development efforts.

Keywords: embedded training, intelligent tutoring, automated assessment, distributed learning

Intelligent Tutoring Methods for Optimizing Learning Outcomes with Embedded Training

The advancing state of the art in virtual and constructive computer based tools provides an opportunity to construct increasingly engaging embedded training. However, even with the benefits of training conducted while immersed in the operating environment and coupled with advanced simulation technology, if the key ingredient of tailored performance feedback is absent then learning outcomes may still be limited. In traditional training settings, human instructors can naturally provide immediate or post-exercise feedback that allows the training audience to understand and learn from their mistakes. Often instructors can also control exercise flow to regulate difficulty levels in tune with the performance of the training participants, to avoid situations of either overwhelming or underwhelming demands that lead to negative training. While these benefits are well documented, this amounts to one of the central challenges for embedded training, where the goal is to approximate the same benefits in a setting where the availability of human instructors is limited or none. Within such a framework, structured training methods provide a means to achieve learning objectives and concept retention, with minimal instructor involvement. Intelligent structured training applies real-time automated evaluation and feedback methods based on Intelligent Tutoring System (ITS) techniques.

This paper reviews two case studies of embedded training prototypes developed for the U.S. Army which employ structured training methods to optimize learning without direct instructor involvement. These prototypes include a man-wearable trainer for dismounted operations, and a robotic vehicle control station trainer. A common theme from the implementation of these prototypes is the need for data exceeding that which is typically available from standard simulation data protocols. Where training objectives include the assessment of trainee decision-making factors, it becomes necessary for evaluation mechanisms to be able to analyze not only the simulation outcomes from trainee actions, but also what specific steps the trainee took in the operating environment which led to those outcomes. Frequently this represents an additional category of human interface data that is not preserved in the simulation data stream. This paper also summarizes results from selected preliminary experiments which give indicators for the promise of the structured training approach in the embedded setting, particularly with respect to acquisition of procedural knowledge. Experimental results also highlight areas of friction where technology obstacles in the embedded setting make some task areas more “trainable” than others. This paper is organized into the following sections:

- Introduction to a state-based approach for implementing automated evaluations
- Review of the Virtual Warrior ITS project, and learner reaction feedback collected in initial experiments
- Review of the Command and Control Vehicle ITS project, followed by experimental results to measure learning
- Summary of potential directions for enhancing ITS approaches with future embedded training development efforts

Automated Evaluation Methodology

The goal for automated evaluation in embedded training is to collect data in real-time from learner interactions with the system, and draw conclusions indexed to specific learning objectives in a manner consistent with how human instructors would perform the same task. As a baseline capability, this can be used to stimulate learning activities, from practice events with real-time coaching, to exercises with after action review. Particularly with free-play virtual environments, it is important to constrain the development task for constructing an evaluation approach. Attempts to model all possible variations of correct performance can be prohibitive in free-play scenarios, and so for such conditions this has led to a state-based approach for identifying training points aligned with learning objectives.

With this approach for embedded training applications involving free action in complex, dynamic scenarios, automated evaluation mechanisms are implemented as Behavior Transition Networks (BTNs), an extension of a Finite State Machine (FSM) approach suited to the complexity of the operational domain. Evaluation BTNs are indexed to relevant principles through a scenario definition interface. This is a generalizable task, which can easily be delegated to an instructor or subject matter expert for training systems with a plurality of exercise scenarios. Thus, when the exercise is executed in the simulation, each evaluation BTN for each principle relevant to the scenario is activated.

The use of BTNs represents a number of enhancements to the traditional FSM model. In particular, the evaluation machines are composed using an authoring and execution tool called SimBionic (Fu, Houlette, Jensen, & Bascara, 2003). SimBionic provides a visual toolset for defining simulation behaviors, with support for multiple chained conditions between states, local and global variables, and the organization of different BTNs in a hierarchically nested structure. This tool can be used to control the behavior of simulation entities, or in instructional applications, to specify the behavior of instructional agents running as virtual entities within a simulation. The run-time engine for these BTNs interprets all active behaviors in parallel and invokes actions and predicates implemented within the application as specified by each behavior. A behavior may be composed of actions, conditions, connectors, and other behaviors. Actions are functions that are invoked to carry out activities, such as sending a message; these often correspond to functionality provided in the API of a given simulation or other environment with which the behaviors are integrated. Conditions perform checks, typically whether an entity is in a specified state, invoking predicates to determine state information from the simulation or testbed. Predicates are functions that return values representing the state information of an entity. Connectors connect actions, conditions and behaviors to specify their relative order of execution.

These BTNs form the structure for defining and executing automated evaluations that monitor states and performance in a virtual training environment. For each simulation entity, a corresponding monitoring entity can be created, which invokes specified evaluation behaviors to be controlled by the behavior execution run-time engine. The notion of underlying forms of knowledge or skills has a counterpart in a hierarchy of principles defined for the domain by the scenario author. In run-time, individual BTNs may perform evaluation functions for either single or multiple principles, depending on whether efficiencies of scale can be realized in terms of data analysis. For example, in a case where two tactical principles would be evaluated via analysis of the same set of source data, then both can be implemented together in the same BTN, even if it produces different conclusions with respect to the different principles in different execution states.

With the BTN-based approach for implementing automated evaluation measures, these evaluations function like instructional agents, identifying conditions that are instructionally significant. For some tactical training domains, there is also value in presenting adaptive opposing forces which can analyze the tactical situation and modify their behavior and tactics accordingly. Particularly in exercises against simulated asymmetric threats, this can help prepare warfighters for the kinds of adaptations that occur on the battlefield. The same kind of agent-based structure described here for instructional evaluations can also be used for adaptive behavior models (Ludwig, Jensen, Proctor, Patrick & Wong, 2008). As a result, triggers from the instructional logic in evaluation BTNs can be used to initiate OPFOR responses to certain conditions or simulation states. When supported by the virtual environment, this can add an extra dimension of realism with the instructional agent BTNs not only monitoring performance, but also directing the behavior of virtual entities through BTNs interfacing with simulation models.

Dismounted Virtual Warrior ITS

Concepts for dismounted embedded training make use of technology advances in helmet-mounted displays, man-wearable computers, and other immersive hardware which can interoperate with operational devices such as handheld Command and Control (C2) systems. Within such a framework, structured training methods provide a means to achieve learning objectives and concept retention, with minimal instructor involvement. The Army Research Development and Engineering Command (RDECOM) Simulation and Training Technology Center (STTC) sponsored a prototype effort to integrate technologies and draw insights about the potential for intelligent tutoring with dismounted operations. First, the effort identified the nature of the data that an integrated structured trainer consumes in order to generate useful real-time feedback for dismounted Soldiers. These data include not only state information direct from the simulation, but also data reflecting Soldier actions in the primary interface and secondary C2 interfaces. Data categories can be generalized and catalogued for future related training efforts. Constructing an operational prototype also afforded an opportunity to collect user feedback in experiments with human participants, providing several indicators for the areas of greatest fit or friction between the dismounted training objectives and a structured training approach.

Dismounted Player Unit Testbed

The Virtual Warrior system (shown in Figure 1) is a prototype embedded trainer which incorporates movement tracking sensors with six degrees of freedom, independently tracking the head, leg, torso, and weapon. These sensors are integrated into the hardware and software interfaces.

Figure 1. Virtual Warrior Prototype

Soldiers experience the virtual environment visually with the use of a helmet mounted display (HMD) which uses the existing bracket used for night vision goggles. Stereo sound is also provided through speakers that have been adapted to the helmet. There are two physical controllers used with the VW; the Weapon User Interface (WUI), and the Soldier Control Unit Interface (SCUI) The WUI is essentially a joystick that allows the user to move forward, backwards, left, right, and toggle through certain tasks such as bore sighting and marking cleared rooms. The SCUI allows the user to view the situational awareness screen, send digital messages, use mouse controls, change posture, and reset the system.

A specially implemented message box was developed as a popup that can appear in the 3D view when the integrated Intelligent Structured Trainer detected conditions meriting immediate feedback. Figure 2 shows a typical view with the popup window present.

Figure 2. Testbed 3D View with Feedback Popup

Although the architecture could support multiple players, scenarios were designed to operate with one human Soldier and SAF behaviors for the other team members. The design of the experimental scenarios with the VW player unit also incorporated the use of the C2 Mobile Intelligent Net-centric Computing System (C2MINCS) into the testbed architecture. This is a dismounted mobile computing platform to provide Soldiers network centric C2 connectivity and situational awareness. Specifically for these scenarios C2MINCS was used to send a variety of pre-formatted tactical reports, including spot reports, position reports, and target handoff reports.

Data Collection Requirements

A common thread in training applications with automated evaluation goals, and particularly in ITS systems, is the need to capture as much data as possible to correctly interpret the decision-making of the trainees. Frequently the cognitive skills required for military operations fall into two categories: conceptual knowledge and procedural knowledge. Conceptual knowledge includes general tactical principles and underlying purposes which provide a framework for applying procedural knowledge. For example, conceptual knowledge includes understanding the tactical principles of clearing a building, and the purposes for tasks like submitting reports. Procedural knowledge is the understanding of the specific sequence of subtasks, such as the proper position within a formation in movement. The major categories of data collection requirements for the implementation of automated evaluations for these forms of knowledge in this domain included:

- Simulation entity state data (e.g., position, orientation, posture, head orientation)
- Operator inputs (e.g., reports sent, vehicle mount/dismount actions, use of controls)
- Line-of-sight data (e.g., calculations for pairs of positions)
- Terrain feature data (e.g., coordinates that lie inside buildings)

This data is consumed in evaluation mechanisms defined in a BTN format described earlier. For a simple example of the data required by a specific evaluation principle, consider the task of evaluating for a designated team member maintaining proper position within a fire team formation during movement. Figure 3 below shows a proper wedge formation for a four person fire team.

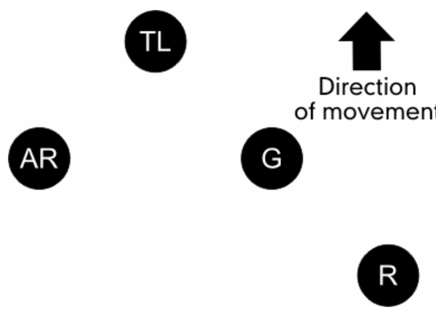

Figure 3. 4-person Fireteam Wedge Formation

For illustration, this example describes a simplified tiered logic that evaluates team member R for proper execution during free-play execution. In the first tier, the evaluation must know when it applies, or more specifically, when the fire team *intends* to be moving in wedge formation. There are many ways to establish this, but a simple method is to simply annotate the terrain. The second tier of the evaluation involves a simple check for relative distance between R and G, without any concern for the positions of other team members, as shown in Figure 4.

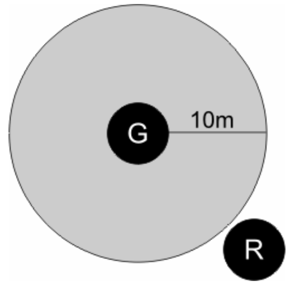

Figure 4. Distance Evaluation

If R is greater than 10m from G, in any direction, then error conditions have been met, and the evaluation triggers a feedback. In Figure 4, R is out of position, being too far from G. However, if this test is passed, then the third tier of the logic is queued. On the assumption that R is within the proper distance radius of G, the next step is to check if R is in the proper section of that radius (area A in Figure 5).

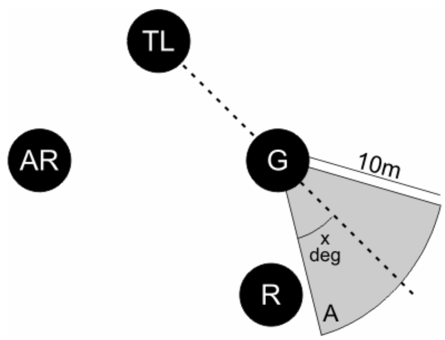

Figure 5. Distance and Position Evaluation

Area A is defined by taking the heading from G to TL, and calculating the space of a pie wedge created by an offset on either side of that heading. In the figure above, R is in a position that would satisfy the second tier test (position within the 10m radius), but fail the third tier test (position outside of area A).

This example contains several simplifications from real-world tactical concepts, and also depends on a specific artifact of how this evaluation would be used in training – namely that the position of G relative to the team leader TL will always be correct because all other entities are controlled by SAF. However, it is useful for illustrating two things. First, data requirements can be simplified with respect to operational definitions based on tactics, techniques and procedures, when the range of use cases establishes constraints that reduce the space of possible interpretations. In the above example, it's not necessary to get an explicit datum for the direction of movement, because this is derived from the SAF positions and movement. Second, the tiered implementation is useful not only for simplifying processing (i.e., when the second tier test is failed, it's not necessary to perform the third tier test), but also for identifying nuances of the conditions in which feedback conditions occurred. Although the same operational principle is violated with a failure of either the second or third tier tests, by distinguishing the conditions in the BTN implementation it becomes possible to generate different (and therefore more informative) feedback content in the two cases. Finally, once the BTN for such a principle is implemented, it can be easily parameterized and abstracted for rapid instantiation in any scenario requiring the same evaluation capability, assuming the same constraints apply.

User Feedback

A key goal for the research effort was to collect feedback from test participants using the VW testbed with the integrated Intelligent Structured Trainer. An experiment was conducted with a specific scenario and set of learning objectives, and used as a basis on which to elicit feedback on usability, suitability, and effectiveness.

The training tasks are primarily individual tasks which the Soldier must accomplish for the collective task to be successful. The dismounted Intelligent Structured Training system evaluates three primary individual tasks. Soldiers must be able to:

- Move as a member of a fire team
- Send reports as appropriate
- Perform movement techniques during the collective task of entering and clearing rooms.

Each of these tasks involves subtasks to be executed at different times and locations in the virtual scenario. For the collective task, the trainer prompts the Soldier on his individual actions within the task. The standards for the task *Move as a member of a fire team* include the evaluated individual observing the proper sector of fire during movement, position in the formation, dispersion in the formation, and maintaining contact with the fire team leader. All of these sub-tasks are critical components to successful completion of the larger task.

Figure 6 shows example feedback for sending an observation report at the proper time, determined from the fact that the enemy vehicle is within the field of view, and therefore the corresponding report should have been sent.

Figure 6. Trainer Feedback for Sending an Observation Report

For the experiments conducted, all test participants held an Infantry Military Occupational Specialty, with previous training in the tasks involved in the scenario. The intent of the assessment of the structured training was to determine if the translation of task standards from the appropriate doctrinal material to the software implementation was a realistic interpretation of the published standard. The data gathered from the participants provides their assessment of the instructional aid to evaluate the performance of the task. The tasks or subtasks that were rated the most favorably were the subtasks of *Enter and Clear a Room,* specifically *Maintaining Sector of Fire in the Room* and *Maintaining Muzzle Awareness*.

The lowest rated tasks were *Enter and Clear a Room* and *Move as a Member of a Fire Team*. All of the tasks with lowest ratings were tasks that required the subject to operate in conjunction with Semi-Automated Forces (SAF) entities to meet the required standard. SAF forces in the OTB simulation can be difficult to work with, and the majority of the Soldiers voiced displeasure with trying to accomplish their mission with the SAF entities. Representative comments about the difficulty of working with the SAF follow:

- “The SAF move too fast in the scenario. I have little time to report before I get left behind.”
- “There is no way for me to communicate with the SAF”
- “The SAF behavior is erratic, they are not consistent from iteration to iteration”
- “The SAF do not follow a logical movement pattern to the target building.”

A common challenge with virtual dismounted training is that existing SAF capabilities may be insufficient to support the operational complexity required in a training exercise. The prominence of user feedback on topics rooted in SAF-related performance also highlights the manner in which such issues can interfere with the primary training objectives.

During the scenario development stage, some realism was compromised to achieve a minimally smooth level of interaction between SAF entities and live players. Most of the participants were wary of using the technology of the system to accomplish their training goals versus using the tried and true live training methods. Most stated that the technology needed to grow, and they would like to have interacted with live players in their team. More recent efforts with more advanced simulations or game engines have the potential to fill this technology gap, in terms of the realism of the environment, and the behaviors of non player entities.

With regard to the usability of the player unit, the majority of the system was rated favorably. Locomotion was the top ranked choice for needed modification. This modification would be the most difficult issue to implement, while still meeting the requirements of an embedded training system. Any device used for locomotion must be small and require minimal accessories to use. The Army has a vision for embedded devices to be used as mission planning and rehearsal tools for combat operations. If the devices associated with the dismounted systems interfere with a unit’s ability to carry the required tools for combat, those devices will not be deployed with the unit.

Command and Control Vehicle ITS

The Army Research, Development & Engineering Command (RDECOM) Simulation & Training Technology Center (STTC) sponsored the development of an integrated architecture linking intelligent evaluation mechanisms with their Command and Control Vehicle (C2V) testbed. This testbed is based on likely embedded virtual simulation common components like OneSAF. In the prototype trainer, automated evaluation methods are applied to monitor simulation and testbed events, and deliver feedback during scenario execution via messages published to the simulation environment. The logic for these intelligent evaluations is captured in hierarchical agent behaviors, and indexed to training principles identified by subject matter experts. In a series of experiments, the improvement in performance over time among test subjects receiving automated feedback was contrasted with subjects receiving other forms of instruction such as an instructor-led after action review.

Overview of Domain and Application

The Soldier manning a Command and Control Vehicle (C2V) crewstation must remotely control robotic platforms to perform reconnaissance and engage the enemy. The operator must maintain situational awareness and apply tactical decision-making principles in a heightened information-rich setting with distributed vehicles and sensors under his command. The increased flexibility allows him multiple ways to detect, track, and identify targets. The following figures show some of the primary screens in the Soldier interface in the C2V crewstation.

Figure 7. Unmanned Ground Vehicle (UGV) Gunner Sensor View. This is one of two 3D visualization screens the operator can use in controlling the UGV remotely.

Figure 8. C2V Robotic Asset Mission Status Tool. This tool is one of the primary control interfaces, where different assets are managed and controlled.

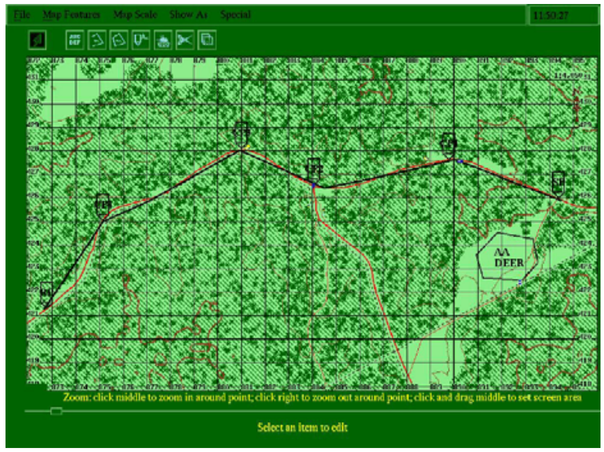

Figure 9. Operational Control Unit (OCU). The OCU provides a 2D view where the operator can create points and routes which will be assigned to vehicles in the simulation.

The ITS technology applied for this project centers around automated evaluation and feedback on operator performance. Example feedback is shown in the following screen from the crewstation.

Figure 10. Immediate Directive Feedback Presented in the Visualization Screen

The automated evaluation mechanisms have the capability to provide feedback either as Immediate Directive Feedback (IDF) in real-time, or as delayed feedback to be incorporated with After Action Review (AAR). With practice in a simulated crewstation environment, students can open a scenario, go through an exercise, and receive customized feedback on their application of the concept of operations. The desired outcome was the training of the robotics operator of the Company Level C2V within a Combined Arms Unit of Action (UA). This was a subordinate task embedded within the larger task of training of the C2V crew, with a focus on the robotics operator. Therefore, the primary intended audience for training is the Robotics officer in his interaction with the company Executive Officer (XO), as well as the Driver and Vehicle Commander, whose secondary duties include the operation of robotic vehicles within the company.

A subject matter expert conducted a task decomposition to identify tasks and actions that other crew members were required to perform to support the Robotics Operator, and additional conditions for the execution of the training scenario. For example, once a target was detected, the Robotics officer was required to select an engagement type, generally consisting of a subset of tasks within the Cooperative Engagement technique. Permission to engage targets would generally flow from the Commanding Officer to the XO. Requests for engagement authority and permission to engage were simplified into a single action, Call for Fire, and automatically handled within the simulation, which accommodates individual training in the embedded setting.

In order to appropriately evaluate for operator performance during an exercise, the training system requires two primary sets of data. First, real-time information about the state of the simulated exercise is necessary in order to assess the conditions under which the student performs different actions. Data in this category includes information about vehicle locations, headings, control status and sensor input, as well as the outcomes of contact with enemy forces. Outcomes can provide useful feedback to the student as to the appropriateness of decisions during an exercise. However, simulation outcomes are not always relevant to student feedback, as a free-play simulation allows for a degree of flexibility where negative outcomes may occur even if the student has performed well on all or most all the relevant principles. The second category of data involves the student's activities within the software and user interfaces. This second category is particularly helpful in diagnosing mistakes that the student has made, since it reveals useful information about the student's intentions.

Experimental Design

Test subjects were given a combination of human tutored and computer aided instructions while occupying the Robotics NCO crew station in a C2V Testbed simulator. This experiment required test subjects to learn both Procedural and Conceptual Knowledge tasks in order to accomplish a tactical mission. Procedural Knowledge tasks included the control of unmanned robotic assets in a virtual environment and learning the correct procedures for conducting: reconnaissance, surveillance and target acquisition tasks, two target engagement techniques, and submitting situation reports. Conceptual Knowledge tasks included learning a defined set of the tactical principles associated with the planning and execution of a tactical reconnaissance mission and their associated supporting tasks. Test Subjects were given a timed tactical scenario which required them to perform a reconnaissance mission in order to demonstrate proficiency in the application of both Procedural and Conceptual Knowledge.

Test subjects were randomly assigned to one of two feedback condition groups: Immediate Directive Feedback (IDF) only and AAR Only (Delayed Feedback). In the Intelligent Structured Trainer developed for this experiment, every action performed by test subjects within the scenario is evaluated to determine if it is the correct response to the current situation. If the response is incorrect, immediate feedback is provided in the form of an error message provided to the test subject. No immediate feedback was provided for following correct procedures. The three types of immediate feedback included:

- Battlefield heuristic feedback – “Conduct a sensor scan before beginning movement”
- Error detection and directive feedback – “Submit a report anytime there is a change to the tactical situation”
- Directive feedback – “You have failed to correctly submit a SITREP. The correct procedure is…”

Immediate Directive Feedback prompts were triggered either when a test subject conducted a procedure incorrectly, or failed to take an appropriate action after receiving an error prompt.

Human tutoring on both Procedural Knowledge and Conceptual Knowledge used open-ended prompts during the AAR to elicit elaboration and self-explanation from the trainees, focused on answering three top level questions: What happened; Why it happened; and How to fix it. It is during the review of tactical principles and the “How to fix it” portion of the AAR that the tutor focuses on Conceptual Knowledge. Conceptual Knowledge includes general tactical principles and definitions of concepts which provide a framework for Procedural Knowledge. For example, Conceptual Knowledge includes understanding the tactical principles of reconnaissance, the various methods for engaging a target, and the purpose for submitting reports; Procedural Knowledge is the understanding of the correct sequence of steps to actually engage a target.

During the AAR, test subjects were required to review concepts and procedures and were then given open-ended, content-neutral prompts to elicit elaboration and feedback. For example, one measure of Conceptual Knowledge required the test subject to define the term “Cooperative Engagement” and give an example. As part of the AAR, the test subject reviewed the definition of Cooperative Engagement and was asked the following questions:

- “Can you explain this concept in simple terms?”
- “Can you give an example?”
- “When have you done something like this?”

If, in answering these questions, the test subject provided an incorrect answer, the tutor identified the answer as incorrect and asked the test subject to try again.

All test subjects were administered an un-timed paper and pencil pre-test to establish a baseline of subject knowledge. Each training trial began with a review of the procedures and training tasks to be accomplished during the training exercise. After this review, each test subject executed a 30 minute timed training scenario that measured their ability to correctly apply these concepts and conduct these procedures in a manned simulator. At the end of each training trial, the test subjects assigned to the IDF Only feedback condition group conducted a self-paced review of the concepts and procedures and completed a subject knowledge test. The IDF Only feedback condition group received no delayed feedback in the form of an AAR. Test subjects assigned to the AAR Only feedback condition group received a human facilitated AAR and then completed a subject knowledge test.

During phase two, a randomly assigned, un-timed subject knowledge paper and pencil test was administered to measure retention of knowledge. Test subjects were then given a transfer task which was identical to the previously learned concepts about the task, Conduct a Route Reconnaissance, and included the same goal, constraints and options for completing the tasks, but the terrain was different than the terrain used during the training tasks. Test subjects had to plan a route on a paper map and identify concepts and procedures when answering questions about the task. Test subjects then re-occupied the C2V simulator and executed the transfer task on a provided scenario.

Experimental Results

The overall results of the experiment support the hypothesis that the timing and type of feedback received during training does effect the acquisition, retention and transfer of both Procedural and Conceptual Knowledge. Significant differences did exist in individual measures, suggesting that immediate directive feedback has a significant effect in reducing the number of errors committed while acquiring new procedural skills during training as well as retention of these procedural skills. Also, delayed feedback, in the form of an AAR which includes open-ended prompts to foster elaboration, has a significant effect on the acquisition, retention and transfer of higher order Conceptual and Procedural Knowledge about a task.

The results of this experiment show that the timing of feedback, immediate and delayed, and the type of feedback, directive and explanatory, has an impact on the acquisition, retention and transfer of knowledge. Immediate feedback to correct detected errors promoted skill acquisition, retention and transfer for procedures. Providing feedback on procedures, whether immediate or delayed, resulted in improvement, as shown in Table 1.

Table 1.
Retention Scores for Procedural Knowledge Following Feedback

Feedback Condition	Mean	SD
IDF Only	3.72	.92
AAR Only	3.80	1.21

However, providing Immediate Directive Feedback significantly reduced the amount of procedural errors committed during training, versus Delayed Feedback, as shown in Table 2.

Table 2.
Procedural Errors Following Feedback

Feedback Condition	Mean	SD
IDF Only	28.22	19.50
AAR Only	42.08	24.59

An analysis of variance (ANOVA) was performed on the data, following the practice of using a combination of between-subjects and within-subjects variables. For this analysis, the between-subjects factor is Feedback Conditions (IDF Only and AAR Only), and the within-subjects factor is Training Trials. The results of the ANOVA found a significant main effect for Feedback Conditions ($F(1,18) = 5.87$), and Training Trials ($F(2,24) = 13.05$), with a $p < .05$. The ANOVA results tell us that there are significant differences in the number of IDF prompts triggered by each test subject and that these differences are explained by both the number of Training Trials (learning occurring over time) and by Feedback Condition, IDF Only and AAR Only. The significantly lower number of error prompts triggered during the execution of the training and transfer scenarios demonstrates that the IDF Only Feedback Condition had a significant effect on the acquisition and transfer of the performance of procedures. This supports earlier studies on the benefits of immediate and directive feedback (Anderson, Corbett, Koedinger, & Pelletier, 1995; Buzhardt and Semb, 2002; Dihoff, Brosvic, Epstein, & Cook, 2004; Guthrie, 1971; Kulik and Kulik, 1998).

Delayed feedback, like that used in the AAR Only Feedback Condition group, promoted retention of new Conceptual Knowledge better than the IDF Only condition, as shown in Table 3.

Table 3.
Retention Scores for Conceptual Knowledge Following Feedback

Feedback Condition	Mean	SD
IDF Only	3.10	2.03
AAR Only	5.60	2.12

Providing feedback on performance, whether immediate or delayed, directive or explanatory, resulted in improvements in the acquisition, retention and transfer of knowledge. These results should prove useful in the pursuit of future embedded training strategies.

Future Applications

Distributed Learning Interoperability

Distributed learning initiatives share a common goal with embedded training, in the aim of delivering training in a variety of settings where it can be of greatest benefit to warfighters. However, advanced forms of training with virtual environments and self-contained intelligent tutoring technologies have traditionally not been available in distributed learning curricula. This has largely been due to the limitations of browser based delivery methods with models such as SCORM (Sharable Content Object Reference Model). More recently, emerging concepts in the distributed learning community provide mechanisms that can be used to construct interoperability with such trainers. Notionally, a SCORM course could incorporate training events on a variety of available platforms, such as an embedded training system or an immersive virtual system at a training facility, which can exercise equivalent learning objectives.

Subsequent to the development of the C2V ITS described earlier, the Joint Advanced Distributed Learning Co-Lab sponsored a prototype effort to construct a mechanism for SCORM interoperability with this same trainer. As a standalone trainer, the C2V testbed cannot be integrated with a browser based learning environment in the traditional manner. Therefore this required the construction of an interoperable architecture, including a mechanism to configure and launch training events based on the instructional sequencing inputs from a SCORM course, while aiming to provide a simple transition for the learner. Performance results compiled automatically and internally in the C2V ITS, are represented in a format that can then be relayed to populate SCORM learner profiles at exercise conclusion. This lays the groundwork for incorporating embedded training events into the experiential component of a distributed learning curriculum, for a wide range of training systems.

Stryker Vehicle Embedded Training

While embedded training is a key transformational requirement for many current and future ground combat vehicle developments, only one fielded embedded training capability exists in today's Army on the family of Stryker ground combat vehicles. This capability was initially an embedded virtual gunnery simulation for the Stryker Remote Weapon System. To extend the utilization of embedded training in the Stryker vehicle requires the expansion of the current capabilities, including feedback and coaching with emphasis on guided practice and reduced coaching. These are essential to moving the Stryker operator from a novice to an expert level of performance on the vehicle. Coaching with on-demand or intelligent tutoring recommendations for corrective or alternative actions supports deeper understanding and better leader decisions in the framework of established doctrine and tactics. These capabilities are being designed into the training management system (TMS) for Stryker embedded training. The TMS will be an integrated compilation of solutions that includes a content authoring system, a learning management system capability, and a performance evaluation tool. This compilation of solutions provides for a flexible and scalable solution for improvement of Stryker embedded training, and is the motivation for a current effort applying and extending ITS methods discussed in this paper. The need to provide IDF and AAR feedback to Stryker operators, with appropriate content, format, and timing, is a current research focus for the vehicle embedded training program.

Areas of Potential Further Enhancement

In the process of implementing ITS methods with embedded training applications, a number of focus areas were identified for potential work expanding on initial findings.

- Develop automated exercise pre-brief or post-brief capabilities to explain the Conceptual Knowledge points within an exercise, to amplify the benefits of immediate feedback for Conceptual Knowledge.
- Automate the AAR protocol to remove humans completely from the feedback process. Include support for trainee self-elaboration during the AAR.

- Abstract principle application conditions, so that evaluations can be applied to different scenarios with minimal adaptation.
- Construct scenario authoring tools to accelerate the development of scenarios that operate as intended on a chosen training environment. Authoring tools can particularly help with the instantiation of evaluation mechanisms with the parameters for a scenario.
- Experiment with alternative or enhanced modes of feedback during exercises.
- Develop enhanced methods for defining and controlling complex OPFOR behavior in scenarios, and coordinate these with instructional evaluation logic.
- Develop capabilities to adjust difficulty levels, either before execution or dynamically during execution in response to the actions of the operator. Ideally, this can be effected through adaptive behaviors controlling non-player entities, and triggered by instructional agents.

Other Intelligent Tutoring System concepts can also be brought to bear, such as student modeling and instructional planning. By combining the performance evaluations across scenarios with information about what knowledge and skills were required for correct performance, an ITS can estimate the mastery of a trainee in each skill and knowledge component. This information can be used to automatically select scenarios that force the trainees to practice the skills and knowledge in which they are the weakest until a specified level of mastery is reached. Scenarios can also be selected to make sure each skill and knowledge component is tested a specified minimum number of times. Additional remedial exercises can be added for trainees who are experiencing continued difficulty in specific areas.

References

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

Buzhardt, J., & Semb, G. B. (2002). Item-by-Item Versus End-of-Test Feedback in a Computer-Based PSI Course. *Journal of Behavioral Education* 11(2): 89-104.

Dihoff, R. E., Brosvic, G. M., Epstein, M. L., & Cook, M. J. (2004). Provision of feedback during preparation for academic testing: learning is enhanced by immediate but not delayed feedback. *The Psychological Record* 54: 207-231.

Fu, D., Houlette, R., Jensen, R., & Bascara, O. (2003). A Visual, Object-Oriented Approach to Simulation Behavior Authoring. *Proceedings of the Industry/Interservice, Training, Simulation & Education Conference (I/ITSEC 2003)*.

Guthrie, J. T. (1971). Feedback and sentence learning. *Journal of Verbal Learning and Verbal Behavior* 10: 23-28.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research* 58(1): 79-97.

Ludwig, J., Jensen, R., Proctor, M., Patrick, J, & Wong, W. (2008). Adaptive Behavior Models for Asymmetric Adversaries. *Proceedings of the Industry/Interservice, Training, Simulation & Education Conference (I/ITSEC 2008)*.