A New Generation of Tactical Action Officer Intelligent Tutoring System (ITS)

Richard Stottler
Stottler Henke Associates, Inc.
San Mateo, CA
stottler@StottlerHenke.com

Susan Panichas
Northrop Grumman Corporation
Middletown
susan.panichas@ngc.com

ABSTRACT

Stottler Henke is developing for the US Navy's Surface Warfare Officer's School (SWOS) a new generation of Tactical Action Officer (TAO) Intelligent Tutoring System (ITS), interfaced to the Generic Reconfigurable Training System (GRTS). The GRTS TAO ITS allows TAO students to interact naturally using spoken language to command and query simulated entities corresponding to other crew members and off-ship personnel. The TAO supervises the utilization of the ship's sensors and weapons and, in general, fights the ship. The majority of the TAO's decisions are manifested by verbal commands and queries. Therefore Stottler Henke is developing the required speech recognition capability to allow the ITS to determine what these decisions are from the spoken words. Those decisions are evaluated for correctness, based on the current tactical situation and performance of other, automated, team members. The TAO's mastery of relevant tactical decision-making principles and ability to apply them in tactical situations is modeled along dozens of dimensions based on the entire history across several scenarios. This student model and the student's immediate performance is used by the ITS to automatically make real-time coaching decisions, assemble a debriefing, choose the next scenario to give the student more practice on his or her weaknesses, and make other instructional decisions.

In the current situation, for simulated scenario practice, one instructor is required for every two students to monitor and evaluate their decisions and to play the roles of other combat team members. The GRTS TAO ITS will be deployed with one instructor for a classroom of 42 students. To allow this, in addition to automating the instructional functions, automated role players (ARPs) to represent the other combat team members are also being developed. Stottler Henke is also developing the required speech recognition for these ARPs so they respond appropriately to the TAO's spoken commands.

ABOUT THE AUTHORS

Richard Stottler co-founded Stottler Henke Associates, Inc., an artificial intelligence consulting firm in San Mateo, California, in 1988 and has been the president of the company since then. He has been the principal investigator on a large number of tactical decision-making intelligent tutoring system projects conducted by Stottler Henke including projects for the Navy, Army, Air Force and Marine Corps. Currently he is working on the GRTS TAO ITS for the US Navy, and a Combined Arms ITS as part of the US Marine Corps Combined Arms Command and Control Training Upgrade System (CACCTUS). He has a Masters degree in Computer Science from Stanford University.

Susan Panichas, Program Manager at Northrop Grumman Corporation, has been the program manager for simulation based training systems for over ten years. She is the program manager for the Generic Reconfigurable Training System (GRTS), which is the foundation for US Surface Navy operator and tactical training programs. She has a Masters degree in Physical Oceanography from the Univ of Rhode Island.

A New Generation of Tactical Action Officer Intelligent Tutoring System (ITS)

Richard Stottler
Stottler Henke Associates, Inc.
San Mateo, CA
stottler@StottlerHenke.com

PROBLEM DESCRIPTION

The mission of the Surface Warfare Officers School (SWOS) in Newport, Rhode Island is to provide professional education and training to prepare officers of the U.S. Surface Navy to serve at sea.

As part of their training, mid-career Surface Warfare Officers learn how to "fight" their ship as a Tactical Action Officer (TAO). The TAO training consists of three months of combined classroom and simulator time wherein students are exposed to all elements of surface warfare; air, surface, subsurface, amphibious, strike, and electronic, as well as support mechanisms. The objective of this training is to prepare the officer to exercise command over the people (watchstanders) who operate the warship's weapons, sensors, navigation, and support systems. This training is also intended to sharpen the tactical decision-making of the skills of the TAOs, enabling them to defend their ship during a potentially hostile situation. The decisions the TAO makes during such situations certainly affects the outcome of the ship's mission and potentially has life or death consequences.

The TAO is supported by a large team of watchstanders in the Combat Information Center (CIC). These watchstanders, the TAO directs to take actions or queries for information verbally, over the internal communication network. The watchstanders in charge of a each warfare area manage the watchstanders under them and primarily act on their own initiative, informing the TAO of what they are about to do. Information and intentions are shared over an internal communications network, and the TAO utilizes command by negation to acknowledge or countermand the stated intentions of their watchstanders. The TAO also performs limited actions at their console, primarily hooking tracks to determine available information, adjusting their tactical display, and closing the Fire Inhibit Switch (FIS) to permit engagements.

The TAO gathers information, analyzes that information, and ensures the correct decisions are made and actions taken based on the tactical situation. A key factor in developing sound tactical abilities is the amount of tutored, tactical decision-making practice the TAO is able to have in realistic, simulated scenarios. To maximize the practice opportunities requires a training system that runs on highly available hardware (i.e. PCs) and includes tactical simulation, automated CIC team members, and automatic tutoring. The automated team members, also called Automated Role Players (ARPs), must not always perform perfectly. Since the TAOs primarily command by negation, if their team members are very good, they have little to do and certainly cannot demonstrate that they have full-grasp of all the important concepts. Therefore, when the student TAO has progressed above the Novice level, the ARPs must purposely make mistakes, both omitting the correct actions and actively committing errors in the form of incorrect decisions and actions.

A final requirement on the training system was that the automated ARP behaviors, automated tutoring behaviors, and scenario descriptions and setup must be relatively simple so that the instructors would always understand what the system was doing and why and what it would do next.

The required simulation and PC hardware were already resident at SWOS and ready to be utilized for this training system. The Generic Reconfigurable Training System (GRTS), developed by Northrop Grumman, had a TAO version which faithfully simulated the TAO console on a standard PC and included a simulation of the naval tactical environment. The system was already used at SWOS to train TAO students in console operation, though an instructor was needed for every two students to play the role of other CIC team members and provide tutoring. SWOS had already set up an electronic classroom that included 42 student PCs for viewing electronic materials networked to a single instructor console. The problem was that for 42 students to concurrently run GRTS scenarios would require 21 instructors. This was the motivation for the development of the GRTS TAO ITS and associated ARPs, so that only one instructor would be required for the 42 TAO students while they practiced tactical decision-making in a realistic manner (verbal commands to ARPs and console actions in GRTS's simulated TAO console).

CHALLENGES

There were several challenges that needed to be addressed by this training system. It must be able perform evaluation of the student's decisions in realistic, free-play simulations. This was further complicated by the fact that ARPs would be performing at varying levels of ability and the tutor would be providing various levels of hinting.

A second challenge was automating CIC team members such that they would behave realistically in a variety of tactical situations. This included responding to spoken commands and queries (speech recognition), taking into account the current tactical situation, and remaining responsive in the face of unexpected or overly proactive TAO requests.

Automatic speech recognition is a always challenging however for this application it was more critical than usual, The ITS would largely base it assessment of the student's performance on what was output from the speech recognition system. Even a small percentage of errors would frustrate the students if they were receiving feedback not based on what they said but based on what the system thought they said. Similarly, the ARPs also use the speech recognition output. Errors in this output would cause the ARPs to respond in a mysterious (to the student) manner.

Just as the primary input mechanism for the system as a whole is the student's TAO verbal utterances, one of the primary outputs from the ARPs is also verbal utterances. These must be relevant to the tactical situation and TAO requests, correct in choice of language, understandable, and not in conflict with each other or the TAO's communications.

In the absence of a human tutor, the ITS must make several instructional decisions for each specific student including whether to provide hints, the content of the debriefing, next practice scenarios, and its current estimate of the ability of the student to apply each of dozens of principles to tactical situations. And these decisions needed to occur in a way that was simple enough that instructors could easily understand the process and predict the decisions the ITS would make.

This simplicity and understandability challenge extended to exercise authoring. Instructors need to be able to create new scenarios in GRTS (initial tactical setup, environment, and initial orders for hostile, neutral, and commercial platforms) and then easily add the information required for ITS and ARPs to function correctly.

A final challenge related to the classroom setup. The instructor needed to be able to quickly and easily monitor the performance of all 42 students simultaneously so that if particular students were having a lot of difficulty the instructor could intervene.

HIGH LEVEL OVERALL SYSTEM DESCRIPTION

Below is the high level architecture for the GRTS TAO ITS. Except for the Instructor console, this setup is duplicated 42 times, for each student PC. The student primarily interacts with the GRTS TAO Console Simulation and speaks over and listens to the Audio networks. Through an interface to GRTS, the Automated Role Players (ARPs) and ITS monitor the tactical situation and the student's actions. Through an interface to the audio networks the ARPs and ITS receive the TAO's verbal orders and queries and the ARPs speak to the TAO and each other (for the TAO's benefit). The ITS also sends a summary of the student's performance to the instructor's console. Independently developed Interactive Multimedia Instruction (IMI) can also be linked to for more detailed explanations of principles that the student is having problems with. The ITS is based on the FlexiTrainer ITS development tool and the ARPs were developed using SimBionic.

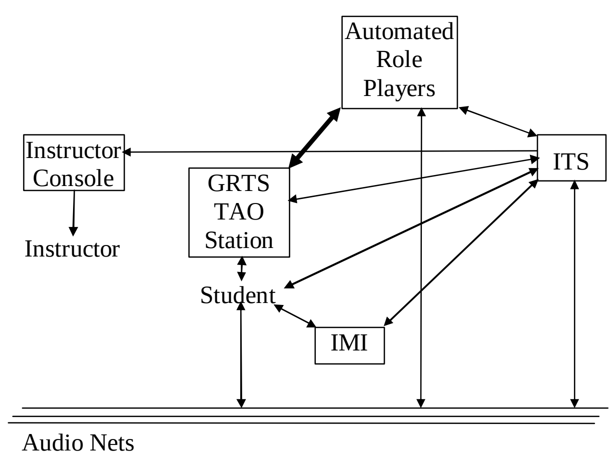

Figure 1. High Level Architecture

GRTS DESCRIPTION

Figure 2 shows the GRTS simulated TAO console. It includes panels for Variable Action Buttons (VABs), display selection (map control keys), radio control, tactical situation map (a scaled version of the large screen display), and Automatic Status Boards that, among other things, display information on the hooked track. The mouse is used to push buttons and select tracks. These displays are driven by a tactical simulation that simulates ownship's sensors and weapons, external platforms, and the environment. GRTS simulations are initialized from a GRTS scenario file created using the graphical GRTS scenario editor. The GRTS versions currently used by SWOS replicate AEGIS system consoles while SWOS trains TAOs for a variety of ship types. Additional GRTS versions for these additional ship types are also being developed.

GRTS also needed to transmit data to the ITS and ARPs including the tactical situation (locations and actions of all platforms in the scenario), the TAO's console actions, and the data the TAO was being provided through the console to make his decisions. This was primarily the list of tracks that were currently being displayed. This information is provided by the GRTS ITS Interface. In developing their interface, Northrop Grumman consulted the ITS/Simulation Interoperability Standard (I/SIS), a draft SISO standard described in [Stottler, et al. 2005]. Similar to many ITS-Simulation systems, the simulation also provides a display mechanism that the ITS uses for real-time hinting and feedback. In addition to data needed by the ITS, I/SIS also describes required ITS-to-simulation transmissions.

Figure 2. GRTS Simulated TAO Console

ITS DESCRIPTION

Automatic Evaluation

The first challenge to overcome was the requirement to evaluate the TAO's decisions and actions in free-play simulations. The inputs to TAO performance evaluation are the simulated tactical situation, the data (e.g. tracks) visible (i.e. detected and displayed) on the simulated TAO console, verbal utterances (and their electronic counterpart) from the ARPs, his verbal commands and queries, and his console actions. The tactical situation, visible tracks, and TAO's console actions were transmitted to the ITS by GRTS. The TAO's verbal commands and queries are sent from the speech recognition and utterance editor (described further below). The evaluation behaviors also use the parameter data included with each scenario such as query distances, warning distances, etc.

The evaluation challenge was addressed using Behavior Transition Networks (BTNs). BTNs are similar to Finite State Machines (FSMs). An FSM is simply a network of states with specific transitions between particular pairs of states, where each transition has a from-state and a to-state. An FSM is in exactly one of its states, the current state, at a time. Associated with each state may be software that executes while the FSM is in that state. Associated with each transition is a condition. If that condition is true when the FSM is in the from-state of the transition, then the FSM will transition to the to-state. An FSM will have one initial current state that it starts in when it first becomes active.

FSMs are useful because the transition conditions can reference simulation events and values, and trainee actions. Typically, for automatic training evaluation, a portion of the FSM is used to monitor events and values in the simulation, looking for a specific type of situation. This type of situation places the FSM in a specific state. Then the second portion of the FSM monitors and evaluates the student's relevant reactions (or lack of them) to this type of situation. Typically, it writes messages to the trainee interface and/or to a log file that will be presented as the AAR that describes why the actions were correct or incorrect.

For purposes of evaluation in realistic free-play simulations, traditional FSMs have been found to be too restrictive and they have therefore been generalized into Behavior Transition Networks (BTNs). BTNs are very similar to FSMs in the sense of having states, transitions, transition conditions, and a current state, but BTNs have additional capabilities. For example, BTNs have variables that are automatically bound to the events and other conditions in the transition. These variables are easily passed between states and transitions and even across BTNs. The best way to employ BTNs to monitor real-time mission execution is to have a large number operating in parallel where each looks at the situation and student's actions from the perspective of how they handle specific types of situations or apply specific types of principles. [Stottler 2003] describes BTNs and their use in performance evaluation in more detail.

A simple example BTN is shown in Figure 3. This BTN evaluates the acknowledgement principle. The BTN starts out in the start state. Whenever the TAO is addressed, the transition is followed to the "TAO Should Acknowledge" state and a new copy of the BTN is created, starting in the Start state. (The fact that following the transition creates a copy is designated by thick oval outline of the transition label.) From the Should Acknowledge state there are 3 possibilities, the TAO correctly acknowledges the communication (and receives credit for passing the principle this time), incorrectly acknowledges, or doesn't verbally respond within 20 seconds. In these latter two cases the student is considered to have failed this application of the Acknowledgement principle.

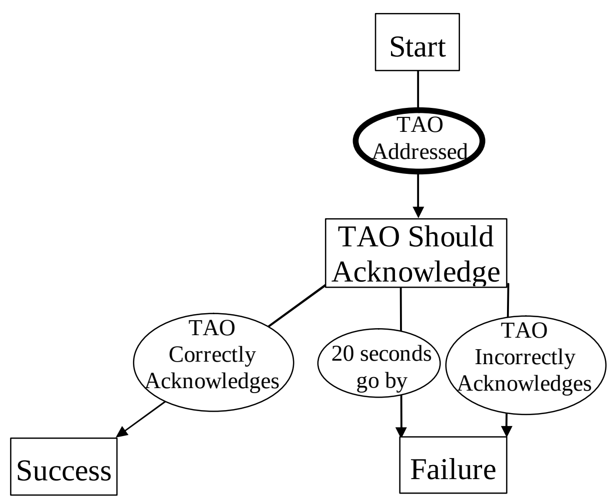

Figure 3. Acknowledgement Evaluation BTN

Student Modeling

The student's mastery is modeled based on their performance in simulated scenarios and their ability to apply 57 different principles (in the first release of GRTS TAO ITS). (The final release will have between 300 and 600 principles). The Student Model receives performance evaluation events from Performance Evaluation BTNs and maintains a record of each principle pass/fail attempt and the outcome. For each Principle a simple formula is executed on this list of attempts and updated immediately after each attempt. The instructor can examine the mastery estimate and attempts-list in real time. The mastery estimate consists of the level designation (novice, intermediate, or advanced) and a real number, between 0.0 and 1.0, which indicates the degree to which the student has mastered the principle in scenarios at that level. Initially the estimate is Novice 0.0. Upon graduation to a new level (to Intermediate or Advanced) the estimate is that level, 0.0. Each new attempt is scored as 0.0 for failed, 1.0 for being immediately successful, and in-between if successful with a hint and averaged in equally with the previous running estimate: new estimate = (old estimate + new attempt score)/2. An estimate above 85% would be considered mastered at that level and allow graduation to the next level.

Real-Time Coaching

The Real Time Coach has two primary purposes. One is to provide hints in situations were the student is not likely to perform well, either because it is a new area or the student has frequently failed previously. The instructional concept is that learning will be more efficient if the student makes the correct decision, even with a hint, than if he or she fails. Preferably the hint should be as vague as possible while still being specific enough to allow the student to perform well. And, of course, the hinting mechanism must be removed when appropriate so that the student can demonstrate independence from it. The second purpose of the real-time coach is to keep the scenario moving along in a way anticipated by the instructor so that opportunities to practice application of tactical principles later in the scenario will still occur. These later opportunities may depend on actions that the TAOs must take earlier in their scenarios. This requires the TAOs take certain actions, even if they don't know when or how to do them.

This hinting strategy was accomplished in a simple way by having 4 levels of successive hinting, when the Instructional Planner has turned hinting on. The first level of hint is just a flag that merely indicates that some TAO action is expected. This is denoted by a blue bar across the bottom of the screen if the TAO has not responded to the situation immediately. If the student continues to fail to respond successfully he or she receives a general hint that merely references the general principle that relates to the current situation. If the student continues to fail to take the correct action, he or she receives a more specific hint which maps the general principle onto the current situation. If the student continues to fail to make the correct decision, he or she receives a prompt saying exactly what should be done. If the student still continues to fail to do take the correct action, he or she receives feedback, the simulation pauses awaiting the correct action, and the student is again instructed what to do. The time intervals between the various levels of hinting depend on the principle involved and sometimes on the pace of the simulation. For example, the principle that the TAO must acknowledge all communications addressed to the TAO has very short time intervals (4 seconds) since a reply should be immediate. The interval between hints relating to querying the identity of an inbound track depends on how fast that track is proceeding inbound. A correct decision with a hint is scored as 0.8 with just the flag, 0.6 with the general hint, 0.4 with the specific hint and 0.2 with the prompt.

Instructional Planning

In addition to "mastered", there are other terms to indicate lesser levels of mastery at any level. These are "begun" (as in "begun novice level") for 0 to 0.4, "partly mastered" for 0.4 to 0.69, and "almost mastered" for 0.7 to 0.84. Of course these and other specific numbers can be easily changed. When a student has a principle in the lowest category (i.e. "Begun") of the current level, then hinting is provided for this principle. The mastery estimate is updated during the scenario and the hinting turned on/off on a principle by principle basis as the estimate moves down/up, respectively. Note that hinting will often be turned on for some principles and turned off for others at the same time. The student will only receive hints on his weak areas.

By defining the mastery formula as the average of the previous estimate and the most recent score (a very simple formula) and by defining the categories as above, with hinting for the lowest category, the desired instructional strategy is achieved in a simple way. A student new to the area who takes the correct action immediately or with only the flag hint will have hinting turned off immediately. The student will have to demonstrate correct performance a total of three times in a row to have been determined to have mastered the principle at the current level. Students responding slower or failing to respond correctly at least approximately 40% of the time (weighting most recent responses heaviest) will continue to receive hinting. They will not be considered mastered until they have decided correctly approximately 85% of the time (weighting most recent responses heaviest). This requires at least two correct decisions applying this principle in a row after hinting has been turned off.

For principles where the mastery estimate is at the "Begun" level, the students will receive positive feedback for correct decisions, under the assumption that if they are either new to the area or have problems with it that they may not be sure when they have done something correctly. The students always receive negative feedback on failed attempts, unless the instructors have determined that delaying feedback until the debriefing is more appropriate for a particular principle. Delayed feedback might be best for principles involving high-level decisions such as the decision to engage a particular track, for example, but not for low level principles, such as proper communication protocols. Important high-level decisions are more likely to be remembered until the end of the scenario, when the debriefing occurs, and also there is often learning value in seeing the ramifications of high-level decision mistakes in the scenario. Low-level decisions are more likely to be forgotten and thus waiting until the debriefing may largely destroy the value of the feedback.

The debriefing includes information on every attempt to apply a tactical principle. For each time in the scenario that the student needed to apply a principle the debriefing includes a description of the situation, significant events leading up to the situation such as communications from team members or important track events, any hints received, the student's actions, a description of whether those actions were correct or not and links to the relevant principles. The situation also includes a snapshot of the TAO's console at the time the principle applied.

The Instructional Planner determines that the student has graduated to the next level when he or she has achieved the highest category (i.e. "Mastered", 85%) for the current level for all tested principles. For students who have not mastered all principles at the current level (e.g. Novice or Intermediate levels) mode, the Instructional Planner recommends for the specific student, scenarios at his/her current level that requires the greatest percentage of his or her weakest, tested principles. The Instructional Planner continues to select these types of scenarios for the student until all principles are mastered. The number of practice scenarios needed by the student to achieve mastery depends on how well they are performed.

ARP DESCRIPTION

Each ARP takes as input the simulated tactical situation, the electronic counterpart to verbal utterances from the other ARPs, and the student-edited translation of his verbal commands and queries and produces actions in the simulation, utterances to the TAO, and utterances and their electronic counterparts to other ARPs. ARPs inject the errors of omission or errors of commission, as directed by the Instructional Planner.

Some behaviors are common to all ARPs. All ARPs share a common acknowledgement behavior, in which they respond with the communication "<own designation>, aye" when addressed by another watchstander, except when a precise response is warranted. This is in addition to any nonverbal response. When an immediate communication is also warranted, the ARP may conjoin the acknowledgement: "AIR, aye, <communication>." For errors of omission, all ARP behaviors can be deactivated for a specified period, or modified to a specified fraction of execution, by the ITS. For errors of commission, most ARP behaviors (or BTNs) have one or more error variants which can be activated by the ITS. (For example, altering the track number in a verbal report to another existing track number.)

The following subsection address the different components and challenges of the ARPs

Speech Recognition

The student's spoken utterances are critical inputs both for performance evaluation and to direct the ARPs. The challenge to achieve essentially perfect performance was addressed by two techniques. The first was based on the fact that this system was not being designed to tutor radio communications syntax and skills. This would be done in an earlier part of the course. Thus communication vocabulary or syntax mistakes would not be caused by a lack of understanding but were more likely to be the results of momentary lapses. Thus under the instructors' direction, the speech recognition system uses only a correct syntax and vocabulary augmented with likely incorrect replacement words (synonyms such as "kill", "engage", "destroy", etc.). The system will try to force any utterance into this defined grammar. The benefit is that syntactically correct utterances by students will be correctly recognized a very high percentage of the time. The instructors felt that this was the only important criterion. When the students say something syntactically incorrect it will either not be recognized at all or recognized as something different from what they said. But since this was just a momentary lapse, either of these results will in effect let them know that they have said something wrongly so they can restate their order or query.

But even a very high rate of correct recognition may be unacceptable given how the resulting text will be used. The solution was to display the recognized text in a simple utterance editor which allows the TAO to rapidly review and possibly edit (or enter) with pull-down menus the text translation of his verbal utterances before sending. This editor includes a submit button that must be pushed before the recognized (and possibly edited) text is sent. This allows students to confirm that their utterance was correctly interpreted before submitting and to either edit it or restate it. Figure 4 shows the Utterance Editor Design and Figure 5 shows one of the pulldown menu sequences. The actual words used as commands in the TAO's utterances have been replaced with the word "Command" since the actual words are sensitive.

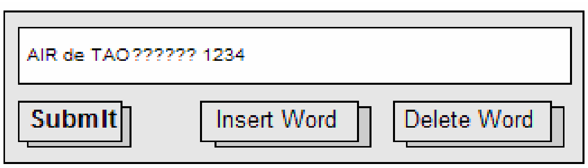

Figure 4. Utterance Editor

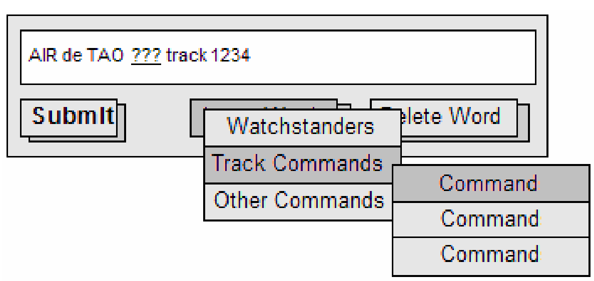

Figure 5. Utterance Editor Pull-Down Menus

Behavior Logic

Behavior Transition Networks (BTNs) were described in the Automatic Evaluation Subsection. BTNs, in addition to being useful for evaluating tactical decision-making in free-play scenarios, are also useful for automating tactical decision making. Similar to evaluation, the first part of the BTN looks for a specific type of situation. However instead of the second part monitoring the student's action, it performs the correct actions itself. For example, the Anti-Air Warfare Coordinator (also called "AAWC" or "AIR") ARP, has a behavior, copied for each incoming air track, which examines the specific air track and, if it meets certain criteria, determines that a query should be issued. It then issues an intention to query utterance, waits for an acknowledgement, and, if received, orders the Identification Supervisor (IDS) ARP to query the track.

There is no significant difference between whether the actions that the ARP performs are verbal, actions taken in the simulation, or a combination of both. Both verbal actions and simulation actions are represented as action primitives in the BTN and ultimately make external calls to interface to the appropriate system (speech generator or simulation interface).

Maintaining ARP responsiveness in a wide range of situations was an important challenge which was tackled in three different ways. First the ARP have BTNs to perform the correct actions in various tactical situations that can occur in the defined set of scenarios, such as the query behavior mentioned earlier. Second, each ARP has a contextual memory of the past activity of the tracks, TAO and other ARPs. This supports the third aspect' that of BTNs that respond to the TAO's utterances. For example if the TAO requests the status of a query on a particular track, the ARP will access its contextual memory and respond with the results of the query if a query was performed, report that the query is in progress if it is, or if a query has not been initiated yet, inform the TAO of that fact.

Injected Errors

As described previously, scenarios above the novice level involve ARPs making mistakes so that students can demonstrate an increased understanding of the tactical principles. Associated with each ARP behavior is a value that specifies whether that ARP should execute properly or omit its action. If omission is called for, then instead of taking the correction actions, a message is sent to Automatic Evaluation that the omission error was just made. An evaluation BTN is then triggered by this message to begin monitoring the student's response (which should correct the error by requesting that the omitted actions be performed). If turned on by the Instructional Planner, appropriate hinting is also activated.

Most of the ARP behaviors also have variants that commit specific types of errors in the form of incorrect actions. An example would be a variant of the query behavior that announces the intent to query the wrong track number. Similar to errors of omission, these variants become active based on a value in the ARP that specifies whether that ARP should commit this specific type of error. When the ARP commits the error, it sends a similar message to Automatic Evaluation that the error was just made so that a BTN can begin monitoring the student's response (which should correct the error by negating the intention and issuing a correcting order).

The ARP determines that at the next opportunity it should commit either an error of omission or an error of commission from the fraction of that type of errors listed in the exercise by the instructor. A simple, deterministic algorithm, rather that one based on a random number generator, was selected so that the errors would be predictable. For each behavior, the ARP keeps track of the number of relevant errors and correct executions. It determines that at the next opportunity it should perform correctly or make an error based on whichever choice will lead to a cumulative fraction that is closest to that specified. For example, if the exercise specifies that the AIR ARP should fail to take the query actions 66% of the time, then it will fail at the first opportunity because 1/1 is closer to 66% than 0/1. At the next opportunity, it will perform correctly because ½ is closer to 66% than 2/2.

Speech Generation

There are several commercially available speech generation systems which will convert text to spoken language. Many offer a variety of voices which are useful for differentiating between the different ARPs. A mixture of male and female voices improves differentiation. Assembling the correct text is relatively straight-forward since the specific words team members are supposed to say are fairly static with the exception of specific parameters such track number, range, speed, etc. The challenge is to prevent the individual ARPs from speaking over each other and the student and do it in a way that doesn't complicate the individual BTNs. The solution was aided by the fact that the headset, in order to mimic the operational equipment, had a push to talk foot pedal which could be monitored by the software.

All the ARP's BTNs use the same action primitive to speak called "Say" that does not directly send the text to the speech generator. Instead it put the text on a queue of things to be said. A simple behavior checked the push-to-talk (PTT) signal before sending text to the speech generator. If PTT was on, it simply waited before sending the next text until PPT was off. Similarly, when text from an ARP was being spoken over the communication network, a flag was set. To keep the student from trying to talk over the ARP who was in the middle of speaking, when the PTT was pushed while the flag was set, the student's microphone was disabled and a message appears telling the student to wait briefly. Since the ARP utterances are short this is usually only a few seconds. When that utterance completes, the behavior sees that the PTT is pushed and waits before sending the next text to the speech generator and resets the flag. This re-enables the student's microphone so that he or she can speak.

EXERCISE AUTHORING

The goal of making exercise authoring as simple as possible was accomplished by conceiving of it as simple entry of field values in a series of screens. The most important of these are the GRTS scenario file (which defines the initial tactical situation), exercise level (Novice, Intermediate, Advanced), evaluated principles (which specify which evaluation BTNs should be running), ARP mistakes, Warning and Weapon Status, and various parameters such as query and warning distances and Surveillance Area and Vital Area radii.

STUDENT MONITORING

Simultaneously monitoring the progress of 42 students each running their own scenario creates a unique challenge. Incorporated is a "Stoplight Display" to allow the instructor to rapidly track (at a macro level) student progress in their scenarios. The "Stoplight Display" is highly intuitive in design and mimics similar tools used by the fleet to monitor training and material readiness. Figure 6 shows this display. With an instructor present, most or all the students will be in "Classroom Mode" running the same scenario as directed by the instructor. Homework mode refers to practicing scenarios, independently and would not normally be done while an instructor was there monitoring students.

Figure 6. Students Performance Summary Display

FUTURE WORK

GRTS TAO ITS is currently being tested and validated for operational use. The first step in the process was a demonstration of a preliminary system to convey the look and feel of the overall system to the instructors.

Next, the combined system is exercised on a large number and variety of test cases designed to exercise each component fully. This is most complex for the ARPs and Evaluations. Both a systematic and statistical testing approach are being employed. The systematic approach takes advantage of the fact that the separate principle evaluations and different ARP behaviors are performed by separate Behavior Transition Networks (BTNs). Correct or incorrect performance along one principle should not impact the evaluation of a different principle or the behavior of different ARPs. The systematic approach makes sure every principle is completely tested by making sure each meaningfully different timing of correct student performance is tested for every principle as well as testing omitted and incorrect responses. This also tests the ARPs on their behavior to all meaningfully different student actions in the full range of scenario situations that occur in the training set. Three different scenarios are used – one with every novice principle represented, one with every intermediate principle, and one with every expert principle.

The statistical approach uses random choice to determine whether a correct, incorrect, or omitted answer will be given and makes sure each principle is tested with this methodology at least 10 times. This will give a reasonable statistical sample for each principle as well as over 600 samples for the entire set of principles, allowing a reasonably powerful method for statistically calculating an error rate. The random choice will be performed at the time the test case is defined, so that the appropriate test inputs and expected program outputs can be documented in advance, thus following standard testing procedures. The statistical testing will use the same three scenarios as the systematic testing.

This testing and accompanying debugging will provide a robust system that the instructors can run and evaluate. Their feedback from this “leave behind” version will be incorporated into the final system which will be used and evaluated with a class of real TAO students at SWOS. Suggestions from this fist class will be incorporated into the operational GRTS TAO ITS. This first operational version covers one area, Air Defense Detect to Engage. Development will also continue on 9 other domain areas. Additionally we will apply the underlying ITS technology, FlexiTrainer, to other application areas. Similarly we will apply the ARP underlying technology, SimBionic, to developing ARPs in other domains.

CONCLUSIONS

BTNs proved to be a natural and effective way to implement the evaluation and ARP behaviors. The Speech Recognition solution of using a strict grammar and an utterance editor works well when the students’ utterances are based on a restricted syntax and the ITS is not tutoring communication skills per se.

REFERENCES

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

Stottler, R., B. Spaulding, R. Richards (2005), *Use Cases, Requirements and a Prototype Standard for an ITS/Simulation Interoperability Standard (I/SIS),* SISO 2005 Spring Simulation Interoperability Workshop San Diego, CA, April, 2005.

Stottler, R. (2003), *Techniques for Automatic AAR for Tactical Simulation Training*, I/ITSEC 2003. Dec., 2003.

Stottler, R. & Vinkavich, M. (2000), *Tactical Action Officer Intelligent Tutoring System*, I/ITSEC 2000.