Autonomy Requirements for Virtual JFO Training

Brian S. Stensrud, Ph.D.
Soar Technology, Inc.
1942 West CR 419, Suite #1060
Oviedo, FL 32766
stensrud@soartech.com

Gino Fragomeni, Pat Garrity
U.S. Army Research Laboratory
SFC Paul Ray Smith Simulation &
Training Technology Center
gino.fragomeni@us.army.mil

ABSTRACT

Over the past decade, the DoD has invested significant resources into developing virtual practice environments for training joint terminal attack controllers (JTACs), forward observers (FOs), and joint fires operators (JFOs). These practice environments feature realistic visuals and functional true-to-life equipment (such as binoculars, mapping tools, and radios). However, executing realistic training scenarios using this technology requires a variety of simulation operators and instructor personnel to provide role players and training support for each student receiving the training. By exploiting recent advances in artificial intelligence technology, however, these environments can be enhanced to provide a force multiplier for these personnel, resulting in a more realistic, individualized experience for the warfighter. In this paper, we describe specific components and functions of virtual JFO/JTAC training that can be enhanced by artificial intelligence technology, as well as the specific algorithms and components that can be brought to bear. Specifically, we describe methodologies for controlling air and ground support assets, such as CAS aircraft and Fire Direction Centers (FDCs), and address the use of natural language processing and generation technologies to connect those assets verbally to the student. Finally, we describe an existing effort sponsored by STTC where these technologies are being integrated and demonstrated into an existing virtual JFO training system.

ABOUT THE AUTHORS

Dr. Brian Stensrud is a lead scientist at Soar Technology, Inc. specializing in artificial intelligence technology for training simulations and related applications. On staff since 2003, Brian has been principal investigator/project manager for over $6M in research contracts, as well as serving in a technical lead role in the development of several autonomous agent platforms and toolsets. He has been selected for two Army SBIR Achievement Awards, in 2010 and 2012, for work performed by the U.S. Army Aviation, Research, Development and Engineering Center (AMRDEC). Brian received a Ph.D. in Computer Engineering from the University of Central Florida (2005), and
B.S. degrees in Mathematics, Computer and Electrical Engineering from the University of Florida (2001). He has over 12 years experience in artificial intelligence, behavior modeling, and simulation.

Mr. Gino Fragomeni serves as a Science & Technology Manager for Dismounted Soldier Technologies at U.S. Army Research Laboratory, Simulation & Training Technology Center (ARL-STTC). He currently works in Ground Simulation Environments Division conducting R&D in the area of dismounted Soldier training & simulation. His current interests include artificial intelligence and immersive environments centric to dismounted training applications. Gino is also a reservist with the United States Army Special Operations Command-Central (SOCCENT) with over 28 years of military experience. He earned a Master of Science from the University of Central Florida (2002) and has specialized training in Systems Engineering and Simulation.

Mr. Pat Garrity is the Chief Engineer for Dismounted Soldier Training Technologies at the Army Research Laboratory's Simulation and Training Technology Center (ARL STTC). He currently works in the Ground Simulation Environments Branch conducting research and development in the area of dismounted Soldier training and simulation where he was the Army's Science and Technology Objective Manager for the Embedded Training for Dismounted Soldiers Science and Technology Objective. His current interests include Human-In-The-Loop (HITL) networked simulators, virtual and augmented reality, and immersive dismounted training applications. Garrity earned his B.S. in Computer Engineering from the University of South Florida in 1985 and his M.S. in Simulation Systems from the University of Central Florida in 1994.

Autonomy Requirements for Virtual JFO Training

Brian S. Stensrud, Ph.D.
Soar Technology, Inc.
1942 West CR 419, Suite #1060
Oviedo, FL 32766
stensrud@soartech.com

Gino Fragomeni, Pat Garrity
U.S. Army Research Laboratory
SFC Paul Ray Smith Simulation &
Training Technology Center
gino.fragomeni@us.army.mil

BACKGROUND

Joint Fires Observers, or JFOs, are US Military personnel that provide fires coordination support, specifically for indirect fire and close air support. In this capacity, JFOs provide a critical link between dismounted personnel on the battlefield and supporting fires elements. In the US Army, at least one JFO element is attached to every maneuver platoon for this purpose. Typically, JFOs take up a forward position in an operating post that optimally provides good visibility on both targets of interest and potential strike ingress vectors. JFOs respond to and interact with units on the ground, first to identify the nature and location of an enemy force. The JFO is then responsible for selecting a course of action and coordinating with fires elements to get rounds on that force. This coordination, called 'terminal guidance operations,' involves interaction with the fires element (through voice or other signals including smoke rounds and laser targeting) to identify and pinpoint the enemy position. Proper coordination by the JFO is critical to the success of these fires missions and the safety of the dismounted force.

Over the past decade, the DoD has invested in a variety of virtual simulation technologies across multiple services to support the training of the JFO and related roles (e.g., the forward observer and joint terminal attack controller/JTAC). These include PEO STRI's Call for Fire Trainer (CFFT), Joint Fires and Effects Training System (JFETS), and CFFT-AV (augmented virtuality); the Air Force's JTC/TRS system (see Figure 1), and others. These systems share several features, including a virtual display of the battlespace (either on a flat/domed screen or viewed through a head-mounted display), realistic equipment (e.g., binoculars, targeting equipment, radio), and a SAF backend for representing the scenario's computer generated forces.

Figure 1. The U.S. Air Force's JTC TRS Trainer

With few exceptions, these training environments require the time and expertise of human role players to drive non-player activities. For JFO/JTAC training, these roles include fixed- and rotary-wing aircraft piloting, indirect fire coordination, squad-level behaviors, and OPFOR. When these role players are unavailable to support the training, then, computer-generated forces (CGFS) must be relied on to simulate each of these roles.

CGF REQUIREMENTS IN DYNAMIC TRAINING ENVIRONMENTS

Historically, constructive and virtual training systems make use of lightweight CGF control mechanisms, commonly known as *scripts*, for situations where human role player support is infeasible. Scripts can be thought of as a set of step-by-step instructions that are carried out by CGFs over the course of a scenario. As such, scripts generate predictable, repeatable CGF behaviors and thus a consistent set of stimuli for the trainee. Depending on the application, CGF scripts may provide an appropriate experience.

However, in many military domains (including JFO training), a scripting approach is far too static and brittle to support training goals. Specifically, scripts are not dynamic or complex enough to handle the non-determinism that is common in training simulations – e.g., from unexpected trainee behaviors, environmental conditions and other stimuli.

There are a few requirements for CGFs to be capable of handling complex training environments such as JFO training. The first, and perhaps most important, of these characteristics is the ability to maintain constant attention of the environment. Robust CGF systems must have constant and frequent attention to their environments regardless of task. For scripts and other legacy CGFs, it is the lack of this constant perception that precludes them from handling the unexpected in a training environment.

Another requirement for robust CGFs is that they *expect* interruption to their behaviors and are modeled accordingly. In contrast to behavior scripts and other simple CGF platforms, which are encoded to follow a rigid sequence of tasks or goals, more capable CGFs must be capable of reacting to unplanned stimuli in the environment without permanently disrupting their primary goals. As such, these systems must respond to interruption in such a way that they can (1) suspend their current behavior to address to the stimuli and (2) properly *re-enter* (not restart) their interrupted behavior after response.

Another important characteristic is the ability of agents to handle the existence of multiple, simultaneous goals within a single exercise or training scenario, specifically if they are intended to replicate a specific warfighter role. Agents must then have the ability to prioritize amongst their various goals and constantly decide which goals to act upon based on the situation and mission needs. By combining the pursuit of multiple goals with constant situation awareness and the ability to interrupt tasking, intelligent agents are able to both act deliberately according to a pre-specified plan and also react to events within a dynamic environment in support of those goals.

As an alternative to scripts or other legacy CGF paradigms, artificially intelligent human behavior models – or *intelligent agents* – can provide a variety of features not present in typical CGF platforms (Stensrud et al, 2012):

- *Goal-based behavior*. Not dependent on a pre-specified script, intelligent agents use real-time situation awareness and can act in pursuit of domain-specific goals.
- *Non-determinism*. Intelligent agents are responsive and sensitive to unpredictable changes and unanticipated events in the environment.
- *Dynamic decision-making*. Intelligent agents are able to pursue, prioritize, and reason about multiple goals or tasks simultaneously, automatically switching goals when necessary.
- *Interactivity*. Intelligent agents are able to communicate information, status, and commands, and interact with both human operators and other agents in the context of the exercise.
- *Transparency*. Intelligent agents can explain complex reasoning behaviors and their motivations for validation and after-action review purposes.

An effective means by which to implement intelligent agents with these features is to use a *cognitive architecture* (see Laird, 2012, Anderson, 1996 Cassimaitis, 2002), though other methodologies and AI techniques may be appropriate given the breadth and depth of the required behavior. Cognitive architectures, specifically Soar and ACT-R, are particularly suited to model these behaviors, primarily because they are designed to model behavior in manners that imitate the way the human brain operates - *unified theories of cognition* as Newell (1990) posited. Recall that the goal is to model behaviors associated with JFO training – aircraft, fires operators, and dismounted units – all controlled in the operational environment by well-trained (and human) soldiers. Any realistic training experience that automates those roles must be capable of modeling their complexity.

CGF REQUIREMENTS FOR JFO TRAINING

In this section, we introduce and describe specific automation requirements for a subset of the roles required for JFO training – aircraft for close-air support, indirect fire coordination, and the JFO role itself. As described above, the JFO's responsibility is to coordinate fires on ground targets by communicating with both the dismounted element and the fires element. For instance, consider a dismounted infantry squadron taking fire from heavily-armed insurgents in an urban area. Depending on the size of the insurgent force and the situation, the squad may not have enough firepower to eliminate the threat, and thus must call in for air or indirect fire support. To do this, the squad will make a radio call into the JFO, detailing their situation and need.

Optimally, the JFO will be situated in an observation post (OP) or base, away from the fight but with good visibility of the ground units, the threat, and the surrounding area. Communicating with the units on the ground, the JFO will first obtain awareness about the situation; specifically the size and nature of the threat, the goals and status of the ground units in contact, and details about the immediate surrounding area.

Figure 2. Boeing AH-64 Apache Helicopter

After initial interaction with the ground unit, the JFO will make a judgment on how to best implement fires. Depending on the situation, the JFO may have a variety of fires options available; including artillery units and close-air support from fixed- or rotary-wing air assets. Choosing the appropriate measure requires consideration of many factors beyond simply the nature of the threat – factors such as ground unit position and the surrounding area (in urban environments, for instance, heavy indirect fire can easily impact the population).

Once a measure is chosen, the JFO must interact with the fires element (either directly or via JTAC) to submit the request details for fires. For artillery strikes, the JFO will make a call into the Fire Direction Center (FDC), a division-level (and higher) element that controls a battery of guns. In that call, the JFO will provide position and target description to the FDC, as well as other parameters and restrictions (e.g. "danger close" modifiers to indicate that the target is nearby friendly/neutral elements). The JFO is also responsible for providing targeting support for the guns, which includes 'spotting' the rounds as they impact and providing adjustment data to the FDC if and when they are off-target (U.S. Army, 1991).

The JFO similarly coordinates indirect fires support from air elements, including rotary wing (e.g., AH-64) and fixed-wing (e.g., AC-130) aircraft. This coordination requires similar data be passed from the JFO; however, tighter coordination is required as the fires elements are in motion and require a direct line to the target. This means that the JFO must not only coordinate the selection and direction of the fires, but also the aircraft's approach towards, and recognition of, the target.

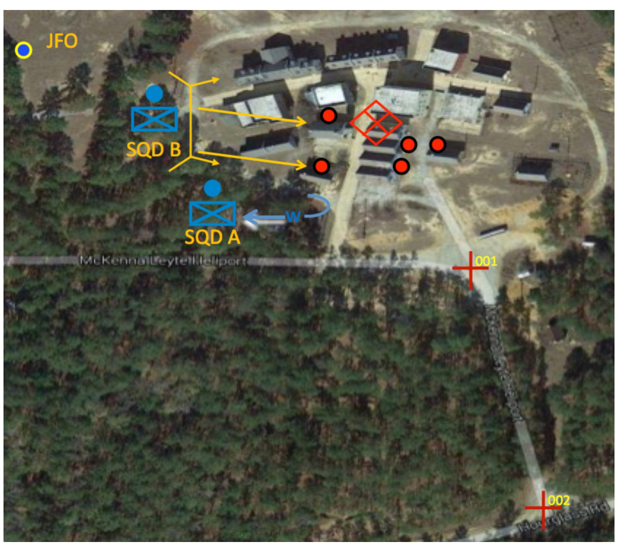

Figure 3. Top-down view of a representative fires mission

Finally, the JFO must coordinate back with the units on the ground to inform them of the incoming fires and their effect. Information of where the fires are coming from and where they are targeting is important for the ground unit in anticipating how the threat may react to the fires (e.g., escape direction). Additionally, this coordination may include advisories on dangers such as blast radii – close-combat situations where the ground units are going to potentially be affected by the incoming fires and must move to safer locations.

As described, the JFO has an important, non-trivial role to play to support ground units with fires. However, the roles of the other elements involved in the process are similarly non-trivial. Automating these roles for a virtual training use case, therefore, requires the encoding of dynamic, interactive behaviors that is not possible using scripts or other static approaches. In the following sub-sections, we enumerate the primary requirements for implementing autonomous characters to serve as dismounted units, air assets, and fire support elements in a JFO training scenario.

Situation Awareness

It is necessary for each of the automated support roles to maintain situation awareness throughout each training scenario. Endsley (1994) defines situation awareness as the perception of the elements in the environment within a volume of time and space, the compression of their meaning, and the projection of their status in the near future. In a JFO context, situation awareness covers the following:

- Awareness of the position, intent, and state of the various friendly, neutral, and enemy forces in the scenario, to the extent that data is available
- Entity-level understanding of the mission and the goals necessary to complete that mission
- Physical and human terrain, and the effects of that terrain on the mission and goals
- Rules of engagement, if applicable

Maintaining accurate situation awareness is key for automated role players to provide realistic behaviors in virtual environments. The actions of the human trainee, by nature, are unpredictable and have a direct effect on how training scenarios unfold. Role players must have the ability to adapt to this non-determinism in a way that is consistent with their own mission and goals.

Speech Recognition and Generation

Most of the interactions that take place between the JFO, the ground units, and the fire support elements take place over voice. On the surface, this means that each of the supporting autonomous elements must be able to recognize and generate speech to simulate these interactions. However, this requirement is often misunderstood as simply the ability of CGFs to *recognize* speech (e.g., converting spoken words into a string of text), leaving out a critical element of *understanding* that speech in terms of those CGF's goals.

The prevalence and variance of speech interaction in complex domains such as JFO training requires that the supporting elements can both *recognize* incoming voice interactions and also *understand* the content of those interactions in proper context (so that they can appropriately respond). To do this, autonomous elements must include the following subsystems:

- A speech recognition engine capable of converting an audio feed into computerized text. Several COTS recognition engines exist that can support this task.
- A grammar-based or natural-language interpretation engine, capable of reading computerized text and converting into parsed machine-readable text (e.g., in XML or JSON format)
- An intelligent agent or reasoning system able to process machine-readable text as interactions and parse their contents in the context of its goals

Put simply, the 'speech recognition' requirement can be broken down into three sub-problems: speech recognition, speech processing, and speech understanding. Similarly, speech generation is not simply the problem of taking text and turning it into audio. Rather, speech generation in a complex training environment requires the speech content to similarly be generative – elements must be able to compose text statements in real-time based on the situation. In contrast to speech recognition, though, outgoing text statements can go directly to a speech generation engine (again, several COTS engines exist for this step) without the need for a parsing/processing step beforehand.

Tactical Behaviors

As described, both the supporting ground units and the fires elements have dynamic roles to play in JFO training scenarios. In addition to the more generic requirements of situation awareness and speech generation/recognition, these elements must also be able to properly execute their roles to support the mission.

Squad-level and other ground units support the training scenario by providing robust and accurate behaviors in support of direct engagement with the enemy force. Specifically, these units must be able to move tactically around the environment, perform direct engagements with the enemy force using proper cover and concealment techniques, and react appropriately in response to interactions with the JFO (e.g., retreating to a covered area to avoid fratricide from incoming fires). Likewise, if realistic OPFOR are to be automated, they must also include representative movement, firing, and coordination techniques.

Similarly, air and artillery elements support the training scenario by executing fires requests delivered by the JFO. Air assets must properly approach the target based on knowledge of the terrain, known target reference points, and the nature of the enemy threat. In some cases, when the target is not easily visible from the aircraft's vantage point, tighter coordination between the pilot and JFO is required (e.g., 'talk-on-target', U.S. Navy, 2009). Finally, each element must select an appropriate method of fire, in coordination with the JFO based on the threat, the location and potential danger to the ground unit, and known rules of engagement.

Many of the behaviors listed above can be simulated by a script or low-fidelity SAF in a static training scenario. However, in a JFO scenario, each of these behaviors relies on real-time information and the capacity to react to a changing situation on the ground. Scenarios in this domain are typically quite fluid; the parameters of an attack plan can quickly change due to any one of a number of occurrences, requiring fires elements to react by adjusting their behaviors. Automating that behavior, therefore, requires a modeling representation that can encode not only the mechanics of these tactics but also *an understanding* of those tactics and how they support the mission. It is this understanding that will allow automated elements to behave in uncertain and constantly changing environments.

USE CASE: AUTONOMY FOR CFFT-AV

The Call for Fire Trainer – Augmented Virtuality (CFFT-AV, Korris and Garrity, 2011) is a virtual training system for JFOs and JTACs and is currently installed at the SFC Jared Monti Hall Mission Simulation Center at the Fires Center of Excellence in Fort Sill, OK. CFFT-AV provides an environment where JFO trainees use physical equipment replicas (including rifles, binoculars, compasses, rangefinders, paper maps, etc.) to conduct fires missions in conjunction with simulated entities in a virtual world.

Figure 4. The CFFT-AV

The CFFT-AV currently relies on manned role players to manage the supporting CGFs (aircraft FDC elements) during training scenarios. The SFC Paul Ray Smith Simulation & Training Technology Center (STTC) is currently sponsoring an effort to develop and integrate autonomy for these roles. A primary goal of this effort is to expand the applicability of the CFFT-AV, allowing the JFO to get training using the system when manned support roles are unavailable.

Intelligent Agents for CFFT-AV Integration

To support this effort, the authors have brought to bear existing CAS behavior models, developed using the Soar cognitive architecture, to control CGF for JFO scenarios (Jones et al, 2009).

- Helo-Soar, an intelligent agent for rotary-wing aircraft behaviors and CAS
- IF-Soar, an intelligent agent for fire direction center
- Insurgent agents for enemy combatant behaviors (Stensrud et al, 2012)

Each of these agents has been adapted from previous use cases and system architectures. Since their behaviors execute in parallel with the standard game loop of a virtual or constructive simulation, a feature common to cognitive architectures, they have the advantage of being portable and not tied to a particular software configuration.

Helo-Soar (Jones et al, 2007) and IF-Soar (Stensrud et al, 2006) both support a variety of behaviors supporting fires missions, and interact directly with the JFO using doctrinal speech. Additionally, both agents consist of goal-level understandings of how to execute CAS activity. As the parameters of the mission unfold in real time, each agent can react and adapt in kind, maintaining situation awareness throughout each training scenario and using that awareness to inform its goals.

System Architecture

CFFT-AV currently uses the Gamebryo engine for its virtual environment display, however STTC is in the process of migrating from Gamebryo to Unity. In anticipation of that switch, the authors' current prototype architecture connects IF-Soar, Helo-Soar and insurgent agent models into the Unity game engine, as illustrated in Figure 5.

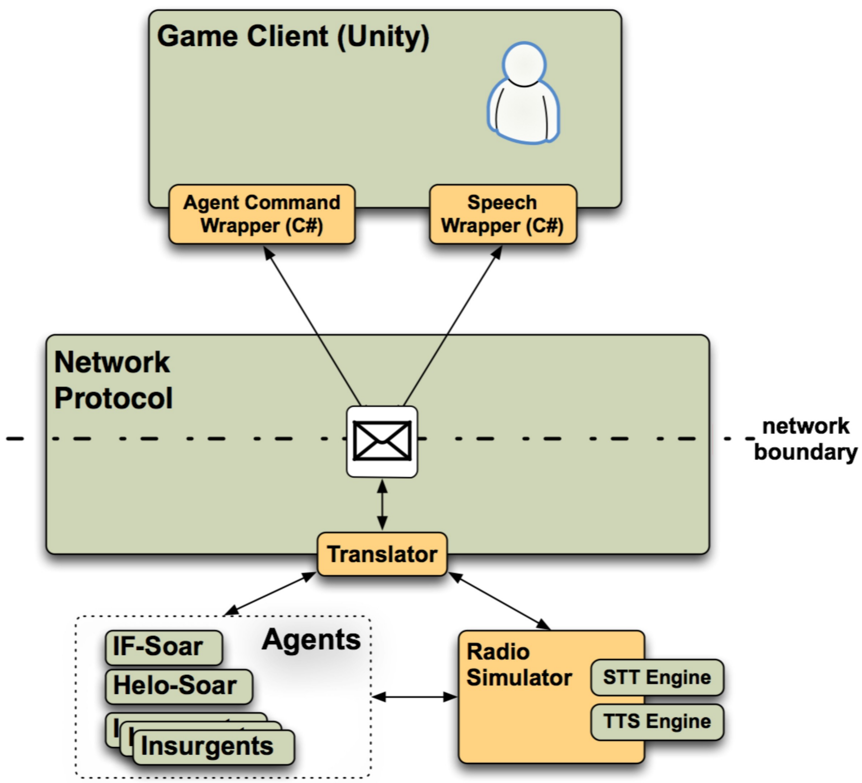

Figure 5. Prototype Architecture.

As shown, each of these agents are connected to the Unity game engine (www.unity3d.com) via a network boundary and messaging protocol – JBUS (Dumanoir, 2008), in this case, and connected to a radio simulator software package which includes COTS speech recognition and generation services. Using this architecture, the agents receive state information from the game environment via the JBUS interface, and send commands back through that interface that correspond to the actions that they wish to execute.

Voice interactions similarly take place via network protocol – speech actions are recognized and parsed by a radio simulator component, using an embedded speech-to-text (STT) engine and supporting software, with machine-readable XML representations of that speech sent to the agents for consumption (e.g., a message from the JFO to an aircraft requesting a fires action). Speech actions generated by the agents are sent back to the radio simulator, which uses a text-to-speech (TTS) engine to produce audio. That audio is then compressed and sent over the network protocol and played to the trainee through his headset/speakers.

SUMMARY AND CONCLUSIONS

JFO operations, as with many military domains, is a complex and dynamic task; requiring coordination and synchronization with a variety of ground, air, and artillery elements. Absent manned role players to support JFO training in a virtual or constructive simulation environment, unmanned or autonomous entities are a hard requirement. In this paper, we outline the features necessary of autonomous entities to fill this requirement, and describe an ongoing effort to integrate and apply such entities to an existing trainer.

Using the architecture described above, we have successfully generated a prototype training capability for JFO elements featuring fully autonomous supporting fires elements. With this capability, JFO trainees can take incoming calls from manned or unmanned ground elements, assess the situation on the ground and then coordinate fires support with either rotary-wing or artillery elements, depending on which is more appropriate. Those fires elements will then carry out the fires request within Unity so that it is visible to the trainee.

Since the agents are interacting with the JFO trainee only through the game environment and speech engine, they are agnostic to the specific technologies that the JFO trainee is using. As a result, the agents and agent architecture developed for this prototype will be able to fold directly into the CFFT-AV once Unity support has been added. Once integrated, JFO trainees will be able to take advantage of the full CFFT-AV system without the added requirement of manned support roles.

Figure 6. Playing of the role of the JFO, calling in a fires mission to be carried out by autonomous support elements in the Unity game engine

Note that, by integrating autonomous entities for each necessary role, we can adapt a training system to support training of other roles. For instance, by swapping in an autonomous JFO agent and swapping out the autonomous ground units, we create the underpinnings of a close-air-support trainer for infantry. With such a trainer, dismounted infantry trainees in a virtual environment (such as the Dismounted Soldier Training System/DSTS) can call autonomous entities for fires support, and perform all of the necessary activities to support that call (e.g., proper communication verbiage, providing directing/support fire on the enemy, moving away from the target area to avoid fratricide, etc.)

ACKNOWLEDGEMENTS

This work described in this paper was supported by the U.S. Army under Grant W911QX-12-C-0143.

REFERENCES

Jones, R. M., Wallace, A. J., & Wessling, J. (2004). An Intelligent Synthetic Wingman for Army Rotary Wing Aircraft. Proceedings of the *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) Conference*, Orlando, FL.

Jones, R.M., Stensrud, B.S. and Taylor, G. (2009), "Knowledge-Rich Intelligent Agents for Special Operations Forces (SOF) Training," Proceedings of the *2009 Special Operations Forces Industry Conference (SOFIC)*, Tampa, FL. June 2-4, 2009.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51, 355–365.

N. L. Cassimatis, "A cognitive architecture for integrating multiple representation and inference schemes," *inMedia Laboratory*. Cambridge,MA: MIT, 2002.

Stensrud, B., Taylor, G., & Crossman, J. (2006). IF- Soar: A Virtual, Speech-Enabled Agent for Indirect Fire Training. Paper presented at the *25th Army Science Conference*, Orlando, FL.

Laird, J.E.: *The Soar Cognitive Architecture*, 2012, MIT Press.

Newell, A. 1990. *Unified theories of cognition.* Cambridge, Mass.: Harvard University Press.

M. R. Endsley and M. D. Rodgers, "Situation awareness information requirements for en route air traffic control," presented at *Human Factors and Ergonomics Society 38th Annual Meeting*, Santa Monica, CA, 1994.

United States Army (1991), *Army Field Manual 6-30*, armypubs.army.mil/doctrine/DR_pubs/dr_a/pdf/fm6_30.pdf

United States Navy (2009), Close Air Support, Joint Publication 3-09.3, http://www.fas.org/irp/doddir/dod/jp3_09_3.pdf

Dumanoir, P., Clinger, B. and Rivera, J. (2008), Evolving Standards in US Army Live Simulations, Proceedings of the *2008 Simulation Interoperability Workshop*, Orlando, FL.