AVATALK Virtual Humans for Training with Computer Generated Forces

Robert C. Hubal, M.S., Ph.D.
Geoffrey A. Frank, Ph.D.
Curry I. Guinn, Ph.D.
Research Triangle Institute
3040 Cornwallis Rd.
Research Triangle Park, NC 27709-2194
919/541-6045
vr-info@rti.org

Keywords:
Intelligent agents, natural language, technology assisted learning, virtual humans.

ABSTRACT: *Research Triangle Institute's (RTI) AVATALK™ technology enables natural interaction with computer generated virtual humans, linking theory of human behavior with virtual reality, knowledge representation, and natural language processing. AVATALK virtual humans act realistically without human intervention--action takes such form as choice of utterances, conversational expectancy, and branching logic within the application. Underlying virtual human action is a dialog processor, a behavior engine, and a virtual environment adapter. AVATALK-enabled applications provide an opportunity for practice with numerous case-based scenarios in a reproducible, objective learning environment prior to the challenge of actual engagement. Training and sustainment benefits include enhanced adaptability, availability, and assessment, and reduced loss of effectiveness for learners at distributed locations. RTI has already developed AVATALK-enabled products for training to conduct data collection interviews, for tank maintenance training, and for Emergency Medical Technician training. AVATALK intelligent agents are a logical extension for training in the communication that occurs among forces.*

1. Introduction

In training using computer generated forces (CGF), the behaviors of the forces themselves and the digital communications that establish situation awareness among forces are well developed. We believe there is a gap, however, in automating voice communications among combined arms team members (e.g., between the trained force and CGF). To close this gap, we see intelligent agents (IA's) as providing for the CGF the ability to communicate command and control via voice.

Research Triangle Institute's (RTI) AVATALK™ technology represents a novel strategy for modeling cognitive and behavioral performance in creating IA's for use with CGF. IA's can be used as mentors, virtual peers, and coaches. As mentors, IA's can tailor scenarios for execution based on a review of plans and learner capabilities. As virtual peers, IA's can portray team members who deliver relevant information during problem solving. As coaches, IA's can observe, record, and evaluate learner actions and conduct off-line after-action reviews (AAR's).

In this paper, we discuss our approach to simulation-based training, describe the gap we see in training voice communications among forces, describe how AVATALK technology can assist in this training, and present a potential application for Fire Support training.

2. Simulation-based Training

Our approach to training is illustrated in Figure 2.1 (adapted from [1]); the approach holds for training technical skills such as maintenance and inspection as well as soft skills such as interaction and interviewing. We have found that for most skills, a combination of learning environments proves most cost-effective [2,3]. Specifically, providing students with a constructive or virtual learning environment enables them to become familiarized with materials, acquire and even practice their skills. This reduces the need for live equipment and on-the-job simulations (which are often costly, time consuming, and manpower intensive) to validation of skill performance.

Commonly, trainers use "knowledge, skills, attitudes" as a description of learning that must occur. We define these terms as follows:

- Acquiring knowledge is gaining the ability to declare it, that is, know the relevant information and be able to reiterate it. It is gained through familiarization. Declarative knowledge is factual, cognitive, well understood, basic information about skills and contexts. It can be taught in lectures, learned through reading or basic interactive multimedia instruction, or gained informally. It forms the basis for skills acquisition.
- Acquiring skills is less the ability to state knowledge than it is the ability to process or perform a task. Procedural knowledge can begin as declarative but gradually becomes automated ("proceduralized"), often as a motor skill, that demands decreasing exertion to accomplish. Procedural abilities are gained during acquisition, and they are automated during practice. By definition, practice requires repeated performance in an environment that alters to reflect performance outcomes. Acquisition of technical skills is efficiently done first in virtual worlds, later in constructive environments (e.g., part-task trainers) [4]. For acquisition of tactical skills, virtual training (e.g., CCTT) and constructive (e.g., JANUS, BBS) environments can be interleaved [5].
- In understanding when and how to apply knowledge, and in realizing gaps in knowledge that need to be filled for particular tasks, a student demonstrates strategic knowledge. Understanding of how knowledge affects attitudes and emotions is known as affective knowledge. In performing the skills under realistic conditions, a student experiences the affect, or emotional content and attitudes, associated with the skill. That is, one strategically applies declarative or procedural knowledge (or both) and experiences consequences. The validation of skills that proves they have been successfully acquired must take place in full-up simulators or using live equipment.

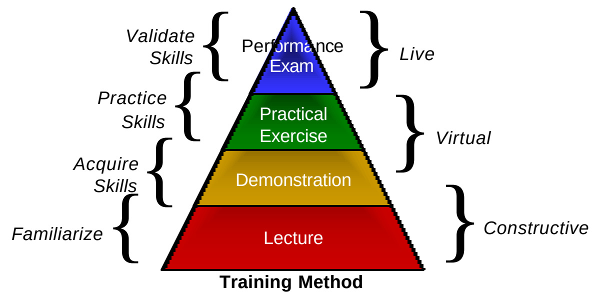

Figure 2.1 Training Triangle

The Training Triangle shows a relationship between level of proficiency and relative amount of time spent in each learning environments: traditional pedagogical (i.e., lecture), constructive, virtual, and live. There are several main points regarding the figure:

- Time spent in learning environments decreases as level of proficiency increases. The most time is spent in a traditional environment being familiarized with the tasks. The least time is spent in live environments validating attitudes and skills. Virtual and constructive environments enable effective acquisition and practice of skills. The right mix of learning environments will change for any given skill set, but this model has proven to be cost-effective for multiple applications.
- There may not be a need to validate all skills. Instead, there may be a need to validate some skills, practice to a lesser degree other skills, and be familiarized with still other skills. For instance, supervisors may only need basic understanding of course material, but broader knowledge, while other students may need in-depth knowledge of a narrower set of material. Similarly, if a set of tasks requires comparable skills, then performance on only a small subset of those tasks needs to be validated. For the remaining set, familiarization and acquisition should prove sufficient, so that the skills will successfully be applied to them on the job. The identification up-front of analogous skills is important for realizing cost-effectiveness from the mix of learning environments.
- The role of instructors changes as students advance in proficiency. Initially they are instructors, disseminating important information to relatively passive learners. Gradually, the students begin constructing and becoming engaged in their learning environments, and instructors transition into facilitators in the learning process. Eventually, the students have acquired the skills and know-how to apply the skills, and the instructors become

validators of those skills, and finally mentors for the proficient students.

In our approach we place an emphasis on establishing the right mix of live, virtual, and constructive (LVC) training, based on cost, efficiency, and effectiveness considerations.

3. The Gap

Experience with Advanced Warfighting Experiments (AWE's) has shown the importance of voice communications even with digital forces [6]. Lessons learned in the Focused Dispatch and Warrior Focus AWE's emphasized that digital communications are particularly valuable during planning before the battle and during mop-up operations and consolidation after the battle. Some functions have been extensively digitized, such as situation awareness (i.e., the tracking of friendly forces) and sensor to shooter connections for fire support. However, during the times of highest stress, commanders and staff need to use voice communications for coordination, command, and control. Realistic battle command simulations must allow coordination between the live staff in training and the friendly virtual forces, including subordinate commanders, superior commanders and staff, combined arms units such as combat support and combat service support, and virtual peer units such as units on either flank. These interactions are typically terse and highly focused on the mission, making them suitable for modeling with AVATALK.

The other important driver for the use of voice communications as part of combined arms training is that most forces that will be fielded will be hybrid units that are only partially digitized. This includes forces that combine Active Component digital forces with Reserve Component analog forces, and joint operations with forces from other nations. Voice communications will be a critical link between the digital and analog forces. Voice radio communications will be essential for establishing the locations of allied units, as well as for command and control.

The current solution for providing voice dialogs with commanders of allied units is to employ a large staff to support the training exercise. These staff requirements reduce the cost-effectiveness of CGF simulations for training. As CGF technology becomes more tactically sophisticated, the ability to provide voice interactions will eliminate a key barrier to achieving the cost-saving potential of CGF simulations for training.

4. AVATALK: Technology to Bridge the Gap

AVATALK technology enables natural interaction with computer generated virtual humans, linking theory of human behavior with virtual reality, knowledge representation, and natural language processing. AVATALK virtual humans act realistically without human intervention. Underlying virtual human action is a dialog processor, a behavior engine, and a virtual environment adapter.

The AVATALK suite of technologies involves:

- *Natural language processing* (NLP), which incorporates the ability to recognize natural, unscripted speech and to understand speech based on the content of the discourse. The application is designed to expect relevant, reasonable speech from the soldier, similar to that which occurs during regular conversation. As in regular conversation, expectations mature as conversation progresses.
- *Emotion and behavior modeling*. AVATALK virtual humans act realistically as if they are angry, frightened, serene, or in pain. Action takes the form of choice of utterances, conversational expectancies, and branching logic within the application.
- *Dynamic virtual worlds* in which activities occur and contextual cues reside. Conversation does not take place in a vacuum; instead, the environment plays a large role in shaping conversational flow.

AVATALK-enabled applications provide an opportunity for practice with numerous case-based scenarios in a reproducible, objective learning environment prior to the challenge of actual engagement. Training and sustainment benefits include enhanced adaptability, availability, and assessment, and reduced loss of effectiveness for learners at distributed locations. RTI has already developed AVATALK-enabled products for training to conduct data collection interviews [7], for tank maintenance training [8], as virtual standardized patients [9], and for Emergency Medical Technician training [10].

We stress that, although the scenarios are pre-defined, the interaction itself is unscripted. The scenario establishes initial conditions, but the soldier's responses to the virtual human, as well as inherent flexibility in how the virtual human is allowed to react, cause the conversational flow to vary from interaction to interaction. We believe this leads to a realistic learning application wherein the soldier must learn to handle each interaction individually.

AVATALK software was designed to be usable and versatile. Its ease of use makes it ideal for both centralized or decentralized training. It is useful for initial training, refresher training, and ongoing assessment of interaction skills. Because the software was designed to run on a relatively inexpensive laptop computer, it can be used on many home-station computers, with distribution via compact disc or a network.

5. Applications of AVATALK

RTI has created or is creating highly effective and engaging training applications using AVATALK technology. We present three examples.

5.1 AMAT

With ACT II funds, RTI developed the Advanced Maintenance Assistant and Trainer (AMAT) for delivery to the Combat Service Support Battle Lab in Fort Lee, VA [8]. AMAT is a spoken-dialogue assistant and trainer for the maintenance of line replaceable units in the M1A1 Abrams tank. It is an extension of the Virtual Maintenance Trainer (VMAT) developed at RTI for the Army National Guard. VMAT enables National Guard mechanics to train and update their skills by working on a virtual M1A1 tank.

In AMAT, the soldier conducts a dialog with a virtual coach, who provides verbal cues on how to find appropriate sections within Technical Manuals. Using AMAT, the mechanic can access important diagnostic information and procedures using voice input and output. During training, AMAT also allows the soldier to speak to the system to manipulate the view in the virtual tank.

RTI installed AMAT on a wearable computer, thus providing hands-free operation with a conversant Integrated Electronic Technical Manual.

5.2 AVATALK-Survey

RTI created a survey interviewer training module, AVATALK-Survey, that addresses a critical training need in survey research: survey nonresponse. Nonresponse to household surveys is increasing, despite extraordinary measures being taken to counter the trend [7].

Research suggests that, to train how to solicit participation in surveys, effective training programs must address respondent concerns, train interviewers to develop strategies to adapt to cues provided by the respondent, and create a realistic learning environment [7]. Current survey practice does not offer a solution, instead leading interviewers to follow complex, standardized interviewing procedures. AVATALK-Survey, on the other hand, is a solution. The application (see Figure 5.2) generates a variety of respondents showing a range of emotions, creates a virtual environment in which contextual cues can be added or changed, and can be used for home study to supplement current training agendas. There is also a coach, another virtual human, who can act as observer, facilitator, or mentor.

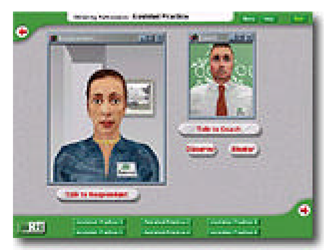

Figure 5.2 AVATALK-Survey Interface

5.3 JUST-TALK

RTI is developing an application for the National Institute of Justice, JUST-TALK, using AVATALK technology to provide a computerized virtual person that interacts with the trainee in a similar way to the role-playing approach to training. It is designed to train civilian police officers to handle mentally disturbed individuals. It will be available on demand for practice, will allow self-paced study, and be usable on a home computer or on a computer at the police station.

An officer responding to a situation where the subject is mentally ill must quickly make a number of difficult decisions. The officer's actions must be consistent with federal, state, and local laws and with police department policies. Officer response has to be immediate and confident; uncertainty can either produce a loss of control or result in inappropriate actions. Effective training is key to providing the required confidence and ensuring that appropriate actions are taken.

The emphasis in JUST-TALK is on applying the technology to support law enforcement training, and determining the best way to deliver the technology to the many small police departments in the country. The technology is applicable to other related training problems where effective interviews with the subjects are a critical skill for successful resolution of incidents.

6. Fire Support Training: A Potential Application

To illustrate how the NLP capabilities of AVATALK could be used to support training with CGF, we describe a potential application for training fire support coordination teams for the field artillery. Figure 6.1 illustrates the concept. The application would use:

- CGF simulators together with the battle simulation system. We assume that an instructor will be commanding the computer-generated OPFOR to provide more challenging responses to student tactics.
- A tactical device (or emulator). For the fire support coordinator, this would be an AFATDS system.
- A digital interface between the HLA simulation environment and the COE/VMF environment used by the tactical device. These interfaces are being used in simulation centers now.
- A voice interface which electronically links the CGF simulators with a voice communications device (such as a radio or radio emulator). This interface contains the AVATALK software and one or more dialog databases that are customized for the particular scenario. This interface is a two-way interface. A particular simulation event (such as contact with an enemy force) triggers a dialog initiated by AVATALK (such as a call for fire dialog). Similarly, AVATALK listens to audio inputs from the student, and will provide appropriate responses based on scripts in the dialog database.

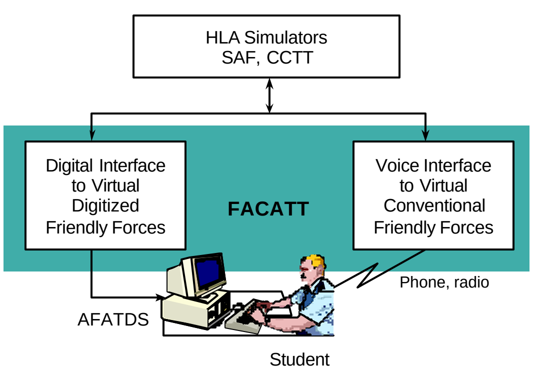

Figure 6.1 Using AVATALK Natural Language Processing to support Fire Support Training.

This training system would provide a computer-based capability for adapting combined arms training scenarios to FA missions in environments that include the schoolhouse, garrison, and armory. If the AVATALK interface was connected to a radio such as SINCGAR, it would allow combined live and constructive simulation-based training.

When a verbal message is received by the communications device, it is translated by the interface into digital form. The natural language processing software uses the dialog databases to maintain the context of the communications and monitors the HLA environment to ensure an appropriate response to the message. The AVATALK software database processes the message received to determine its content and then prepares a verbal (natural language) response which is sent via the interface to the station's voice communications device for transmission to the appropriate station. These transmissions continue until the need for communications is satisfied.

The capabilities of a system like this will provide a path for the transition to embedded training by allowing FA personnel and units to use existing equipment, command and control systems, electronic training environments, and simulations to achieve realistic training with less support infrastructure. They will:

- Link via a HLA compatible CGF simulator such as CCTT to an interactive visual terrain environment and electronically download combined arms training scenarios
- Reduce the time and manpower required for preparing, coordinating, and electronically downloading combined arms tactical training scenarios.
- Provide scenario and verbal communications inputs and responses for members of combined arms team that are necessary to conduct meaningful FA training; including enemy and friendly higher, adjacent and subordinate units, terrain data, advanced weapon systems models, and distributed simulations.
- Support data collection, capturing inputs, building and maintaining files, and after-action assessments. The system would automatically transcribe voice communications into digital form for analysis during AAR's and for experimental analysis of training and doctrine.
- Support the use of CGF simulations for seamless

garrison, field and contingency training.

7. Conclusion

We have described a gap in automating voice communications among combined arms team members. To close this gap, where CGF are used, we see AVATALK intelligent agents as providing the ability to communicate command and control via voice. Within the training environment itself, IA's can also act as coaches and mentors. We are prepared to apply the technology to computer generated forces.

8. References

[1] Hubal, R.C., & Helms, R.F. (1998). Advanced Learning Environments. Modern Simulation & Training, 5, 40-45.

[2] Helms, R.F., Hubal, R.C., & Triplett, S.E. (1997). Evaluation of the Conduct of Individual Maintenance Training in Live, Virtual, and Constructive (LVC) Training Environments and their Effectiveness in a Single Program of Instruction. Final Report, September 30, 1997. Submitted to Battelle RTP Office, Subcontract # TCN 97031, Delivery Order #0027, Dated April 16, 1997.

[3] Frank, G.A., Helms, R.F., & Hubal, R.C. (1999). Benefits of an Advanced Learning Environment for Longbow AH-64D Maintenance Training. Information Paper developed for Army Training & Doctrine Command. April 22, 1999.

[4] Field, S.S., Frank, G.A., Helms, R.F., & Hubal, R.C. (1999). Army Learning & Training Effectiveness Symposium. Final Report, March 30, 1999. Submitted to Battelle RTP Office, Subcontract #DAAH04-96-C-0086, Agreement #99020, Delivery Order #0366, Dated February 12, 1999.

[5] Frank, G.A., & Helms, R.F. (2000). Army Learning & Training Effectiveness Symposium II (Collective). Final Report, January 9, 2000. Submitted to Battelle RTP Office, Subcontract #DAAH04-96-C-0086, Agreement #99134, Delivery Order #0481.

[6] Training System Manager Force XXI. (April 1997). Force XXI Overview. Ft. Knox, KY.

[7] Camburn, D.P., Gunther-Mohr, C., & Lessler, J.T., (1999). Developing New Models of Interviewer Training. International Conference on Survey Nonresponse, Portland, OR, October 28-31, 1999.

[8] Guinn, C.I., & Montoya, R.J. (1998). Natural Language Processing in Virtual Reality. Modern Simulation & Training, 6, 44-55.

[9] Hubal, R.C., Kizakevich, P.N., Guinn, C.I., Merino, K.D., & West, S.L. (2000). The Virtual Standardized Patient–Simulated Patient-Practitioner Dialogue for Patient Interview Training. In J.D. Westwood, H.M. Hoffman, G.T. Mogel, R.A. Robb, & D. Stredney (Eds.), Envisioning Healing: Interactive Technology and the Patient-Practitioner Dialogue, 133-138. IOS Press: Amsterdam.

[10] Kizakevich, P.N., McCartney, M.L., Nissman, D.B., Starko, K., & Smith, N.T. (1998). Virtual Medical Trainer: Patient Assessment and Trauma Care Simulator. In J.D. Westwood, H.M. Hoffman, D. Stredney, & S.J. Weghorst (Eds.), Medicine Meets Virtual Reality - Art, Science, Technology: Healthcare (R)evolution, 309-315. IOS Press: Amsterdam.

[11] Helms, R.F., & Frank, G.A. (1997). Final Design Report, Advanced Experimental Environment. Submitted to U.S. Army Depth and Simultaneous Attack Battle Lab, Ft. Sill, OK, Contract #DABT60-97-D-0007.

Author Biographies

ROBERT HUBAL conducts research on technology assisted learning, focusing on development, presentation, and evaluation of materials and identifying approaches to improve learning and training effectiveness. He was technical lead for an Army National Guard experiment to examine the cost-effectiveness of live, virtual, and constructive training environments [2]. He has investigated the interaction effects among cognitive task demands, expertise, and presentation of information. Most recently, as co-developer of RTI's AVATALK™ technology, he has developed behavioral software that enables virtual humans to act and behave realistically in controlled learning contexts. Dr. Hubal holds undergraduate and masters degrees in Computer Science and a doctorate in Cognitive Psychology.

GEOFFREY FRANK is co-author of a chapter on virtual reality training in the American Society of Training Developer's Handbook, and has written a military handbook for the Army Material Command on electronic systems design. He was the Project Leader for the Army Classroom XXI for Leader Development project, for the ACT II WarLab XXI; and for the ACT II MPRTS. He also was task leader for establishing electronic linkages between the Advanced Experiment Environment and DIS-compliant JANUS at the D&SABL at Ft Sill, OK [11]. He has nearly 15 years of experience in simulation, software engineering, and software tool development. Dr. Frank holds a doctorate in Computer Science.

CURRY GUINN is a research engineer and co-developer of AVATALK™. Using pioneering spoken human-computer dialogue algorithms, he has integrated advanced spoken dialogue capabilities in virtual reality based trainers on a variety of platforms: off-the-shelf PCs, wearable computers, hand-held computers, and high-end workstations. His natural language processing engine is currently being used by Duke University in its Pascal Programming Tutor project, and his work on a multimedia tutoring system for teaching introductory computer science is currently in use as part of the curriculum, also at Duke University. Dr. Guinn holds a doctorate in Computer Science.