Toward a New Generation of Virtual Humans for Interactive Experiences

Jeff Rickel and Stacy Marsella, *USC Information Sciences Institute*

Jonathan Gratch, Randall Hill, David Traum, and William Swartout, *USC Institute for Creative Technologies*

Virtual humans—autonomous agents that support face-to-face interaction in a variety of roles—can enrich interactive virtual worlds. Toward that end, the Mission Rehearsal Exercise project involves an ambitious integration of core technologies centered on a common representation of task knowledge.

Imagine yourself as a young lieutenant in the US Army on your first peacekeeping mission. You must help another group, called Eagle1-6, inspect a suspected weapons cache. You arrive at a rendezvous point, anxious to proceed with the mission, only to see your platoon sergeant looking upset as smoke rises from one of your platoon's vehicles and a civilian car. A seriously injured child lies on the ground, surrounded by a distraught woman and a medic from your team. You ask what happened and your sergeant reacts defensively. He casts an angry glance at the mother and says, "They rammed into us, sir. They just shot out from the side street and our driver couldn't see them." Before you can think, an urgent radio call breaks in: "This is Eagle1-6. Where are you guys? Things are heating up. We need you here now!" From the side street, a CNN camera team appears. What do you do now, lieutenant?

Interactive virtual worlds provide a powerful medium for entertainment and experiential learning. Army lieutenants can gain valuable experience in decision-making in scenarios like the example above. Others can use the same technology for entertaining role-playing even if they never have to face such situations in real life. Similarly, students can learn about, say, ancient Greece by walking through its virtual streets, visiting its buildings, and interacting with its people. Scientists and science fiction fans alike can experience life in a colony on Mars long before the required infrastructure is in place. The range of worlds that people can explore and experience with virtual-world technology is unlimited, ranging from factual to fantasy and set in the past, present, or future.

Our goal is to enrich such worlds with virtual humans—autonomous agents that support face-to-face interaction with people in these environments in a variety of roles, such as the sergeant, medic, or even the distraught mother in the example above. Existing virtual worlds, such as military simulations and computer games, often incorporate virtual humans with varying degrees of intelligence. However, these characters' ability to interact with human users is usually very limited: Typically, users can shoot at them and they can shoot back. Those characters that support more collegial interactions, such as in children's educational software, are usually very scripted and offer human users no ability to carry on a dialogue. In contrast, we envision virtual humans that cohabit virtual worlds with people and support face-to-face dialogues situated in those worlds, serving as guides, mentors, and teammates.

Although our goals are ambitious, we argue here that many key building blocks are already in place. Early work on embodied conversational agents[1] and animated pedagogical agents[2] has laid the groundwork for face-to-face dialogues with users. Our prior work on Steve[3,4] (see Figure 1) is particularly relevant. Steve is unique among interactive animated agents because it can collaborate with people in 3D virtual worlds as an instructor or teammate. Our goal with Steve is to integrate the latest advances from separate research communities into a single agent architecture. While we continue to add more sophistication, we have already implemented a core set of capabilities and applied it to the Army peacekeeping example described earlier.

Mission Rehearsal Exercise

Steve supports many capabilities required for face-to-face collaboration with people in virtual worlds. Like earlier intelligent tutoring systems, he can help students by providing feedback on their actions and by answering questions, such as "What should I do next?" and "Why?" However, because Steve has an animated body and cohabits the virtual world with students, he can interact with them in ways that previous disembodied tutors could not. For example, he can lead students around the virtual world, demonstrate tasks, guide their attention through his gaze and pointing gestures, and play the role of a teammate whose activities they can monitor.

Steve's behavior is not scripted. Rather, Steve consists of a set of general, domain-independent capabilities operating over a declarative representation of domain tasks. We can apply Steve to a new domain simply by giving him declarative knowledge of the virtual world—its objects, their relevant simulator state variables, and their spatial properties—and the tasks that he can perform in that world. We give task knowledge to Steve using a relatively standard hierarchical plan representation. Each task model consists of a set of steps (each a primitive action or another task), a set of ordering constraints on those steps, and a set of causal links. The causal links describe each step's role in the task; each link specifies that one step achieves a particular goal that is a precondition for a second step (or for the task's termination). Steve's general capabilities use such knowledge to construct a plan for completing a task from any given state of the world, revise the plan when the world changes unexpectedly, and maintain a collaborative dialogue with his student and teammates.

Steve's capabilities are well-suited for training people on complex, well-defined tasks, such as equipment operation and maintenance. However, virtual worlds like the peacekeeping example introduce new requirements. To create a more engaging and emotional experience for users, agents must have realistic bodies, emotions, and distinct personalities. To support more flexible interaction with users, agents must use sophisticated natural language capabilities. Finally, to provide more realistic perceptual capabilities and limitations for dynamic virtual worlds, agents such as Steve need a human-like model of perception.

We are addressing these requirements in an ambitious new project called the Mission Rehearsal Exercise (MRE).[5] In addition to extending Steve with several new domain-independent capabilities, we implemented the peacekeeping scenario as an example application to guide our research. Figure 2 shows a screen shot from our current implementation. The system displays the visual scene on an eight-foot-tall screen that wraps around the user in a 150-degree arc with a 12-foot radius. Immersive audio software uses 10 audio channels and two subwoofer channels to envelop a participant in spatialized sounds that include general ambience (such as crowd noise) and triggered effects (such as explosions or helicopter flyovers). We render the graphics, including static scene elements and special effects, with Multigen/Paradigm's Vega. The simulator itself—or a human operator using a graphical interface—triggers external events, such as radio transmissions from Eagle1-6, a medevac helicopter, and a command center.

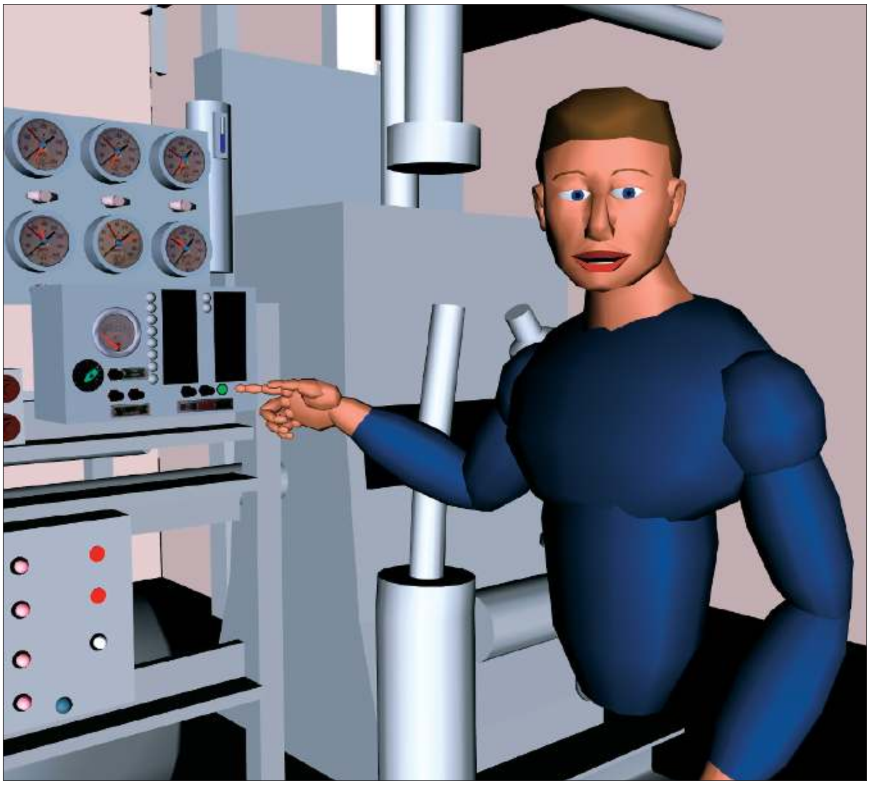

Figure 1. Steve, an interactive agent that functions as a collaborative instructor or teammate in a virtual world, describes a power light.

Three Steve agents interact with the human user (who plays the role of lieutenant): the sergeant, the medic, and the mother. All other virtual humans (a crowd of locals and four squads of soldiers) are scripted characters implemented in Boston Dynamics' PeopleShop. The lieutenant talks with the sergeant to assess the situation, issue orders (which the sergeant carries out through four squads of soldiers), and ask for suggestions. The lieutenant's decisions influence the way the situation unfolds, culminating in a glowing news story praising the user's actions or a scathing news story exposing decision flaws and describing their sad consequences.

This sort of interactive experience clearly has both entertainment and training applications. The US Army is well aware of the difficulty of preparing officers to face such dilemmas in foreign cultures under stressful conditions. By training in immersive, realistic worlds, officers can gain valuable experience. The same technology could power a new generation of games or educational software, letting people experience exciting adventures in roles that are richer and more interactive than those that current software supports.

In the remainder of this article, we describe the key areas in which we are extending Steve. Although there has been extensive prior research in each of these areas, we cannot simply plug in different modules representing the state of the art from the different research communities. Researchers devel-

Figure 2. An interactive peacekeeping scenario featuring (from left to right) a sergeant, a mother, and a medic.

oped the state of the art in each area independently from the others, so our fundamental research challenge is to understand the dependencies among them. Our integration revolves around a common representation for task knowledge. In addition to providing a hub for integration, this approach provides generality: By changing the task knowledge, we can apply our virtual humans to new virtual worlds in different domains.

Virtual human bodies

Research in computer graphics has made great strides in modeling human body motion. Most relevant to our objectives is the research focusing on real-time control of human figures. Within that area, some work uses forward and inverse kinematics to synthesize body motions dynamically so that they can achieve desired end positions for body parts while avoiding collisions with objects along the way. Other work focuses on dynamically sequencing motion segments created with key-frame animation or motion capture. This approach achieves more realistic body motions at the expense of flexibility. Both approaches have reached a sophisticated level of maturity, and much current research focuses on combining them to achieve both realism and flexibility.

We designed Steve to accommodate different bodies. His motor-control module accepts abstract motor commands from his cognition module and sends detailed commands to his body through a generic API. Integrating a new body into a Steve agent simply requires adding a layer of code to map that API onto the body's API. Steve's original body, developed by Marcus Thiébaux at the USC Information Sciences Institute, generates all motions dynamically using an efficient set of algorithms. However, the body does not have legs—Steve moves by floating around—and its face has a limited range of expressions that do not support synchronizing lip movements to speech. Furthermore, Steve's repertoire of arm gestures is limited to pointing. By integrating a more advanced body onto Steve, we can achieve more realistic motion with little or no modification to Steve's other modules.

For MRE, we integrated new bodies developed by Boston Dynamics Incorporated, as shown in Figure 2. The bodies use the human figure models and animation algorithms from PeopleShop, but BDI extended them in several ways to suit our needs. First, BDI integrated new faces (developed by Haptek Incorporated) to support lip-movement synchronization and facial expressions. Second, while the basic PeopleShop software primarily supports dynamic sequencing of primitive motion fragments, BDI combined their motion-capture approach with procedural animation to provide more flexibility, primarily in the areas of gaze and arm gestures. The new gaze capability lets Steve direct his body to look at any arbitrary object or point in the virtual world. In addition to specifying the gaze target, we can specify the manner of the gaze, including the speed of different body parts (eyes, head, and torso) and the degree to which they orient toward the target.

The gesture capability uses an interesting hybrid of motion capture and procedural animation. Motion capture helped create the basic repertoire of gestures. However, we decompose gestures into stages (such as preparation, stroke, and retraction), and can change the timing of each gesture stage to achieve different effects. Moreover, gestures typically have two extremes: a small, restrained version of the gesture and a large, emphatic version. A Steve agent can dynamically generate any gesture between these extremes by specifying a weighted combination of the two. Thus, these new bodies leverage the realism of motion capture while providing the flexibility of procedural animation.

Task-oriented dialogue

Spoken dialogue is crucial for collaboration. Students must be able to ask their virtual human instructors a wide range of questions. Teammates must communicate to coordinate their activities, including giving commands and requests, asking for and offering status reports, and discussing options. Without spoken-dialogue capabilities, virtual humans cannot fully collaborate with people in virtual worlds.

Steve used commercial speech recognition and synthesis products to communicate with human students and teammates. However, like many virtual humans, it had no true natural-language-understanding capabilities and understood only a relatively small set of preselected phrases. There are two problems with this approach that make it impractical for ambitious applications such as MRE. First, it is very labor intensive to add broad coverage, since each phrase and semantic representation must be individually added. For example, in a domain such as the MRE, the same concept (such as "the third squad") could be referred to in multiple ways (such as "third squad," "where are they," "what are they doing"). It would be very cumbersome to add a separate phrase for each referring expression and predicate about the concept. Moreover, adding a new similar concept (such as, "fourth squad") requires duplicating all this effort. In contrast, a grammar-based approach allows one to achieve the same coverage by merely adding individual lexical entries to an existing grammar. In addition to increasing efficiency, this approach can help achieve better understanding when the system recognizes only parts of a complete phrase.

The second problem with using a phrase-based speech recognizer without full natural-language dialogue capabilities is that all the recognition is done by the speech recognizer itself rather than by modules that

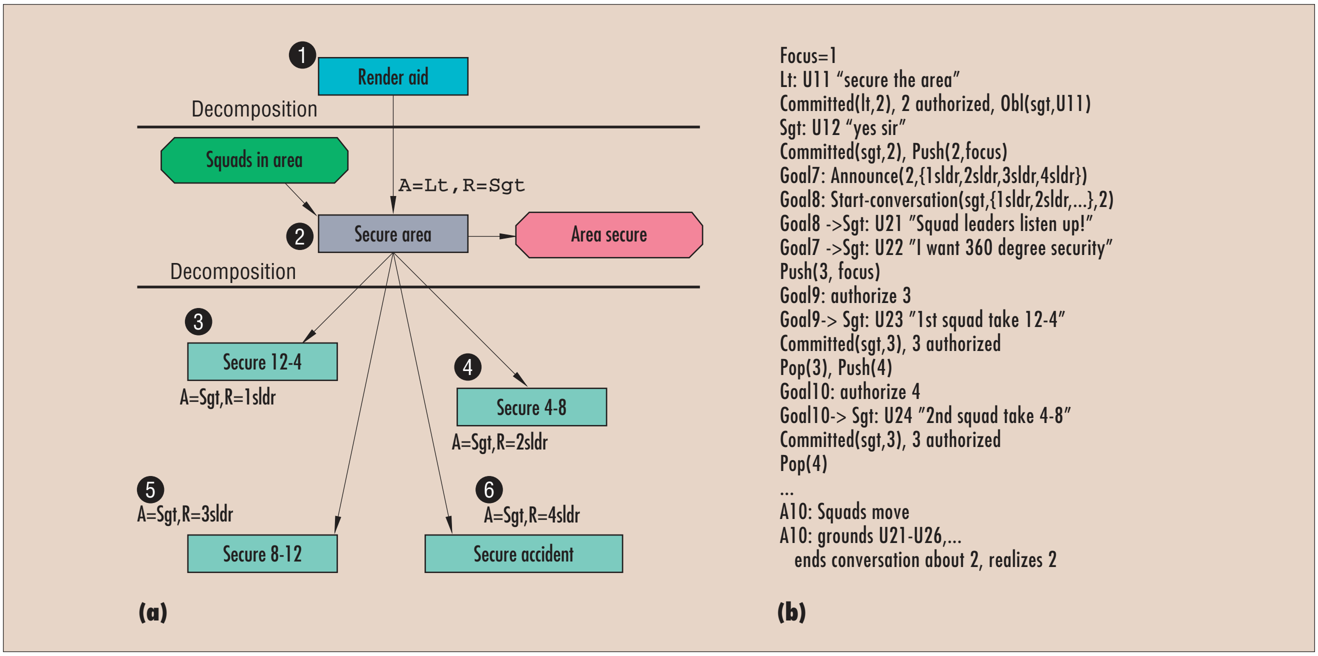

Figure 3. A task model fragment with dialogue-related behavior.

have access to the evolving task and dialogue context. A rich task model—such as the one Steve uses—can help choose a more sensible interpretation than a speech recognizer alone.

For these reasons, in the MRE project we extended the spoken dialogue capabilities to include

- A domain-specific finite-state speech recognizer with a vocabulary of several hundred words, allowing recognition of thousands of distinct utterances
- A finite-state semantic parser that produces partial semantic representations of information expressed in text strings
- A dialogue model that explicitly represents aspects of the social context[6,7] and is oriented toward multiple participants and face-to-face communication[8]
- A dialogue manager that recognizes dialogue acts from utterances, updates the dialogue model, and selects new content for the system to say
- A natural language generator that can produce nuanced English expressions, depending on the agent's personality and emotional state as well as the selected content[9]
- An expressive speech synthesizer capable of speaking in different voice modes depending on factors such as proximity (speaking or shouting) and illocutionary force (command or normal speech)

These new components, along with Steve's task model, give our agents a very rich and flexible dialogue capability: The agent can answer questions and perform or negotiate about directives following user- or mixed-initiative strategies rather than forcing the user to follow a strict system-guided script.

Dialogue behavior and reasoning make crucial use of Steve's task model, both for determining the full intent of an underspecified command and for deciding what to do next in a given situation. Consider the task model fragment in Figure 3, which is annotated with a sequence of dialogue actions, dialogue state updates, and the sergeant's dialogue goals. The task model encodes social information for team tasks, including both who is responsible for the task (R) and who has authority to allow the task to happen (A). In this task model, the lieutenant has the responsibility for rendering aid. However, some subtasks, such as securing the local area, can be delegated to the sergeant, who, in turn, can delegate subtasks to squad leaders.

In Figure 3, the sergeant's task focus is initially on the Render Aid task. When the lieutenant issues the command to secure the area (utterance U11), the sergeant recognizes the command as referring to a subaction of Render Aid in the current task model (Task 2). As a direct effect of the lieutenant issuing a command to perform this task, the lieutenant becomes committed to the task, the sergeant has an obligation to perform the task, and the task becomes authorized. Because the sergeant already agrees that this is an appropriate next step, he is able to accept it with utterance U12, which also commits him to perform the action. The sergeant then pushes this task into his task model focus and begins plans to carry it out. In this case, because it is a team task requiring actions of other teammates, the sergeant, as team leader, must announce the task to the other team members. Thus, the system forms a communicative goal to make this announcement. Before the sergeant can issue this announcement, he must make sure he has the squad leaders' attention and has them engaged in conversation. He forms a goal to open a new conversation so that he can produce the announcement. Then his focus can turn to the individual tasks for each squad leader. As each one enters the sergeant's focus, he issues the command that commits the sergeant and authorizes the troops to carry it out. When the troops move into action, it signals that they understood the sergeant's order and adopted his plan. When the task completes, the conversation between sergeant and squad leaders finishes and the sergeant turns his attention to other matters.

Emotions

It is hard to imagine an entertaining or compelling character that doesn't express

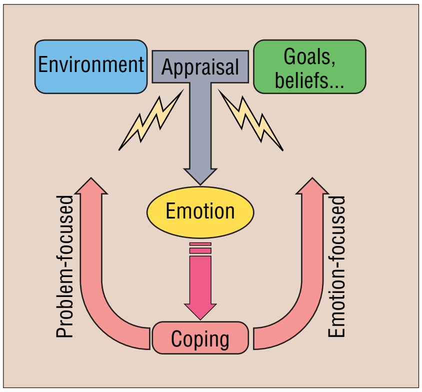

Figure 4. The model of emotion in which an agent appraises the emotional significance of events in terms of their impact on the agent's goals and plans.

emotion. Skilled animators use emotional behaviors to create a sense of empathy and drama and to fill their creations with a rich mental life. The growth in learning games builds on the theory that entertainment value translates into greater student enthusiasm for instruction and better learning. But beyond creating a sense of engagement, emotion appears to play a central role in teaching. Tutors frequently convey emotion to motivate or reprimand students.

Furthermore, unemotional agents such as the original version of Steve are simply unrealistically rational as teammates. In many training situations like the MRE, students must learn how their teammates are likely to react under stress, because learning to monitor teammates and adapt to their errors is an important aspect of team training. So, Steve's lack of emotions hampers his performance as an instructor and teammate, and it makes him less engaging for interactive entertainment applications.

Fortunately, research on computational emotion models has exploded in recent years. Some of that work is particularly well suited to the type of task reasoning that forms the basis of the Steve system.[10,11] We integrated these models with Steve and significantly broadened their scope.[12,13] This work is motivated by psychological theories of emotion that emphasize the relationship between emotions, cognition, and behavior.

Figure 4 illustrates the basic model. The appraisal process at the emotional model's core results in an emotional state that changes in response to changes in the environment or to changes in an agent's beliefs, desires, or intentions. Verbal and nonverbal cues manifest this emotional state through facial displays, gestures, and other kinds of body language, such as fidgeting, gaze aversion, or shoulder rubbing. But emotions don't serve merely to modulate surface behavior. They are also powerful motivators. People typically cope with emotions by acting on the world or by acting internally to change their goals or beliefs. For example, the sergeant's defensive response at the beginning of this article is a classic example of emotion-focused coping—shifting blame in response to strong negative emotions. The agent's coping mechanism models this kind of behavior by forming intentions to act or altering internal beliefs and desires in response to strong emotions.

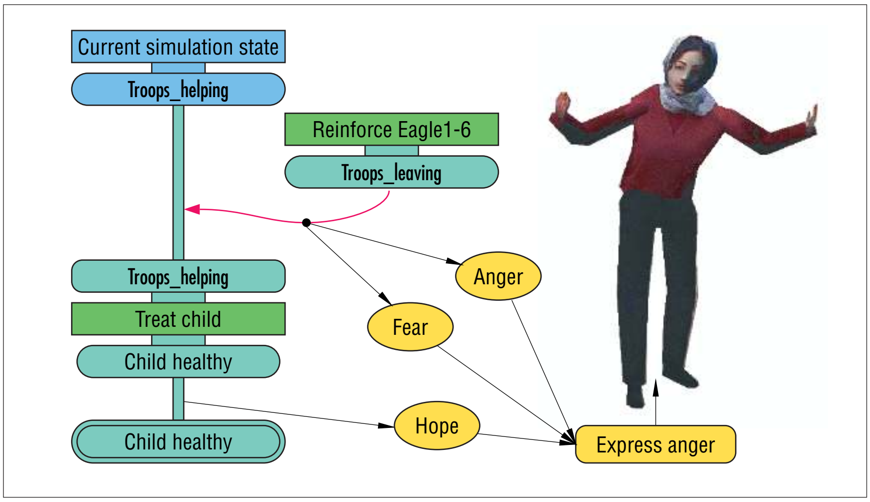

Figure 5. An emotional appraisal that builds on Steve's explicit task models. The appraisal mechanism assesses the relationship between the events and the agent's goals.

Figure 5 illustrates some details of the appraisal process, which characterizes events in terms of several features—such as valence, intensity, and responsibility—that are mapped to specific emotions. For example, an action in the world that threatens an agent's goals would cause an agent to appraise the action as undesirable. If the action might occur in the future, an agent would appraise it as an unconfirmed threat, which would be subsequently mapped to fear. If the action has already occurred, the confirmed threat would be mapped to distress. Figure 5 shows a task representation from the mother's perspective in the peacekeeping scenario. She wants her child to be healthy and believes that will occur if the troops stay and treat the child. This potentially beneficial action leads to an expression of hope. If the troops leave, their leaving threatens her plans and leads to appraisals of distress and anger.

In contrast, we can look at coping as the inverse of appraisal. For example, to discharge a strong emotion about some situation, one obvious strategy would be to change one or more of the appraised factors that contribute to the emotion. Coping operates on the same representations as the appraisals—the agent's beliefs, goals, and plans—but in reverse to make a direct or indirect change that would have a desirable impact on the original appraisal. For example, the sergeant feels distress because he is potentially responsible for an action with an undesirable outcome (the boy is injured). He could cope with the stress by reversing the undesirable outcome (perhaps by forming an intention to help the boy) or by shifting blame. Currently, we use a crude personality-trait model to assert preferences over alternative coping strategies. We are working on extending this model.

Appraisal and coping work together to create dynamic external behavior—appraisal leads to coping behaviors that in turn lead to a reappraisal of the agent-environment relationship. This model of appraisal, coping, and reappraisal begins to approach the richness and subtle dynamics necessary to create entertaining and engaging characters.

Humanlike perception

One obstacle to creating believable interaction with a virtual human is omniscience. When human participants realize that a character in a computer game or a virtual world can see through walls or instantly knows events that have occurred well outside its perceptual range, they lose the illusion of reality and become frustrated with a system that gives the virtual humans an unfair advantage. Many current applications of virtual humans finesse the issue of how to model perception and spatial reasoning. Steve's original version was no exception. Steve was omniscient; he received messages from the virtual world simulator describing every state-change relevant to his task model, regardless of his current location or attention state. Without a realistic model of human attention and perception, we had no principled basis to limit Steve's access to these state changes.

Recent research might provide that principled basis. Randall Hill, for example, has developed a model of perceptual resolution based on psychological theories of human perception.[14,15] Hill's model predicts the level of detail at which an agent will perceive objects and their properties in the virtual world. He applied his model to synthetic fighter pilots in simulated war exercises. Complementary research by Sonu Chopra-Khullar and Norman Badler provides a model of visual attention for virtual humans.[16] Their work, which is also based on psychological research, specifies the types of visual attention required for several basic tasks (such as locomotion, object manipulation, or visual search), as well as the mechanisms for dividing attention among multiple tasks.

We have begun to put these principles into practice, beginning with making Steve's model of perception more realistic. We implemented a model that simulates many of the limitations of human perception, both visual and aural, and limited Steve's simulated visual perception to 190 horizontal degrees and 90 vertical degrees. The level of detail Steve perceives about objects is high, medium, or low, depending on where the object is in Steve's field of view and whether Steve is giving attention to it. Steve can perceive both dynamic and static objects in the environment. Steve perceives dynamic objects, under the control of a simulator, by filtering updates that the system periodically broadcasts.

The simulator does not represent some objects, such as buildings and trees, in the same way that it represents dynamic objects, which means Steve won't perceive them in the same manner that he perceives dynamic objects. Instead, Steve perceives the locations of buildings, trees, and other static objects by using the scene graph and an edge-detection algorithm to determine the locations of these objects. The system encodes this information in a cognitive map along with the locations of exits, which Steve can infer using a space-representation algorithm.[17] We will eventually use the cognitive map for way-finding and other spatial tasks.[18]

We model aural perception by estimating the sound pressure levels of objects in the environment and determining their individual and cumulative effects on each listener on the basis of the distances and directions of the sources. This lets the agents perceive aural events involving objects not in the visual field of view. For example, the sergeant can perceive that a vehicle is approaching from behind, prompting him to turn around and look to see whether it is the lieutenant. Another effect of modeling aural perception is that some sound events can mask others. A helicopter flying overhead can make it impossible to hear someone speaking in normal tones a few feet away. The noise might prompt the virtual speaker to shout and might also prompt the Steve agent to cup his ear to indicate that he cannot hear.

Virtual humans that support rich interactions with people pave the way toward a new generation of interactive systems for entertainment and experiential learning.

We have made great progress toward virtual humans that collaborate with people in virtual worlds, but much work remains. The implemented peacekeeping scenario serves as a valuable test bed for our research. It provides concrete examples of challenging research issues and serves as a basis for evaluating our virtual humans with real users. We continue to extend Steve's individual capabilities, but we are especially focusing on the interdependencies among them, such as how emotions and personality affect each capability and how Steve's external behavior reflects its internal cognitive and emotional state. While our goals are ambitious, the potential payoff is high: Virtual humans that support rich interactions with people pave the way toward a new generation of interactive systems for entertainment and experiential learning.

Acknowledgments

The Office of Naval Research funded the original research on Steve under grant N00014-95-C-0179 and AASERT grant N00014-97-1-0598. The Department of the Army funds the MRE project under contract DAAD 19-99-D-0046. We thank our many colleagues who are contributing to the project. Shri Narayanan leads a team working on speech recognition. Mike van Lent, Changhee Han, and Youngjun Kim are working with Randall Hill on perception. Ed Hovy, Deepak Ravichandran, and Michael Fleischman are working on natural-language understanding and generation. Lewis Johnson, Kate LaBore, Shri Narayanan, and Richard Whitney are working on speech synthesis. Kate LaBore is developing the behaviors of the scripted characters. Larry Tuch wrote the story line with creative input from Richard Lindheim and technical input on Army procedures from Elke Hutto and General Pat O'Neal. Sean Dunn, Sheryl Kwak, Ben Moore, and Marcus Thiébaux created the simulation infrastructure. Chris Kyriakakis and Dave Miraglia are developing the immersive sound effects. Jacki Morie, Erika Sass, Michael Murguia, and Kari Birkeland created the graphics for the Bosnian town. Any opinions, findings, and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the Department of the Army.

References

1. J. Cassell et al., *Embodied Conversational Agents*, MIT Press, Cambridge, Mass., 2000.

2. W.L. Johnson, J.W. Rickel, and J.C. Lester, "Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments," *Int'l J. Artificial Intelligence in Education,* vol. 11, 2000, pp. 47–78.

3. J. Rickel and W.L. Johnson, "Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor Control," *Applied Artificial Intelligence*, vol. 13, nos. 4–5, June–Aug. 1999, pp. 343–382.

4. J. Rickel and W.L. Johnson, "Virtual Humans for Team Training in Virtual Reality," *Proc. 9th Int'l Conf. Artificial Intelligence in Education*, IOS Press, Amsterdam, 1999, pp. 578–585.

The Authors

Jeff Rickel is a project leader at the University of Southern California's Information Sciences Institute and a research assistant professor in the Department of Computer Science. His research interests include intelligent agents for education and training, especially animated agents that collaborate with people in virtual reality. He received his PhD in computer science from the University of Texas at Austin. He is a member of the ACM and the AAAI. Contact him at the USC Information Sciences Inst., 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292; rickel@isi.edu; www.isi.edu/~rickel.

Stacy Marsella is a project leader at the University of Southern California's Information Sciences Institute. His research interests include computational models of emotion, nonverbal behavior, social interaction, interactive drama, and the use of simulation in education. He received his PhD in computer science from Rutgers University. Contact him at the USC Information Sciences Inst., 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292; marsella@isi.edu; www.isi.edu/~marsella.

Jonathan Gratch heads up the stress and emotion project at the University of Southern California's Institute for Creative Technology and is a research assistant professor in the department of computer science. His research interests include cognitive science, emotion, planning, and the use of simulation in training. He received his PhD in computer science from the University of Illinois. Contact him at the USC Inst. for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA 90292; gratch@ict.usc.edu; www.ict.usc.edu/~gratch.

Randall Hill is the deputy director of technology at the University of Southern California's Institute for Creative Technologies. His research interests include modeling perceptual attention and cognitive mapping in virtual humans. He received his PhD in computer science from the University of Southern California. He is a member of the AAAI. Contact him at the USC Inst. for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA 90292; hill@ict.usc.edu.

David Traum is a research scientist at the University of Southern California's Institute for Creative Technologies and a research assistant professor of computer science. His research interests include collaboration and dialogue communication between human and artificial agents. He received his PhD in computer science from the University of Rochester. Contact him at the USC Inst. for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA 90292; traum@ict.usc.edu; www.ict.usc.edu/~traum.

William Swartout is the director of technology at the University of Southern California's Institute for Creative Technologies and a research associate professor of computer science at USC. His research interests include virtual humans, immersive virtual reality, knowledge-based systems, knowledge representation, knowledge acquisition, and natural language generation. He received his PhD in computer science from the Massachusetts Institute of Technology. He is a fellow of the AAAI and a member of the *IEEE Intelligent Systems* editorial board. Contact him at the USC Inst. for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA 90292; swartout@ict.usc.edu.

5. W. Swartout et al., "Toward the Holodeck: Integrating Graphics, Sound, Character, and Story," *Proc. 5th Int'l Conf. Autonomous Agents*, ACM Press, New York, 2001, pp. 409–416.
6. D.R. Traum, *A Computational Theory of Grounding in Natural Language Conversation*, doctoral dissertation, Dept. of Computer Science, Univ. of Rochester, Rochester, N.Y., 1994.
7. C. Matheson, M. Poesio, and D. Traum, "Modeling Grounding and Discourse Obligations Using Update Rules," *Proc. 1st Conf. North Am. Chapter of the Assoc. for Computational Linguistics*, Morgan Kaufmann, San Francisco, 2000, pp.1–8.
8. D. Traum and J. Rickel, "Embodied Agents for Multiparty Dialogue in Immersive Virtual Worlds," to be published in *Proc. 1st Int'l Joint Conf. Autonomous Agents and Multi-agent Systems*, ACM Press, New York, 2002.
9. M. Fleischman and E. Hovy, "Emotional Variation in Speech-Based Natural Language Generation," to be published in *Proc. 2nd Int'l Natural Language Generation Conf.*, 2002.
10. J. Gratch, "Émile: Marshalling Passions in Training and Education," *Proc. 4th Int'l Conf. Autonomous Agents*, ACM Press, New York, 2000, pp. 325–332.
11. S.C. Marsella, W.L. Johnson, and C. LaBore, "Interactive Pedagogical Drama," *Proc. 4th Int'l Conf. Autonomous Agents,* ACM Press, New York, 2000, pp. 301–308.
12. J. Gratch and S. Marsella, "Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents," *Proc. 5th Int'l Conf. Autonomous Agents*, ACM Press, New York, 2001, pp. 278–285.
13. S. Marsella and J. Gratch, "A Step Toward Irrationality: Using Emotion to Change Belief," to be published in *Proc. 1st Int'l Joint Conf. Autonomous Agents and Multiagent Systems*, ACM Press, New York, 2002.
14. R. Hill, "Modeling Perceptual Attention in Virtual Humans," *Proc. 8th Conf. Computer Generated Forces and Behavioral Representation*, SISO, Orlando, Fla., 1999, pp. 563–573.
15. R. Hill, "Perceptual Attention in Virtual Humans: Toward Realistic and Believable Gaze Behaviors," *Proc. AAAI Fall Symp. Simulating Human Agents*, AAAI Press, Menlo Park, Calif., 2000, pp. 46–52.
16. S. Chopra-Khullar and N.I. Badler, "Where to Look? Automating Attending Behaviors of Virtual Human Characters," *Proc. 3rd Int'l Conf. Autonomous Agents*, ACM Press, New York, 1999, pp. 16–23.
17. W.K. Yeap and M.E. Jefferies, "Computing a Representation of the Local Environment," *Artificial Intelligence*, vol. 107, no. 2, Feb. 1999, pp. 265–301.
18. R. Hill, C. Han, and M. van Lent, "Applying Perceptually Driven Cognitive Mapping to Virtual Urban Environments," to be published in *Proc. 14th Annual Conf. Innovative Applications of Artificial Intelligence,* AAAI Press, Menlo Park, Calif., 2002.