Intelligent Virtual Agents for Education and Training: Opportunities and Challenges

Jeff Rickel

USC Information Sciences Institute
4676 Admiralty Way, Suite 1001
Marina del Rey, CA, 90292
rickel@isi.edu
http://www.isi.edu/isd/rickel

Abstract. Interactive virtual worlds provide a powerful medium for experiential learning. Intelligent virtual agents can cohabit virtual worlds with people and facilitate such learning as guides, mentors, and teammates. This paper reviews the main pedagogical advantages of animated agents in virtual worlds, discusses two key research challenges, and outlines an ambitious new project addressing those challenges.

1 Introduction

Interactive virtual worlds provide a powerful medium for experiential learning. Navy personnel can become familiar with the layout and operation of a ship to which they will be assigned before they ever set foot on it. History students can learn about ancient Greece by walking its streets, visiting its buildings, and interacting with its people. Biology students can learn about anatomy and physiology through adventures inside the human body. The range of worlds that people can explore and experience is unlimited, ranging from factual to fantasy, set in the past, present, or future.

Our goal is to enrich such worlds with intelligent virtual agents – autonomous, animated agents that support face-to-face interaction with people in these environments in a variety of roles. Existing virtual worlds such as military simulations and computer games often incorporate virtual agents with varying degrees of intelligence. However, the ability of these characters to interact with human users is usually very limited; most 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 typically very scripted, and offer human users no ability to carry on a dialogue. In contrast, we envision virtual agents that cohabit virtual worlds with people and support face-to-face dialogues situated in those worlds, serving as guides, mentors, and teammates. We call this new generation of computer characters *animated pedagogical agents* [1].

A. de Antonio, R. Aylett, and D. Ballin (Eds.): IVA 2001, LNAI 2190, pp. 15–22, 2001.

2 Roles for Animated Pedagogical Agents

Research on animated pedagogical agents can draw on a long history of work in computer-aided learning, especially work on intelligent tutoring systems [2]. However, as discussed in detail by Johnson, Rickel, and Lester [1], animated pedagogical agents offer a variety of distinct new capabilities:

Interactive Demonstrations. A simulated mockup of a student's real work environment, coupled with an animated agent that inhabits the virtual world, provides new opportunities for teaching the student how to perform tasks in that environment. Perhaps the most compelling advantage is that the agent can demonstrate physical tasks, such as operation and repair of equipment. For example, Rickel and Johnson's Steve agent [3,4] cohabits a 3D mockup of a US Navy ship with students, and can demonstrate procedures while providing spoken commentary describing his objectives and actions. Steve can respond to interruptions from students and use planning to adapt the demonstration to the state of the virtual world, thereby providing more interactive demonstrations than alternatives such as video.

Navigational Guidance. When a student's work environment is large and complex, such as a ship, one of the primary advantages of a virtual mockup is to teach the student where things are and how to get around. In this context, animated agents are valuable as navigational guides, leading students around and preventing them from becoming lost. For example, as Steve demonstrates tasks, he uses a path planning algorithm to lead students around a complicated shipboard environment. Similarly, the WhizLow agent [5] uses path planning to guide students through the internals of a virtual computer.

Gaze and Gesture as Attentional Guides. To draw students' attention to a specific aspect of a chart, graphic or animation, tutoring systems make use of many devices, such as arrows and highlighting by color. An animated agent, however, can guide a student's attention with the most common and natural methods: gaze and deictic (e.g., pointing) gestures. For example, Steve uses gaze and pointing gestures to draw a student's attention to objects in the virtual world as well as to people and agents who are responsible for the next step in a task. Among other animated agents, Presenter Jack [6] is notable for its broad repertoire of deictic gestures, and Cosmo [7] is notable for its sophisticated criteria for choosing deictic gestures and accompanying referring expressions.

Nonverbal Feedback. One primary role of a tutor is to provide feedback on a student's actions. In addition to providing verbal feedback, an animated agent can also use nonverbal communication to influence the student, which allows more varied degrees of feedback than earlier tutoring systems. For example, the Adele agent [8,9] nods or smiles to indicate agreement with a student's actions, presents a look of puzzlement when the student makes an error, and shows pleasant surprise when the student finishes a task, while Herman the Bug [10] sometimes congratulates students by cartwheeling across the screen.

Conversational Signals. When people carry on face-to-face dialogues, they employ a wide variety of nonverbal signals to regulate the conversation and complement their verbal utterances. While tutorial dialogue in most previous tutoring systems resembles Internet chat or a phone conversation, animated pedagogical agents allow us to more closely model the face-to-face interactions to which people are most accustomed. For example, the agents of Cassell and her colleagues [11,12,13] use gaze and gestures to regulate turn taking, head nods to provide backchannel feedback, and gaze, eyebrow movements, head movements, and gestures to emphasize words.

Emotions and Personality. Motivation is a key ingredient in learning, and emotions play an important role in motivation. By modeling and expressing emotions, animated agents may improve student motivation by conveying enthusiasm for subject matter and appearing to care about the student's progress, and an animated agent with a rich and interesting personality may simply make learning more fun [14]. For example, Cosmo [15] employs a repertoire of emotive behaviors to advise, encourage, and (appear to) empathize with students.

Virtual Teammates. When students must learn to perform effectively in a team, they must master their individual role while also learning to coordinate their actions with their teammates. In such team training, animated agents can play two valuable roles: they can serve as instructors for individual students, and they can substitute for missing team members, allowing students to practice team tasks when some or all human instructors and teammates are unavailable. Steve supports such team training; a team can consist of any combination of Steve agents and human students, each assigned a particular role in the team [16]. Different Steve agents can be configured with different appearances and voices, and students can monitor the activities of their agent teammates as well as communicate with them through spoken dialogue. This capability for multiple animated agents to cohabit a virtual world with multiple people provides a rich environment for team training.

Story and Character. Engaging stories with interesting characters have a powerful ability to capture and hold our attention and leave a memorable impression. If such stories could be harnessed for education and training, and made interactive to allow active learning by students, the result could provide a potent tool for learning. For example, the Interactive Pedagogical Drama system [17] combines a rich story line and emotionally expressive animated agents to allow mothers of children with cancer to interactively control a character in a similar situation. Through a therapist character in the story, who reacts dynamically to the human mother's choices, the mother gains exposure to a problem-solving methodology developed by clinical psychologists to help such people cope with common problems.

3 Towards Broader Support for Experiential Learning

Despite exciting progress on animated pedagogical agents over the last five years, two main challenges remain. First, while such agents incorporate sophisticated capabilities from other areas of artificial intelligence, such as spoken dialogue and models of emotion, they do not represent the state of the art in those areas. More work is required to fully integrate the latest advances into animated agents. Second, each animated pedagogical agent developed to date incorporates only some of the capabilities described in the last section. We do not yet have a single agent architecture that integrates all of these capabilities, addressing all of their interdependencies.

We are currently addressing these two challenges in an ambitious new project [18,19]. In contrast to our prior work on Steve, which focused on teaching well-defined tasks, our new project focuses on leadership training and decision making in stressful situations. Figure 1 shows a screen shot from a prototype implementation of an example application, in which a young lieutenant (human user) is being trained for a peacekeeping mission. In the current implementation, there are three Steve agents that interact with the lieutenant: a medic (front right) that serves as his teammate, a Bosnian mother (front center) whose boy has been accidentally injured by one of the lieutenant's vehicles, and a sergeant (front left) who serves as both a teammate and mentor. All other characters (soldiers and an angry crowd of locals) are currently simple scripted agents. The bodies and animation algorithms for all the Steve agents and scripted characters were developed by Boston Dynamics Incorporated, and the expressive faces for the sergeant and medic were developed by Haptek Incorporated.

Fig. 1. An interactive peacekeeping scenario featuring (left to right) a sergeant, a mother, and a medic

This new type of training exploits all of Steve's prior capabilities, but it is pushing us in several new directions. First, we are addressing issues of story and character to make the learning experience more engaging and memorable. The basic story line was created by a Hollywood script writer, in consultation with Army training experts. As the simulation begins, a human user, playing the role of a U.S. Army lieutenant, finds himself in the passenger seat of a simulated vehicle speeding towards a Bosnian village to help a platoon in trouble. Suddenly, he rounds a corner to find that one of his platoon's vehicles has crashed into a civilian vehicle, injuring a local boy (Figure 1). The boy's mother and an Army medic are hunched over him, and a sergeant approaches the lieutenant to brief him on the situation. Urgent radio calls from the platoon downtown, as well as occasional explosions and weapons fire from that direction, suggest that the lieutenant send his troops to help them. Emotional pleas from the boy's mother, as well as a grim assessment by the medic that the boy needs a medevac immediately, suggest that the lieutenant instead use his troops to secure a landing zone for the medevac helicopter. The lieutenant carries on a dialogue with the sergeant and medic to assess the situation, issue orders (which are carried out by the sergeant through four squads of soldiers), and ask for suggestions. His decisions influence the way the situation unfolds, culminating in a glowing news story praising his actions or a scathing news story exposing the flaws in his decisions and describing their sad consequences.

Second, since the human user must collaborate with his agent teammates to formulate novel plans, rather than simply execute well-defined procedures, we are integrating state-of-the-art natural language understanding and generation algorithms into Steve, as well as extending those algorithms to handle multi-party conversations in immersive virtual worlds [20]. On one hand, virtual worlds are an ideal application for current spoken language technology: they provide a microworld where conversation can legitimately be restricted to the events and objects within its confines. On the other hand, they raise issues that have received relatively little attention in computational linguistics. First, face-to-face communication in virtual worlds requires attention to all the nonverbal signals (e.g., gaze, gestures, and facial displays) that accompany human speech. Second, conversations that are situated in a 3D world raise a host of issues, including the attentional focus of the conversants, whether and to what degree they can see and hear one another, and the relative locations of conversants and the objects they are discussing. Finally, since there will typically be multiple real and virtual people, virtual worlds require support for multi-party conversations, including the ability to reason about the active participants in a conversation as well as who else might be listening. While there has been some early work in the area of embodied conversational agents [21], and some of this work has addressed human-agent dialogues situated in 3D virtual worlds [3], there is currently no general model of such dialogues.

Third, to model the behavior of teammates in stressful situations, as well as create virtual humans that can induce stress in the human user by reacting emotionally, we have integrated a computational model of emotions into Steve [22]. For example, the mother in our peacekeeping scenario becomes increasingly angry at the lieutenant if his decisions thwart her goal of getting assistance for her boy. To handle the full range of emotions that arise in our scenario, we are incorporating a model of emotions arising from goals and plans [23] as well as a model of emotions arising from interpersonal relationships [17], and we are also working towards an outward expression of emotions that will be both believable and interpretable [24].

Our goal in this project is to create an architecture for animated pedagogical agents that integrates a broad range of capabilities. It is not sufficient to simply plug in a number of different modules representing the state of the art in spoken dialogue, affective reasoning, models of teamwork, motor control, etc. The state of the art in each area was developed independent of the others, so the fundamental research challenge is to understand the interdependencies among them. Our goal is an animated agent architecture that can be reused in a wide variety of applications and easily configured to model different styles and personalities.

4 Conclusion

Interactive virtual worlds combined with animated pedagogical agents offer an exciting new tool for education and training. While this tool builds on a long line of work in simulation-based learning and intelligent tutoring systems, animated pedagogical agents provide a variety of new capabilities. The last five years have seen exciting progress in this new technology, but most existing agents still have a limited range and depth of capabilities. To address this challenge, we are currently working on a single agent architecture that will incorporate a broader range of capabilities as well as integrating state-of-the-art modules for core capabilities such as spoken dialogue and affective reasoning. If successful, our work will lead to more sophisticated applications for animated pedagogical agents as well as a better understanding of the interdependencies among these different agent capabilities.

Acknowledgments. My research on animated pedagogical agents would not have been possible without the help of many talented people. Lewis Johnson collaborated with me on the original design of Steve's brain, and Marcus Thiebaux developed the models and animation algorithms for Steve's original body. Randy Stiles and his team at Lockheed Martin created the virtual world for Steve's shipboard training, and Allen Munro and his team at the USC Behavioral Technology Laboratories developed the simulation models for that virtual world. In our new project, Bill Swartout serves as the overall project director, Shri Narayanan is working on speech recognition, Randy Hill is working on perception, Ed Hovy is working on natural language understanding and generation, David Traum is working on dialogue management, Jon Gratch and Stacy Marsella are working on affective reasoning, Lewis Johnson 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, Jay Douglas is working on interactive narrative issues, Ben Moore and Marcus Thiebaux created the simulation infrastructure, Chris Kyriakakis and Dave Miraglia are developing the immersive sound effects, and most of the graphics for the Bosnian town were created by Jacki Morie, Erika Sass, and Michael Murguia. The original research on Steve was funded by the Office of Naval Research under grant N00014-95-C-0179 and AASERT grant N00014-97-1-0598, and the new work is funded by the Army Research Office through the USC Institute for Creative Technologies under contract DAAD19-99-C-0046.

References

1. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial Intelligence in Education **11** (2000) 47–78
2. Wenger, E.: Artificial Intelligence and Tutoring Systems. Morgan Kaufmann, Los Altos, CA (1987)
3. Rickel, J., Johnson, W.L.: Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. Applied Artificial Intelligence **13** (1999) 343–382
4. Rickel, J., Johnson, W.L.: Task-oriented collaboration with embodied agents in virtual worlds. In Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: Embodied Conversational Agents. MIT Press, Cambridge, MA (2000)
5. Lester, J.C., Zettlemoyer, L.S., Gregoire, J., Bares, W.H.: Explanatory lifelike avatars: Performing user-designed tasks in 3d learning environments. In: Proceedings of the Third International Conference on Autonomous Agents, New York, ACM Press (1999)
6. Noma, T., Zhao, L., Badler, N.I.: Design of a virtual human presenter. IEEE Computer Graphics and Applications **20** (2000) 79–85
7. Lester, J.C., Voerman, J.L., Towns, S.G., Callaway, C.B.: Deictic believability: Coordinating gesture, locomotion, and speech in lifelike pedagogical agents. Applied Artificial Intelligence **13** (1999) 383–414
8. Shaw, E., Johnson, W.L., Ganeshan, R.: Pedagogical agents on the web. In: Proceedings of the Third International Conference on Autonomous Agents, New York, ACM Press (1999) 283–290
9. Shaw, E., Ganeshan, R., Johnson, W.L., Millar, D.: Building a case for agent-assisted learning as a catalyst for curriculum reform in medical education. In: Proceedings of the Ninth International Conference on Artificial Intelligence in Education, IOS Press (1999)
10. Lester, J.C., Stone, B.A., Stelling, G.D.: Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments. User Modeling and User-Adapted Interaction **9** (1999) 1–44
11. Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., Stone, M.: Animated conversation: Rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents. In: Proceedings of ACM SIGGRAPH '94, Reading, MA, Addison-Wesley (1994) 413–420
12. Cassell, J., Thórisson, K.R.: The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence **13** (1999) 519–538
13. Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., Yan, H.: Conversation as a system framework: Designing embodied conversational agents. In Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: Embodied Conversational Agents. MIT Press, Cambridge, MA (2000)
14. Elliott, C., Rickel, J., Lester, J.: Lifelike pedagogical agents and affective computing: An exploratory synthesis. In Wooldridge, M., Veloso, M., eds.: Artificial Intelligence Today. Volume 1600 of Lecture Notes in Computer Science. Springer-Verlag, Berlin (1999) 195–212

15. Lester, J.C., Towns, S.G., Callaway, C.B., Voerman, J.L., FitzGerald, P.J.: Deictic and emotive communication in animation pedagogical agents. In Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: Embodied Conversational Agents. MIT Press, Cambridge, MA (2000)
16. Rickel, J., Johnson, W.L.: Virtual humans for team training in virtual reality. In: Proceedings of the Ninth International Conference on Artificial Intelligence in Education, IOS Press (1999) 578–585
17. Marsella, S.C., Johnson, W.L., LaBore, C.: Interactive pedagogical drama. In: Proceedings of the Fourth International Conference on Autonomous Agents, New York, ACM Press (2000) 301–308
18. Rickel, J., Gratch, J., Hill, R., Marsella, S., Swartout, W.: Steve goes to Bosnia: Towards a new generation of virtual humans for interactive experiences. In: AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment. (2001)
19. Swartout, W., Hill, R., Gratch, J., Johnson, W., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiébaux, M., Tuch, L., Whitney, R., Douglas, J.: Toward the holodeck: Integrating graphics, sound, character and story. In: Proceedings of the Fifth International Conference on Autonomous Agents, New York, ACM Press (2001) 409–416
20. Traum, D., Rickel, J.: Embodied agents for multi-party dialogue in immersive virtual worlds. In: Agents 2001 Workshop on Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents, Montreal, Canada (2001) 27–34
21. Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: Embodied Conversational Agents. MIT Press, Cambridge, MA (2000)
22. Gratch, J., Marsella, S.: Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In: Proceedings of the Fifth International Conference on Autonomous Agents, New York, ACM Press (2001) 278–285
23. Gratch, J.: Émile: Marshalling passions in training and education. In: Proceedings of the Fourth International Conference on Autonomous Agents, New York, ACM Press (2000) 325–332
24. Marsella, S., Gratch, J., Rickel, J.: The effect of affect: Modeling the impact of emotional state on the behavior of interactive virtual humans. In: Agents 2001 Workshop on Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents, Montreal, Canada (2001) 47–52