It's Virtually Pedagogical: Pedagogical Agents in Mixed Reality Learning Environments

Jayfus Tucker Doswell
George Mason University
doswellj@hotmail.com

Abstract
Pedagogical Embodied Conversational Agents (PECA) that autonomously behave in mixed reality environments, respond to muti-modal input across computer networks, interact with human learners using context aware intelligence, and apply proven pedagogical techniques during instruction have the potential to improve and accelerate human learning performance anytime, anywhere, and at any-pace. This paper discusses the PECA Product Line Architecture (PPLA) model for building interactive pedagogical agent systems and discusses a prototype system from the architecture.

CR Categories: D.2.11 *Domain-specific architectures; H.5.2 Graphical user interfaces (GUI; H.5.1 Artificial, augmented, and virtual realities*

Keywords: virtual instructor, pedagogical agent, embodied conversational agent, pedagogical embodied conversational agent, virtual reality, software architecture

1. Introduction

The Pedagogical Embodied Conversational Agent (PECA) is defined as a pedagogical agent or virtual instructor that provides a personalized human learning experience by applying empirically evaluated and tested instructional techniques for improving human learning [Doswell 2004]. These instructional techniques, combining the art and science of teaching (i.e., pedagogy), are exemplified by three dimensional (3D) animated characters that intelligently consider multiple variables for improving and potentially accelerating the human learning process. These variables include, but are not limited to learning styles, human emotion, culture, gender, pedagogical techniques, and andragogical techniques (i.e., for adult learners). Additionally, these 3D-animated characters are designed to behave autonomously in mixed reality (e.g., virtual reality, augmented reality, augmented virtuality, etc.) environments, respond to human verbal/non-verbal input across distributed and wireless computer networks, and naturally interact with human learners using context-aware intelligence across varying cultures. By combining state of the art technologies and instructional/learning techniques, the ultimate goal for PECAs is to improve and accelerate human learning performance anytime, anywhere, and at any pace [Doswell 2005]. However, in order for a PECA, virtual instructor, or pedagogical agent to improve and accelerate human learning performance, they must combine the strengths of "master" instructors that possess expertise in specific academic/knowledge domains. Additionally, they must exceed human pedagogical capabilities required to effectively guide learners through complex concepts/task and clarify misunderstanding, while at the same time, become intimately involved understanding the learner and knowledge being learned. These are computational challenges for all virtual instructor developers and require interdisciplinary expertise in areas such as cognitive science, sociology, computer software engineering, computational humanities, educational technology, artificial intelligence, 3D computer graphics, linguistics, and interactive display technologies. A system/software engineering approach to solving these interdisciplinary challenges for designing PECAs with capabilities to provide instruction in mixed reality environments and potential to match or exceed instructional capabilities of human instructors is the focus of this paper.

2. Background

The building blocks of the 3D computer animated PECAs are virtual humans that have started to penetrate our daily operations, starting by inhabiting our auditory world [Plantec 2004]. For example, if people call British Airways, they can have a satisfactory conversation with their virtual reservation agent. Through a combination of state of the art vocabulary, over the phone speech recognition, and natural language processing, one can talk with a pleasantly mannered virtual human about anything within the domain of a British Airway reservation. On the web, virtual taking heads are starting to emerge with definitive personalities incorporating face animation, and avatars representing virtual reality participants. The international noted author and "futurist", Ray Kurzweil stated that by year 2010, virtual humans will have the ability to pass the Turning Test. In his prediction, people will not mistake virtual humans for real ones, but will interact naturally with them as information assistants, virtual coaches, virtual sales clerks, entertainers, and even for love replacement therapy [Plantec 2004]. Perhaps, one of the most important applications of virtual human technology will be in the teaching domain. Because virtual teachers may be realistically designed to operate separately from school system politics and policies, it would be difficult to attribute alternative motives to an animated character. Additionally, a well designed virtual instructor will work just as efficiently with no pay and teach with its main to improve the knowledge of the human learner with options to teach to the traditional curriculum or extend it. Currently, education is going through an evolutionary stage

[Plantec 2004] and more people are realizing that specialized knowledge acquisition rather than traditional grade level elevation contributes greatly to human "success". With an innovative model of virtual instructor directed learning, real-humans may be effectively taught at their own pace and receive personalized learning by their own personal PECA and personalized instruction based on empirically tested pedagogical techniques [Doswell 2005].

The simulated environments PECAs inhabit and from which they provide instruction also deserves discussion. Typically, virtual instructors and pedagogical agents have occupied virtual reality environments [Cassell 2000; Norma 2000; Rickel 2000; Lester 2000]. However, to support various environmental contexts and the mobile learner, augmented reality, volumetric displays, hand-held displays, and holographic environments should be considered and evaluated during research. Therefore, an additional challenge is raised in how to design and develop PECAs or virtual instructors in mixed reality environments [Doswell 2005]. The term *mixed reality environments* is used to categorize various types of virtual reality (VR) type display systems [Milgram 1994]. Hence, researchers have defined a continuum of real-to-virtual environments, in which VR and AR are parts of the general area of what is now considered, "mixed reality" [Milgram, 1994]. Figure 1 illustrates this mixed reality continuum. In augmented reality, digital objects are added to the real environment. In *augmented virtuality,* real objects are added to virtual ones. In *virtual environments* (or virtual reality), the surrounding environment is virtual.

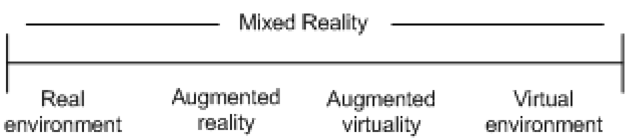

Figure 1: Milgram's reality-virtuality continuum.

The advantage of designing PECAs to operate in mixed reality environments is that they may be programmed to intelligently adapt and use the best simulated environment for igniting learner motivation, reinforcing concepts, guiding learners through new and complex task, clarifying ambiguity, and providing a more enhanced and natural e-learning environment. For example, if a learner wanted to learn about a historical building, during an outdoor student tour, the PECA may be loaded in an augmented reality environment and supplement the tour guide by tailoring the experience to the learners personal interest and learning strengths. To learn more about the tour, the student would then be able to interact with the PECA in VR.

3. PECA Product Line Architecture

To build PECAs that operate in mixed reality environments and effectively provide instruction to human learners requires an extensible, interoperable, modular, and scalable software architecture model. Comprehensive research has been conducted to design such an architecture model, the PECA Product Line Architecture (PPLA). The PPLA is an improvement over previous research on the Joint Embedded Pedagogical Agent Architecture (JEPAA), and it is designed to generate families of PECAs with capabilities to provide instruction from mixed reality environments and adapt to evolving virtual environment technologies [Doswell 2004]. Like JEPAA, the PPLA was developed to address limitations in dynamic 3D animation, knowledge domain dependence, autonomous operation within mixed reality environments, realistic interaction with human end-users, and pedagogy intelligence. The PPLA supports the "plug-in-play" integration of various system/software components [Gomaa 2005] ranging from speech recognition and speech synthesis components to 3D animation algorithms, pedagogical models, and learning management systems.

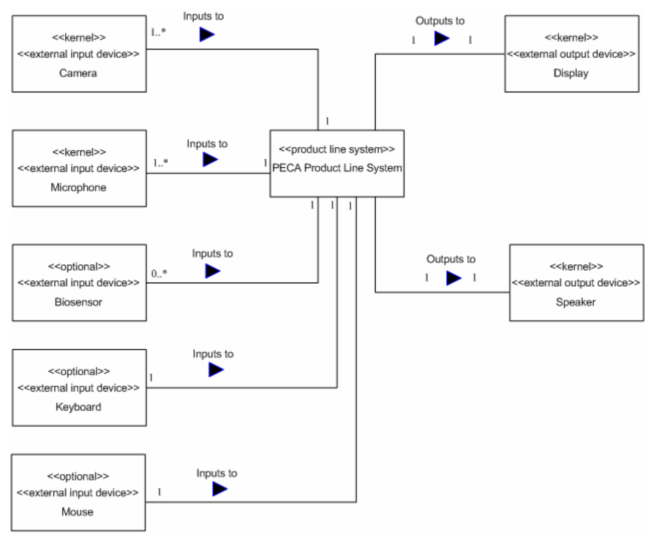

Figure 2: PPLA Context Diagram

The PECA Product Line supports kernel (i.e., required) multimodal external inputs illustrated in Figure 2. Multimodal inputs supported in all PECA systems include human gestures (i.e., face and body); objects in the human environment (e.g., toys) captured by digital camera devices; human verbal input and environment noise captured by microphone devices.

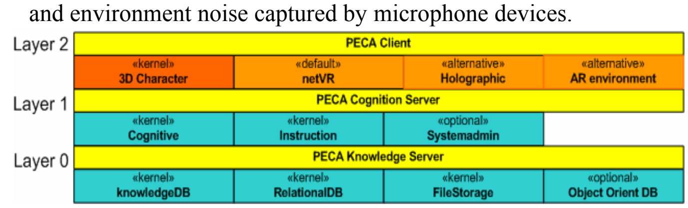

Figure 3: PPLA Layered Architecture View

Optionally, a PECA system created from the PPLA may support the capture and interpretation of human physiological data captured from biosensors to help measure and evaluate the learning progress of the human learner.

Figure 3 illustrates the multi-tier view of the PECA Product Line Architecture. This view is based on the *Layers of*

Abstraction pattern (also known as the Hierarchical Layers or Levels of Abstraction pattern) [Gomaa, 2005]. In this view, and PECA Knowledge Server, PECA Cognition Server, and PECA client tiers are depicted as Layer 0, Layer 1, Layer 2, respectively. Subsystems and corresponding components each reside in these individual layers. In the PECA Client tier, the interactive 3D Character model (i.e., PECA) is the *kernel* (i.e., required) graphical component with which the end user (i.e., learner) interacts. The netVR, Holographic, and AR environment are *alternative* mixed reality environments in which a PECA may be integrated. These environments are denoted as *alternate* because one and only one mixed reality environment may be selected in which PECA will inhabit for each PPLA system implementation. The *default* environment is the netVR. In the PECA Cognition Server, several subsystems exist including, *Cognitive, Instruction, and Systemadmin*. All are kernel subsystem except for the Systemadmin subsystem. The PECA Cognition Server is the most dynamic aspect of the three layers. The PECA Cognition Server contains all components that provide the PECA with its instructional capabilities, behavior/personality, speech intelligence, and visual capabilities. Additionally, this PPLA component is more likely to be updated (besides the data). Consequently, this tier is subjected to expansion faster than the other tiers. The PECA Knowledge Server has several data repositories with which the PECA Cognition Server interfaces. Data repositories including knowledge databases/expert systems, relational database systems, and file storage systems are *kernel* data repositories supported by the PPLA. Optional data repositories may be object-oriented databases among other data repositories or data management systems.

4. PECA System in Virtual Reality

Figure 4: PPLA Prototype

The PPLA was used to develop and prototype a 3D animated PECA and all of its corresponding software components. The 3D character was designed as a "cute" and non-authoritative figure in order to sustain a comfortable learning medium for children. The resulting PECA demonstrates the ability to assess a learner's strength and learning style in a particular academic domain and instruct using a pedagogical technique, scaffolding. The prototype demonstrates a PECA's ability to naturally communicate with human end-users utilizing speech recognition, speech synthesis, and natural language recognition components. Additionally, the prototype demonstrates the ability for the PECA to recognize human gestures. The PECA prototype was deigned with the knowledge to provide basic numeric concepts and basic newtonian physics instruction to children ages 8-10. To prepare the PECA model for operation within a netVE, a 3D animated character was developed in 3D Studio Max with a controllable skeleton and bone structure. Subsystems within the PECA Product Line read the 3D file to analyze and store the PECA's body and face structure in order to prepare it for dynamic animation within a virtual reality environment. For human input interpretation, the PECA Product Line, and thus, the resulting prototype used a combination of Java, C++, share-ware, and Commercial-Off-The Shelf (COTS) software. Rules for storing conversational and instructional knowledge were accomplished using the Artificial Intelligence Markup Language (AIML); an XML type of rule based repository of knowledge. Microsoft SQL Server was used to store dynamic learning tracking data and a learning profile for each learner. In order for the PECA to see learners and other objects in their environment, Intel's OpenCV image algorithms were integrated for specifically recognizing human face gestures, body gestures (e.g., hand gestures), and the presence or the learner. For recognizing and responding to human speech, the Java Speech Application Programming Interface (JSAPI) was used to develop speech synthesis and speech recognition components on the PECA Cognition tier. Custom software components were developed in Java to enable the PECA to apply pedagogy techniques during its instruction. Pedagogical rules were stored in both in AIML and the relational database. To provide a more pervasive interface when verbally communicating, a wireless Bluetooth enabled headset was configured to enable the learner to roam about while communicating with the PECA and listening to its instructional responses.

5. Conclusion and Future Direction

The PPLA and resulting prototype demonstrates the complexity of building virtual instructors, pedagogical agents, and PECAs that provide instruction to humans from mixed reality environments. The implication of autonomous interactive 3D-character instructors that provide instruction based on proven pedagogical models and knowledge of individual learning needs is paramount to increasing human learning performance. This type of interactive learning medium demonstrates a fun and effective learning environment for persons of all ages and especially addresses K-12 learning needs. Continued research is planned to extend the PPLA with capabilities for it to provide instruction from Augmented Reality, Holographic environments, and volumetric displays.

References

CASSELL, J. ET. AL. 2000. Human Conversation as a System Framework: Designing Embodied Conversational Agents. *Embodied Conversational Agents*, MIT Press, 29-61.

DOSWELL, J. 2005. Pervasive Pedagogical Storyteller: Interactive History in Mixed Reality Environments. *In Proceedings of 2nd International Digimedia conference 2005. theme: Multimedia uses in education and entertainment for children*, Cairo, Egypt.

DOSWELL, J. 2004. Pedagogical Embodied Conversational Agents. *In Proceedings of IEEE International Conference on Advanced Learning Technologies 2005*, Joensuu, Finland, 774-776.

GOMAA, H. 2005. *Designing Software Product Lines with UML: From Use Case to Pattern Based Software Architectures*. Addison-Wesley.

LESTER, J. et. al. 2000. Deictic and Emotive Communication and Animated Pedagogical Agents. *Embodied Conversational Agents*, MIT Press, 123-154.

MILGRAM, P. and F. KISHINO. 1994. A Taxonomy of Mixed Reality Visual Displays. In Proceedings of *IEICE Transactions on Information Systems* E77-D, 12, 1321-1329.

NORMA, T. 2000. Design of a Virtual Human Presenter. *IEEE Computer Graphics and Applications*, 79-83.

PLANTEC, P. 2005. *Virtual Humans: A Build-It-Yourself Kit*. AMACON books.

RICKEL et. al. 2000. Task-Oriented Collaboration With Embodied Agents in Virtual Worlds. *Embodied Conversational Agents*, MIT Press, 96-122.