

Interactive Characters for Cultural Training of Small Military Units

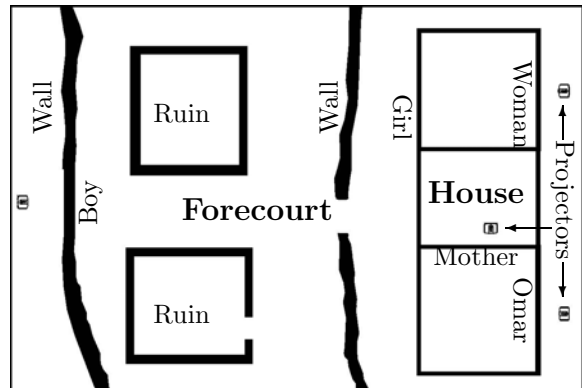
Priti Aggarwal, Kevin Feeley, Fabrizio Morbini, Ron Artstein, Anton Leuski, David Traum, and Julia Kim

Institute for Creative Technologies, University of Southern California,
12015 Waterfront Drive, Playa Vista CA 90094-2536, USA

CHAOS, the Combat Hunter Action and Observation Simulation, is an immersive simulation training environment which gives small military units the experience of interacting with local Afghan villagers during a patrol. It is a physical build-out of a housing compound in a mock Afghan village, with several life-size reactive and interactive animated Pashto-speaking virtual characters. The exercise requires an infantry squad to locate and interview a character named Omar, communicating through a live human interpreter and attending to proper protocol regarding Omar's family. Character animation and behavior is based on extensive interviews with Afghan experts to provide a realistic setting of the intended locale. The system combines virtual human technology, story engineering, and physical set building to provide a compelling training environment that can handle a full squad, requiring trainees to integrate tasks such as working with an interpreter, dealing with non-English speakers from another culture, and assessing information and disposition to make decisions in a mission context.

The simulation is part of the Future Immersive Training Environment Joint Capabilities Technology Demonstration (FITE JCTD), located in a reproduction of an Afghan village at Camp Pendleton, California.

It is an enclosed area consisting of a forecourt and a house with two rooms. As a unit enters the compound it encounters Farhan, a boy playing in the forecourt, projected life-size on a screen at the edge of the compound. Farhan receives signals from a radio frequency locator system which tracks the position of trainees



throughout the FITE installation; as soon as Farhan senses the squad's presence he runs away, providing a distraction from the main mission. A flat screen monitor positioned as the window of the women's residence shows a girl named Tasleem, who scuttles inside when the squad approaches the house; if the squad makes the mistake of entering the women's room then Nasira, an adult woman, admonishes them and gestures to them to leave. All three characters are merely reactive, responding only to the presence of trainees in their vicinity.

The main interaction takes place in the men's residence in the house. Two characters are projected at life size on screens on adjacent walls: Omar, the head of the household, and Asala, Omar's mother. The characters only speak and understand Pashto, a local language of Afghanistan and Pakistan. The trainees interact with the characters through a live human interpreter, who translates the English questions into Pashto and then translates the characters' responses back into English. The interpreter's speech is transformed into Pashto text with customized acoustic, language, and dictionary models using the OtoSense speech recognition engine. The characters are driven by NPCEditor [2], an engine that selects appropriate character responses to Natural Language input based on a statistical learned mapping between input and output utterances. NPCEditor is trained using English and Pashto text in both the questions and responses. The characters can respond automatically to the input speech, or run in semi-automatic mode, where an operator can override the system's selected responses.

To allow a coherent interaction in the face of noisy speech recognition, the characters follow a structured story-driven interaction similar to the Gunslinger architecture [1]. The interaction consists of four "beats", intended to progressively raise the stress level of the trainees and get them out of their comfort zone. Advancement from one beat to the next can happen either in response to a question by the squad, or at the character's initiative if the squad fails to move the interview in the desired direction. The first beat is small talk and greetings. Conflict emerges in the second beat, when the squad tries to find information about a generator that Omar controls while he tries to get the squad to help with the generator's maintenance. Tensions rise in the third beat when the mother gets involved, interrupting the conversation and shouting at both the trainees and her son. Resolution is reached in the fourth beat; at this point a signal is sent to operators at mission control, who radio the squad with their next assignment.

We developed and tested the system with nine infantry squads at Camp Pendleton. This demonstrated that a mixed-reality, multiple character environment can successfully engage a small military unit, allowing them to practice tactical questioning and decision-making skills in a safe, consistent, and controlled environment that realistically depicts situations they will encounter in deployment. Innovations in this system include Pashto language interaction, combining virtual characters with a live human interpreter, use of story for scaffolding the interaction, and integration with a locator system to drive multiple virtual characters to engage (and distract) a whole military unit in a single exercise.

References

1. Hartholt, A., Gratch, J., Weiss, L., The Gunslinger Team: At the Virtual Frontier: Introducing Gunslinger, a Multi-Character, Mixed-Reality, Story-Driven Experience. In: Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjálmsón, H.H. (eds.) IVA 2009. LNCS, vol. 5773, pp. 500–501. Springer, Heidelberg (2009)
2. Leuski, A., Traum, D.: Practical language processing for virtual humans. In: Proceedings of the Twenty-Second Innovative Applications of Artificial Intelligence Conference (IAAI 2010), July 2010, pp. 1740–1747. AAAI Press, Atlanta (2010)