Shared Reality: Physical Collaboration with a Virtual Peer

J. Cassell, M. Ananny, A. Basu, T. Bickmore, P. Chong, D. Mellis,
K. Ryokai, J. Smith, H. Vilhjálmsson, H. Yan
Gesture and Narrative Language Group, MIT Media Laboratory
E15-315, 20 Ames St., Cambridge MA 02139
+1 617 253 4899
{justine, ananny, anindita, bickmore, horn, stendahl, kimiko, jsmith, hannes, yanhao}@media.mit.edu

ABSTRACT

We describe a novel interface, in which a human and embodied conversational agent share a seamlessly integrated virtual and physical environment. This type of interface, in which objects are passed from the real to the virtual world, has potential applications in unsupervised learning, collaborative work, and entertainment. We introduce Sam, our first implementation of such an interface, which allows children to engage in natural storytelling play with real objects, in collaboration with a virtual playmate who shares access to those real objects.

Keywords

Shared reality, embodied conversational agent, tangible interface, storytelling, children, peer, collaboration

INTRODUCTION

In telling stories, children actively produce their own understanding of the world around them and of themselves, while practicing valuable language and thinking skills. Children engage in this activity naturally while playing with toys and play sets, especially when there are other children who are willing to collaborate to make a story and playfully introduce new skills. But current technology does not support these important skills in children. And children get more value from the activity when playmates engage in the activity collaboratively, which is not always within the capabilities of real-life younger children. An embodied conversational character has the educational potential and emotional engagement for this goal. However, while animated characters are familiar to children from television, current versions of children's animation encourage children to be passive consumers of narrative, rather than active producers of their own stories. There is a place, then, for a new type of interface that lets children engage in natural storytelling behavior and language play with a collaborative virtual peer, and real physical toys. Sam is an interface that meets these criteria.

SAM

The Sam system has two components: an embodied conversational agent, and a toy castle with a plastic figurine. The agent is projected on a screen behind the castle, and can both tell stories and listen to the user's stories. The physical toy castle is only half of the castle; an image on the screen gives the appearance that the castle continues into the screen. The figurines can exist in either the physical world or on the screen, so that the agent and the user can pass it back and forth between their worlds.

For example, Josh wanders into the play room. He hears "does anyone want to play with me?" and looks over to see a castle play set with a life-sized animated child behind it. The child seems to be playing with a little figurine. Josh goes over to sit in front of the castle. The animated child looks at him, smiles, and says "Hi, I'm Sam!". After Josh introduces himself, Sam says "let's make up a story. I'm going to start." Sam then tells the beginning of a story, moving the figurine around the castle, occasionally looking up at Josh to draw him in to the story. Then Sam says, "I'll put the toy in the magic tower so you can tell a story", and places the figurine inside the back of a tower. When Josh opens the door at the front of the tower, he finds the toy Sam had been playing with. He picks it up and begins to tell a story, and while he does so, Sam watches him and the toy, nodding, smiling, and prompting "what happens next?" when Josh hesitates. When he's done, Josh gives the toy to Sam, and the interaction continues.

This system combines work in embodied conversational agents [2], story listening systems [1], and tangible interfaces [4], and introduces a novel way of seamlessly integrating the tangible and the virtual. The result is a combined virtual/physical interface that uses children's

existing social knowledge and interests to draw them into engaging storytelling play.

The key and novel aspects of this system are (a) the shared reality of the playset, (b) a responsive embodied conversational character, and (c) the model of story-listening that scaffolds children's narrative play.

SHARED REALITY

The virtual playmate shares the same world as the child, and can pass real toys back and forth within that shared reality. We believe that this seamless integration of the virtual and the physical plays an important role in supporting children's everyday storytelling play, and can also have substantial applications in other areas. Since the child and Sam share access to the physical world, the interaction is emotionally engaging, and helps the child to suspend disbelief. When the virtual world extends into the physical world, children can engage their bodies in the interaction while manipulating objects (which is natural in children's play) and discussing them with Sam. This allows the system to support the child's voice more effectively and playfully. This shared interface also encourages collaboration, enhancing Sam's ability to inspire and scaffold children in their narration.

EMBODIED CONVERSATIONAL PLAYMATE

The system's success also depends on the simulated peer, whose behavior is modeled on children's social and conversational skills. When a child approaches the play set, and sees a life-sized animated child sitting behind it, she is primed to behave in a particular way, based on years of experience of playing and talking with other children. No extra learning is required – a child can simply sit down and begin playing and talking in familiar ways. Sam's voice and appearance as a co-equal peer creates an unintimidating play space to facilitate more creative and open-ended play [3]. Finally, in other applications, such a system can effectively demonstrate how to use physical objects since it has a body like the user's.

STORY LISTENING

Our earlier research has shown that in playing with a story-listening interface children tell stories that are more imaginative and more narratively complex than when they play by themselves, and that are as imaginative and complex as when they play with other children [1]. Including a virtual embodied peer and toys that children can play share with the character may enhance the believability and engagement of the interaction.

IMPLEMENTATION

Sam detects a child's presence through a microphone, and two pressure-sensitive mats placed in front of the castle. Once the child sits in front of the castle, the agent's gaze at is guided by where the child sits on the mats. When the child is playing with the toys and narrating, the system uses audio threshold detection to determine when to give feedback (backchannels such as "uh-huh" nods, and explicit prompts such as "and then what happened?"). A Swatch RFID tag embedded in the figurine's cape tells the system when the figurine is in the magic tower. A switch in the door tells the system when the magic tower door is opened, so that the child will never see the physical and virtual instantiations of the toy simultaneously (when the door is opened and Sam has the figurine, it disappears instantly and Sam expresses surprise). Sam's stories and other utterances were recorded from a real child, since the quality of children's synthesized voices is poor. The software is written in Java and C++ and can run on a single PC with a graphics acceleration card. The animation is displayed on a back-projection screen behind the castle.

Children (and adults!) who have tried the system spend considerable time playing and improvising stories. We are currently preparing for more substantial user testing.

FUTURE WORK

In the near future, we will include more sensors in the castle, so the system knows where the toy is. This will allow Sam to look at the toy as the child moves it, and to base the story segment selection on the toy's location. We will also add more toys to enrich the experience. The timing and quality of Sam's feedback behavior will be improved by adding intonation detection. Finally, we are investigating using keyword spotting to give the system some understanding of the content of the user's speech, to make the experience even more collaborative.

CONCLUSION

This paper introduced a novel combined virtual / physical interface in which a human shares a physical space and objects with an animated conversational agent. We described Sam, our first implementation of this interface, a system designed to elicit children's collaborative storytelling in a engaging, playful way. This interface is especially suited to supporting children's storytelling play, but it can also be successfully used in a wide range of collaborative and educational applications.

ACKNOWLEDGEMENTS

We thank Ling Bao and Nina Yu. Also Javaneh Mikhak, Sam's voice and storyteller extraordinaire in her own right.

REFERENCES

1. Cassell, J., and Ryokai, K. (forthcoming). "Making Space for Voice: Technologies to Support Children's Fantasy and Storytelling." *Personal Technologies* 3 (4)
2. Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., and Yan, H. (1999). "Embodiment in Conversational Interfaces: Rea." *Proceedings of CHI '99*. ACM, 1999.
3. Fein, G., "Pretend Play: Creativity and Consciousness," in *Curiosity, Imagination, and Play*, Gorlitz, D. & J. Wohlwill, Eds. Hillsdale, NJ: Erlbaum, 1987.
4. Ishii, H. & Ullmer, B. "Tangible bits: Towards seamless interfaces between people, bits, and atoms," in *Proceedings of CHI '97*: ACM, 1997.