Extending Sam:
Utilizing Multiple Toys to Enhance Interactive Storytelling

by

Linda Lin

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in
Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

May 25, 2001
Copyright 2001 Linda Lin. All rights reserved.

The author hereby grants to M.I.T. permission to reproduce and
distribute publicly paper and electronic copies of this thesis
and to grant others the right to do so.

Author

Department of Electrical Engineering and Computer Science
May 25, 2001

Certified by

Justine Cassell
AT&T Career Development Associate Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by

Arthur C. Smith
Chairman, Department Committee on Graduate Theses
Extending Sam:
Utilizing Multiple Toys to Enhance Interactive Storytelling
by
Linda Lin

Submitted to the
Department of Electrical Engineering and Computer Science

May 25, 2001

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Sam is a life-sized on-screen character who shares a physical castle and a plastic figurine
with children, and takes turns with children to tell stories using the figurine. In this thesis,
the original Sam system is improved. Now Sam shares two figurines with the user, and
tells fantasy stories from the narrator’s perspective. When Sam takes turn to tell story,
Sam identifies the toys in the virtual world, plays with the virtual figurines, and tells
stories involving the specific characters. When the child tells a story, Sam displays
appropriate listening behaviors according to the physical location of each toy. The new
system encourages children to tell more interesting stories in a narrator’s perspective.

Thesis supervisor: Justine Cassell
Title: AT&T Career Development Associate Professor of Media Arts and Sciences, MIT
Acknowledgements

First and foremost, I thank Professor Justine Cassell for her kind advice and continuous encouragement throughout my work. I would also like to extend thanks to everyone in the GNL group. I would like to thank Jennifer Smith, who first introduced Sam to me. Thanks to Hannes, for always being cheerful, patient, and helpful, and always answering my random technical questions. Thank Kimiko and Cati, for providing me a large amount of resources throughout my research and implementation of the project. Thanks to Tim and Lee, for their willingness to discuss random ideas with me. Thanks to Liz, Yukiko, and Maruko, for giving me great suggestions about Sam’s interactions. Thanks to Ian, for your promptness in modeling the new castle, and for your wonderful jokes that brought laughter to all of us. Thanks to Jean, who has given Sam a new voice and a new personality. Thanks to Jane, who constantly exchanged encouragements and great ideas with me over ICQ messages.

Next, I would like to thank a few other Media Lab students who made my experience at the lab more enjoyable. Thank Xin Wang for having interesting chats with me and providing me great insights when I take breaks. Thank Pengkai Pan for being a cheerful and humorous friend, and lending me valuable equipments. Thank Cathy Lin for being supportive.

I have to thank all the friends who have been an integral part of my MIT life. Particularly, thank Gina for being a nice roommate, a considerate neighbor, a collaborative work buddy, and a supportive friend; thank Yu Chen for being a cheerful and encouraging friend and a helpful roommate; thank Lei Wang for being a sweet and understanding person to talk to; thank Sabrina for always being energetic, supportive, and fun-seeking; thank Alice for organizing lots of Karaoke trips; thank Lucy for being a nice roommate. I would also like to thank all the friends that I have made at MIT, but I don’t have room to list everyone’s name here. Every friend has made me a better person in some way, and my MIT experience would have been meaningless without all of you.

I would like to extend infinite gratitude to Minghao, for supporting my work throughout the entire year. Thank you for bringing late-night snacks to the lab, providing rides when I am exhausted, cooking me dinners when I am busy, keeping me company at the lab when I pull all-nighters, and, most of all, thank you for your constant encouragement and always being there for me. I thank you with all my heart.

Last, but not least, I thank my parents for making me who I am today. Thank my mother, who constantly cares about my happiness and well-being and strives to give me the best advices. Thank my father, who always had great confidence about my success. Thank you for your spiritual encouragement during the countless phone conversations in the past five years. I would not have succeeded well without your comfort and support.
# Table of Contents

Chapter 1. Introduction ................................................................. 5

Chapter 2. Background................................................................. 7

2.1 Pretend play and storytelling............................................... 7

2.2 Peer influence on children's play........................................... 13

2.3 How toys influence children's play....................................... 16

Chapter 3. Context: Sam............................................................... 20

Chapter 4. Limitations of the previous Sam system....................... 22

Chapter 5. Solutions................................................................. 24

5.1 Enabling Sam to use multiple toys........................................ 24

5.2 Technical Implementation..................................................... 28

Chapter 6. Evaluation............................................................... 43

Chapter 7. Conclusion and future works.................................... 45

References................................................................................... 47
Chapter 1. Introduction

Childhood may be too long ago for us to remember clearly, but most people can probably still recall a few favorite toys that they played with in their childhood years. Some toys may have brought us great joy and inspired our imagination and creativity, and these toys may never fade away from our memory.

After being isolated from toys for many years, I walked into a toy store again this year, and discovered numerous toys that I had never seen as a child. Some toys can talk, move, smile, react to sound and touch, and even make funny faces! The modern technology has revolutionized the children’s world, as the new generation of children have the opportunity to play with toys that incorporate the latest technologies.

Engaging in play is children’s nature. All children like to play, to fantasize, to talk, and to tell stories. They develop their mental, social, and language skills through playing and telling stories. Sam is a system that incorporates modern technology and encourages children to play and to tell stories. As a friendly virtual companion, Sam shares a castle and a plastic figurine with a child, and through interactive play, Sam and the child can take turns to tell each other stories. By playing with Sam, children exercise their creativity and imagination from the fantasy play and storytelling experience.

In the past year, Sam engaged in numerous play activities with children. Children showed great excitement and joy when they played with Sam, and they seem to love to share the figurine with Sam while taking turns telling stories. However, children have consistently sought more figurines to act out their play. Most stories involve more than one character, and the characters have to interact with each other. Using only one toy, Sam tells stories in the perspective of one character, and the other characters in the story
are imaginary. In turn, the child is also forced to invent stories from one character’s point of view.

After observing children’s play patterns and storytelling sessions, and reading background material about toys’ roles in fantasy play, I decided to improve Sam to enhance children’s playing experience with this companion. By allowing Sam to share two toys with the child and to tell stories from the narrator’s perspective, the child’s fantasy play experience would be more smooth and enjoyable, and the child would have the freedom to use more toys to act out the fantasy stories and practice his or her linguistic skills using different voices.
Chapter 2. Background

The interactive storytelling system, Sam, is a virtual peer designed to encourage children between the age of 5 and 7 to tell stories through pretend play. In this chapter, I will first provide background of children’s storytelling and pretend play. Next, I will show the importance of collaborative peers in such storytelling and pretend play. Finally, I will explain the role of the toys used during pretend play and storytelling.

2.1 Pretend play and storytelling

2.2.1. Pretend Play

Children start engaging in pretend play from the age of 2. (Piaget, 1962). However, the pretend play of the 2-year-olds is solitary. Children play alone with toys, and do not attempt to interact with other children. They are preoccupied with only their own points of view, and are not able to take on different perspectives.

At the age of 3, children are focused on “parallel” play. (Peller, 1955). A child attempts to play in a similar way as other children in proximity. However, the children do not interact in their play, and their play activity is parallel rather than intersecting.

Children between the ages of 4 and 7 are actively engaged in more social play. (Vygotsky, 1967). During this period, a child develops skills to take another person’s perspective and share the world with others. As children get older, they become more competent in cooperating with other children in their play, and they are able to engage in make-believe activities.

In children’s spontaneous play, they integrate their thought, actions and language, and display in language and in thought a sense of possibility, the concept of “what might
Children transform real objects into imaginative objects and characters, and define their play based on the newly defined objects and roles. By fantasizing and pretending, children are able to move their perception and thought away from the concrete reality to explore new possibilities in the newly defined world.

Young children under 5 years of age are more likely to construct plays involving everyday reality and assume their everyday personal identity, while older children construct staged fiction and assume fictional role identity during their role play. (Auwarter, 1986). As a child’s age increase, the level of completely staged fiction also increases. In pretend play, younger children are more likely assume their everyday personal identity and construct realistic play involving everyday situation, while older children over 5 years of age are able to assume fictional roles and construct appropriate structured fictional context and produce fictional narratives.

2.2.2. Storytelling

In fantasy play, children practice their language skills, and their language style becomes more sophisticated with development. As children grow older, their fantasy play becomes more storytelling. Children are able to integrate roles, events, and consequences (Scarlett and Wolf, 1979) and generate stories during their play. Children learn to exercise story language as they explore different possibilities with language in their fantasy play. For example, children may learn to use story language such as incorporating beginnings with an opening phrase such as “Once upon a time” and endings with a closing “The end” (Sacks, 1972).
While expressing their ideas and actions using language, young children predominantly speak as themselves, while older children over 4 years of age are able to speak from another character’s perspective. As children become older, their ability to tell stories from different perspectives increases, and they develop the ability to generate stories from the neutral perspective (narrator’s perspective.) (Auwarter, 1986). While younger children are able to narrate from their own perspective as in “I’m gonna swim over there!,” older children are able to narrate from the narrator’s point of view, as in, “He is swimming in the pool!”

Make-believe / fantasy play is also called “symbolic play” in Piagetian terms because during the fantasy play, a child represents objects or actions with something that stands for them. Language is one of the tools that children symbolically use. (Piaget, 1962). In symbolic play, children move in and out of a stream of enacted events. Symbolic play lacks the structures about how episodes connect and how problems are solved. On the other hand, in storytelling, each action must follow from the previous narrative events. The narrative world must appear autonomous and self-sufficient. The narrator must structure how his or her audience will interpret the events within the story. (Scarlett and Wolf, 1979.)

2.2.3 Benefits of pretend play and storytelling

Researchers have found abundant evidence suggesting that make-believe play and storytelling are beneficial to children’s development in many aspects, including mental and linguistic development.
When children engage in symbolic games and pretend plays, they are practicing mental skills that will help them later on in their life, just like practicing walking, balancing, or swimming will foster development in motor skills. Through pretend play, the child tries to understand and imitate the actions and speech in the adult world. (Singer & Singer, 1990).

In pretend play, a child learns to represent objects as other possibilities (Leslie, 1987). For example, a wooden block may represent a car, a train, or a bed. Engaging in fantasy play helps the child to develop awareness of the distinction between fantasy and reality (Singer & Singer, 1990). Singer (1990) claims that make-believe is critical for developing in the child a full-fledged theory of mind, an awareness that one can manipulate toys or imaginary playmates and transform them into more general ideas. Make-believe allows children to create a fantasy world in their own consciousness, and to understand that real objects can be transformed to serve other purposes in the imaginary world. In play, children are able to ignore the ordinary uses of objects and actions, and free themselves from the constraints in time and space (Rogoff, 1990) to develop their own activities.

In fantasy play and storytelling, children are in control of each character’s actions and emotions. They can explore different possibilities without the risk of failure in real life. No matter how unreasonable and seemingly impossible the actions are, they are only part of a fantasy story in an imaginary world. Fantasy play and storytelling provide children a place where they have control in the fantasy world, and they can experience the characters’ emotions through each character’s emotional actions and reactions. (Fein, 1987; Csikszentmihalyi, 1981).
Studies have shown that pretend play helps the children’s well-being. Research studies of 3- and 4-year-old children followed for 1 year suggest that the children who played more often at make-believe or who developed imaginary playmates are in better mental state and are more socially apt. For example, children who are frequently engaged in pretend plays and fantasies are found to smile and laugh more frequently in school, and are more persistent, more cooperative, and less likely to be angry, aggressive, or sad. (Singer & Singer, 1981; Singer & Singer, 1990). In studies of imaginative older children, the results showed that they are less aggressive, less impulsive, and better able to discriminate reality from fantasy. (Singer & Singer, 1981; Singer & Singer, 1990.) In a longitudinal study, Tower and Singer (1980) found that children who engage in imaginative play had higher score in positive emotions such as liveliness, excitement, elation, and joy than those who did not. In another study, Connolly, Doyle, and Reznick (1988) found more positive and less negative affect during social pretend play than during non-pretend social activities such as bowling games. Moreover, research have also shown that children with imaginary playmates are often better behaved and later become more creative (Singer & Singer, 1990.)

One of the most important benefits of pretend play and storytelling is that they foster children’s linguistic skills. Fantasy play provides children various possibilities for language usage. In fantasy play, children often imitate adults’ conversations. When playing different roles, the child would change the tone to express different emotions of different characters. Consequently, children would learn appropriate emotional tone for different contexts. (Singer, 1990).
In social fantasy play, children learn to use decontextualized language. Decontextualized language conveys meaning through lexical and syntactic means, and minimizes the reliance on contextual cues and shared information. For example, decontextualized language may involve using language to define a play theme or to redefine a toy, and does not assume any prior information. Pretend play provides children the opportunity to use language to redefine play props and toys, and enables children to develop the ability to use decontextualized forms of language. By practicing decontextualized language in their play, children learn that language can be meaningful representations of something else. (Pellegrini and Jones, 1994).

When children incorporate storytelling in their play, they develop their language skills even further. As children tell their own stories during fantasy play, they also practice their language skills (Nicolich, 1977). A child learns to speak from different characters’ perspectives as well as the narrator’s perspective. (Auwerter, 1986). Bamberg (1997) has also argued that children use more mature language in narrative than in everyday conversation. In narrative storytelling, children practice their ability to represent objects, actions, and feelings in language (Nicolich, 1997). Therefore, fantasy play and storytelling lead to an effective use of more complicated and sophisticated narrative language during children’s play.
2.2 Peer influence in children’s storytelling

Peers play an important role in children’s development. By interacting with peers, children learn to cooperate with each other. Without control and rules posed by adults, peer play and storytelling are more creative and enjoyable. The playful nature allows children the freedom to explore more possibilities, thus leads to a better learning environment and more interesting and sophisticated pretend play and storytelling.

It has been shown that in some domains, children learn better with peers than they do with their parents. (Corsaro, 1985). Children’s interaction with peers is reciprocal, while their interaction with adults is complementary (Piaget, 1965). In fantasy play, the complementary interactions such as the teacher-student relationships can inhibit sophisticated fantasy. On the other hand, the reciprocal relationship among peers facilitates social fantasy. (Pellegrini and Jones, 1994). With children older than 3.5 years of age, adult’s presence seems to inhibit fantasy play (Pellegrini, 1984). For these children, sophisticated forms of play and uses of language occur in the presence of peers, not adults (Pellegrini, 1983). For children that are older than 3 years of age, adults inhibit social fantasy and corresponding forms of language, whereas peers facilitate them (Pellegrini & Perlmutter, 1989). In the peer context, children generate verbal fantasy related to social perspective taking, whereas the fantasy with parents generally did not. (Perlmutter & Pellegrini, 1987). Dunn and Dale (1984) found that children’s play pattern with their siblings are different than with their parents. When playing with siblings, children feel more free to engage in pretend games involving transformations of role.
identity, location, or psychological state, while their play with mothers mostly involved labeling or acting on realistic objects.

Engaging in pretend play and storytelling with peers also helps children’s logical and language skills, and allows children to produce more complicated stories. Language is the tool used to create, clarify, maintain, and negotiate children’s pretend experience. (Garvey & Berndt, 1985). In conversations with peers, children’s moral and logical reasoning skills are better than when they talk with adults. (Kruger & Tomasello, 1986). It has also been found that children are more likely to tell stories and exercise sophisticated narrative skills when other children are present than when they are alone (Cassell & Ryokai, in press.) When engaged in storytelling with other children, peer stories serve as new suggestions, and encourage children to tell their stories within an appropriate dramatic frame. (Baker-Sennett, 1992). Children incorporate elements and phrases of other children’s stories into their own, and create stories that are more complex (Corsaro, 1992; Preece, 1992; Paley, 1990). Even if one partner’s attention has lagged in telling a story, it could be revived by the others’ suggestions and new views, thus take children’s stories even further (Bos, 1937).

Preece (1992) found that peer audience plays and important role in supporting children’s storytelling. When children listen to their peers’ stories, children act as active, alert, and engaged listeners. Their peer narratives are usually received with interest, are granted close attention, and are carefully monitored for content and manner of delivery. Children on occasion also invite peers to tell a story, and the peers are generally receptive to such invitations. The presence of an attentive and receptive audience encourages the practice of narrative skills. By creating original stories of their own, the children model
storytelling for each other; by inviting each other to tell stories, they confirm each other as narrators; by being demanding, they stretch each other’s linguistic resources. Consequently, children learn from, and teach, each other during their conversation and storytelling.

Some resources have suggested that children’s siblings can encourage their creative pretend play. Singer & Singer (1990) have found that childhood memories of pretend play are often associated with a special person who encouraged play, told fantastic stories, inspired imitation and showed a trusting, loving acceptance of the children and their playfulness. This important person is sometimes a sibling. Siblings often foster each other’s imaginations and accept one another’s plots, and at the same time, they foster a playful spirit in each other.

Through telling a story to a peer, children realize that their partner may not always share their perspective. After they interact with an age-mate, the children would realize that peers are not as responsive and attentive as the caregivers, and do not share the same background knowledge with them. Thus children would attempt to achieve a shared understanding. Through use of language, the child has to construct stories that are meaningful to others, thus would learn to use different perspectives. (Budwig, Strage, and Bamberg, 1986).

It also has been suggested that children prefer same-gender peers during pretend play. When children engage in fantasy play, they play longer and in more complex ways with same-gender playmates and expend considerable cognitive and linguistic resources to negotiate the conflicts. (Huston, 1983; Rubin et al., 1983). Pellegrini & Perlmutter
(1989) found that cross-gender play actually inhibits girl’s exhibition of competence during fantasy play.

2.3 How toys influence children’s play and storytelling

When children engage in fantasy play and storytelling, the toys or props they play with are important objects that inspire children’s creativity. Toys can spark interest and engagement, and initiate elaborations in play and in language. Puppet theatres, toy soldiers, dollhouses and dolls have brought children great joy and inspired children to invent stories involving the characters. Children name their toy figurines, give them personalities and occupations, and use the toys to act out their fantasy stories. (Singer & Singer, 1990)

Some studies show that very young children need structured and realistic toys representing everyday objects in order to engage in pretend play. However, as children get older, their representational skills are more developed, and they enjoy transforming less realistic toys into fantastic objects in their play. (Singer, J., 1994; Connolly et al., 1988). As their use of language increases, they are able to communicate more about their pretend acts, and their interest shifts to less realistic toys. (Singer, J., 1994).

Researchers have found that different types of toys encourage children to engage in different types of play, and consequently elicit different types of language use. Realistic toys and fantastic toys have different effects on children’s play and language. When children play with functionally explicit toys or realistic toys, such as constructive toys like art materials, their play is less imaginative and more constructive, and their language tends to be contextualized. (Singer & Singer, 1990; Pellegrini, 1982, 1983).
When children play with toys that display fantastic themes, such as dolls or dinosaurs, they engage in fantastic play and use decontextualized language. The ambiguity inherent in fantasy elicits decontextualized language from the participants. To initiate and maintain fantasy play, children must use decontextualized language to explicitly explain the scenario, thus allowing other players to understand the play theme. (Singer & Singer, 1990).

A number of studies have suggested that compared with highly structured toys, minimally structured and low-realism toys would elicit significantly more varied themes and richer fantasy from older children’s play. When children play with realistic toys, their play is more context bound; on the other hand, when children play with less realistic props, their play tends to explore a greater variety of themes, and the language used in the play is more decontextualized. (Pellegrini and Jones, 1994). Pulaski (1973) presented two groups of toys to children from kindergarten throughout second grade. The first group of toys included, among other objects, a rag doll, costumes, and wooden blocks, while the second group of toys included Barbie dolls, GI Joe doll, and specific outfits for these dolls – a nurse’s uniform, a bride’s dress, an army uniform, and an astronaut’s suit. The results showed that the first group of minimally structured toys elicited significantly more varied themes and richer fantasy. Other studies have shown that high-fantasy children typically express greater interest and delight in toys, and generally respond more strongly to low-realism than to high-realism toys, while low-fantasy children prefer the latter. (Almqvist, 1994). It has also been suggested that toys that serve multiple functions seem to be used more extensively in children’s development of story materials. (Singer J., 1994)
Children have been continuously attracted to groups of figurines during their pretend play and storytelling. Miniature figures such as people, family figures, or even animals are useful to elicit fantasy play (Singer D., 1994). Clusters of related toy figures, as well as puppet theatres, are especially conductive to the development of pretend story lines, and play a special role in enhancing narrative and imaginative skills. (Singer J., 1994). By acting out creative scenarios with a group of figurines, a child can manipulate these figures, take on different roles and perspectives, and change voices, thus learns and practices a variety of ways to use language. (Pellegrini and Jones, 1994) Through the interaction among the miniature characters, a child can make sense out of a confusing larger world. (Singer D., 1994).

A study has suggested that when a child's story involves a conflict or problem, the narrative is better structured. (Benson, 1993). Compared to sequential narratives, plotted narratives are more likely to include a problem or conflict. With only one toy, it is possible to generate stories involving problems and conflicts, and consequently produce a solution to the problem. However, the problem has to occur between the character and an imaginary role or an object. Multiple toys are more likely to elicit a narrative involving conflicts. With more than one figurine, the child can easily manipulate the minds and actions of different roles, create a problem among the characters, and consequently have the characters solve the problem, leading to a better plotted and structured narrative.

Depends on the look and nature of each toy, different toys are preferred by different gender groups. Children generally play longer and in more complex ways when they play with gender-preferred toys (Huston, 1983; Rubin et al., 1983). Girls engage in fantasy more frequently than boys with both the male-preferred and neutral toys,
(Pellegrini & Perlmutter, 1989), and girls are usually more liberal in their attitude toward opposite-gender toys than boys are (Almqvist, 1994). Because of gender differences in play, Singer J. (1994) suggested that sets of toy figures should either include both male and female characters in costumes that cut across traditional sex-role lines or should be androgynous so that boys and girls will use them flexibly.
Chapter 3. Context: Sam

Although peers are good at encouraging children to tell stories and practice their language skills (Preece, 1982), peers are not perfect learning partners. Children do not always share or play collaboratively. They are not always capable of scaffolding each other’s narratives to help peers’ language acquisition, and they do not always listen quietly to allow other children to revise and reorganize their narrative constructions. (Cassell, 2001).

Sam is a technological system that fosters children’s imaginative mind and encourages children to engage in fantasy play and to tell stories. Sam is a life-sized on-screen character who shares a physical castle and a plastic figurine with children, and takes turns with children to tell stories using the figurine. (Figure 1). When the child finishes his or her turn, the user would place the plastic figurine inside a “magic portal” in the castle, and Sam would take turn to play with the figurine. When Sam takes turn, Sam manipulates the virtual version of the figurine to tell fantasy stories, and provides the child with ideas for incorporations. After Sam finishes a story, Sam places the figurine back into the “magic tower” to allow the child to play with the physical toy. While the child tells stories, Sam displays listening behaviors and gives appropriate verbal and nonverbal back-channel feedback and follow-up questions. (Cassell, Ananny, Basu, Bickmore, Chong, Mellis, Ryokai, Smith, Vilhjálmsson, Yan, 2000). These interactions support children to practice their narrative skills. Early results from a study suggest that children enjoy interacting with Sam, and incorporate elements of Sam’s stories when they invent their own narratives. Sam seems to have the effect of stretching...
the limit of children’s narrative skills, and allows children to utilize their creativity during fantasy play. (Cassell, 2001).

Figure 1. Old Sam System
Chapter 4. Limitations of the Previous version of Sam

In the previous version of the system, Sam recognized the presence of one specific figurine in the castle. When the toy is in the magic tower and the portal door is closed, Sam understands that the toy is in the virtual world and takes turn to tell stories; when the toy is in another room in the castle, Sam knows its location, and her eye gaze follows the toy when she is listening to the user’s story.

However, this scheme limits the user and Sam to play with only one toy throughout the entire interaction. The turn taking is determined solely by the possession of this toy. The toy that Sam and the user shared is a plastic figurine that resembles a prince or princess in a cape, thus it is a highly structured toy. There were a few other toys in the castle, including a king and a queen, but Sam does not know their presence, and would not play with them. When the user plays with these other figurines, Sam does not respond accordingly.

Since Sam is telling stories with only one figurine, the child must also tell stories with only one figurine to enable Sam and the child to smoothly share the same physical objects. Thus during storytelling, the child would generally take on one perspective of the main character, and does not have the opportunity to practice language skills through different voices.

All of Sam’s earlier stories involved mainly one character, a prince / princess living in a large castle. The stories are told in the character’s point of view, and the word “I” is used throughout the stories. Sam is immersed in only one character’s role, and does not speak from different perspectives or the neutral perspective. The virtual toy in Sam’s
hand represents the main character, and other characters involved in the stories are imaginary. These stories that Sam told were all about one prince/princess, hence emphasized the fact that the figurine is highly structured. The figurine has taken on one specific role, and may limit the child's imagination and creativity.

In the previous version, Sam's stories were highly contextualized. Sam used first-person point of view in the stories, and implied that he/she is a prince/princess living in a castle. However, Sam had never introduced himself/herself or the location and context to the user while telling the stories. Since children model their language after their peer's stories, it would help children to use decontextualized language if Sam uses decontextualized language.

From my observation of the videotapes featuring children playing with toys and the castle, I noticed that children like to play with different toys at different times, and sometimes a child would even play with more than one toy at the same time. A child would sometimes hold a toy in each hand to tell a story that involves more than one characters. When children played with Sam in the past, they have consistently requested more figurines to play with. Based on the background research and past data on children's play patterns, I believe that allowing Sam and the child to share more than one minimally structured figurines, along with adding to the system new stories involving multiple characters, would enhance the child's imagination during storytelling.
Chapter 5. Solution

5.1 Enabling Sam to use multiple toys

The data analysis of videotapes of children’s playing and storytelling suggested that children tend to play with more than one toys at a time. Frequently, a child would move several toys around the castle while telling a story that involves multiple characters. The characters in the story would extensively interact with each other. While one child is telling a story, the listener’s eye gaze usually follows the toy that has most recently moved.

I decided to incorporate this pattern of interaction into Sam so that she would be able to tell stories involving multiple characters, and be able to direct the appropriate eye gaze while the user is taking a turn telling a story. In addition to changing the interaction, we also changed Sam’s stories to involve more than one characters, and made Sam use decontextualized language in the stories.

We have gotten a new wooden dollhouse that has a large “magic portal”, which enables the child to place two figurines inside the portal (Figure 2). These figurines include a boy character and a girl character, allowing both girls and boys to engage in creative story telling. The new figurines that we have chosen for Sam are not highly structured. They look like young children with no specific profession or magic power (Figure 3).
Figure 2. Sam’s new castle
Sam is now able to differentiate between the two toys. When one of the toys is placed inside the magic tower and the portal door is closed, Sam takes a virtual figurine that resembles the specific toy in the virtual world, and tells a story that is related to this character. The story may also involve more than one character, and in that case, Sam’s figurine represents the main character in the story while the other characters are imaginary. During Sam’s story telling, the animation of Sam and the virtual character corresponds to the specific story being told.

When two toys are both placed in the magic tower and the portal door is closed, Sam is able to recognize both figurines’ presence in the virtual world, and would tell stories that involve at least two characters. In the animation, Sam is able to hold more than one toy at the same time, and move them around the castle while telling stories.
involving these characters. The two figurines in Sam’s hand now represent the two main characters involved in each story.

While Sam is listening to the user’s story, her eye gaze is directed appropriately toward the toys. If the user is playing with more than one figurine, Sam knows each figurine’s location and movement around the castle. Her eye gaze follows the character that moved most recently. If a figurine moves to a different room in the castle, Sam directs her eye gaze to the figurine’s new location. If the figurine has moved out of the castle, Sam tilts her head up and directs her eye gaze toward the user.

Sam’s new stories now involve different characters and settings. In Sam’s stories, the dollhouse became school, spacecraft, farmhouse, tree house, ordinary house, and castle. This would encourage the user to be creative and transform the playhouse to be any imaginative structure. The characters involved in the stories include boys, girls, kids, animals, monster, and fairy, thus showing the user that she can make the simple figurines play a wide range of characters. Because previous research suggest that stories involving problems and conflicts tend to be better structured narratives, most of the new stories involve problems or conflicts, and solutions to the problems. This is to encourage the child to creatively produce structured stories that involve problems and solutions.

The new stories take place in specific rooms in the dollhouse, thus during storytelling, Sam moves the virtual toys to appropriate rooms in the virtual castle. After the user finishes a story in a specific room of the physical castle, Sam takes the virtual toy and tries to tell an untold story that starts in the same room that the user finished the story.
Most stories involve more than one character. When a user passes Sam both toys, Sam tells a story involving at least two characters and utilizes both figurines to represent the characters. When Sam possesses only one toy, she picks an untold story that starts at the location where the user’s last story ends, and tells the story while manipulating one virtual figurine that resembles the toy in the portal.

The new stories are read by one of our undergraduate students, and each story is read with proper emotional tone. All of the new stories are told from the neutral perspective. Sam is the narrator, and uses decontextualized language in the stories. At the beginning of each story, Sam introduces the setting and the characters. Occasionally, Sam plays different character’s roles when she imitates their voices, and changes her voice and tone to express the emotions of different characters.

5.2 Technical implementation

The technical implementation of new Sam involves a wide range of sensors and new software design. The sensors include a number of RF sensors to detect the location of the figurines, a door sensor to facilitate turn taking, and a motion sensor to detect user presence. The software design enables the software to take inputs from the sensors and correctly interpret the signals, and in turn directs Sam to generate appropriate behaviors, animation, and speech.

5.2.1 Overall setup
Two computers are used to run Sam. One computer runs the main brain of Sam as well as animation and speech output, and also takes signals from the user presence sensor and the door sensor. The auxiliary computer takes signals from the toy location sensors and audio sensor, and processes and sends data to the main computer over the network.

5.2.2 Overall software architecture

Figure 4 shows the overall software architecture of Sam.

![Software Architecture Diagram](image)

**Figure 4. Software architecture of Sam system**

There are four sensor modules, including Audio Sensor module, Motion Sensor module, Toy Location Sensor module, and Portal Sensor module, that directly interact with the hardware sensors through serial port, and report the status of each sensor to Understanding Module. Understanding Module resolves the signals from the sensor modules using multimodal integration, and interprets the current state of the user, the castle, and the toys. For example, if the Toy Location Sensor sends a signal that a toy is present in the portal, and the Portal Sensor sends a signal indicating that the portal door is

---

29
closed, the Understanding Module would resolve the two signals as “one toy is in virtual world.” Understanding Module sends the input state to the Reaction Module, which essentially behaves as a state machine. In each state, Sam exhibits a different behavior. Depending on the input from Understanding Module, Reaction Module decides what the next state is. For example, if Sam were in the “idle” state, and Understanding Module signals “User Present,” the Reaction Module would take Sam to the “greeting” state. To generate specific behaviors, Reaction Module sends events to Generation Module. For example, in the “greeting” state, Reaction Module sends events to Generation Module display the “greeting” behavior. Generation Module would in turn run a script which dictates when to raise the right arm and wave, when to say the phrase “Hi, I’m Sam!” and when to raise the eyebrows. Generation Module sends events to Action Scheduler to realize each atomic action. Action Scheduler is responsible for outputting each speech and action in the lowest level, and execute each command in a synchronized way. Action Scheduler controls the Speech Out Module and the Animation Module, which are each responsible for speech output and animation output.

5.2.3 Toy Location Sensors

Six Easy Key R/W TTL Multireaders (figure 5) and two transponder tags are modified and set up to detect the location of each toy. Each toy has a transponder tag attached to it, and when a tag is in the vicinity of the antenna of a sensor, the sensor detects the tag’s presence and identity.
An Easy Key sensor consists of two parts: a reader circuit in the center and a solenoid antenna connected to the reader. When a swatch tag is placed on top of the antenna, the reader circuit reads the identification code from the tag, and via a serial port on a computer, it reports signals indicating the tag presence and the tag identification. Two original Easy Key Multireaders are placed underneath the portal of the castle to detect toy presence. When the user places one or two toys in the “magic portal” for Sam to tell stories, these sensors send appropriate signals.

Sam also needs to react when a toy is inside any of the rooms of the castle. Since the original circular antenna on the Multireader sensor is too small for the large rooms in the new wooden castle, I made large antennas for the sensors to detect toy presence in the four castle rooms. For each of the four sensors, I used copper wire to wind around a fiber cylinder approximately 20cm in diameter to make a new antenna, and replaced the old antenna on the Multireader sensor with the newly made ones. I cut off the connection of the old antenna from the reader, and soldered two ends of the new antenna to the corresponding nodes in the original reader. (See figure 6.) This new sensor is guaranteed to report tag presence when a tag is placed anywhere on top of the new large antenna,
thus enlarging the area of detection. Each of these four sensors are placed underneath a room of the wooden castle and attached to the castle by Velcro.

![Figure 6. New sensors](image)

Each of the six reader circuits of the Multireader sensors needs 5 volts of DC voltage. I used two 9 volts DC adapters and a few resistors to provide 5 volts for each tag reader. The circuit diagram is shown in Figure 7.

![Figure 7. Circuit diagram of the wiring of tag readers](image)
In order to access the signals generated by each of these sensors, I used a Digi Acceleport USB to serial adapter. This adapter is connected to the USB port of Hyperion, and provides eight additional serial ports to the computer. Each of the six Multireader sensors, two for the portal and four for the castle rooms, are connected to the serial ports on the USB-serial hub, and their signals are reported to Hyperion through the serial connection.

Since Java code does not recognize the presence of the new serial ports on the hub, C++ code is used to communicate with these serial ports to retrieve the signals sent by the Multireader sensors. The C++ code encodes the information read from the serial port as an integer, and passes this integer to Java code through JNI. The Java program, Toy Location Server, decodes the integer and interprets the signal, and sends corresponding information about the tag / toy identification and toy location through the network to Toy Location Sensor module on the main computer.

5.2.4 Portal sensor

The Portal sensor detects whether the portal door is opened or closed. The hardware consists of a hinge fixed to the castle, a small piece of metal fixed to the door, and wires connecting the hinge and the metal piece to a serial port on Sprite. The hinge on the castle and the metal on the portal door are connected to pin 4 (DTR) and pin 8 (CTS) of a DB-9 serial connector respectively. A resistor (51KΩ) is connected between pin 8 (CTS) and pin 5 (GND). See figure 8 for the circuit diagram.
Figure 8. Circuit diagram of portal sensor

The software detects the state of the serial port, and determines whether the portal door is open or closed.

5.2.5 Motion sensor

The hardware of the motion sensor consists of a disassembled motion detector, wires, and a serial port connector. When motion is detected in front of the lens, a 120V AC voltage difference is present across the ends of two wires. I soldered these two wires to the input of an AC adapter that converts 120V AC voltage into 10V DC voltage, and soldered two additional wires from the output of the AC adapter to two pins of a serial port connector. This serial port connector is attached to the serial port of Sprite, the computer that runs most of Sam’s software.

Similar to the software of the door sensor, the software of the motion sensor detects the state of the serial port, and determines the presence of a user in front of the castle.
5.2.6 Software

The major software changes involved in this new version includes modifications to Java and C++ programs that interpreted the signals from the toy location sensors (Toy Location Server and Toy Location Sensor), the Java programs that processed signals from the motion sensor (Motion Sensor Module), the Understanding Module that resolves different signals, the Reaction Module that decides the next state based the status of the toys, the Generation Module which allows Sam to do different actions, and the text and data files that Sam uses to generate appropriate animation.

The software that interacts with the toy location sensors is on the auxiliary computer. In these programs, the C++ code queries a tag reader by sending a string “a” through a serial port. If there is a tag on top of the antenna attached to the reader that is connected to this specific serial port, a string including the tag ID is sent back to the serial port. The C++ code encodes the serial port number and the tag ID into one integer and returns the integer. The Java code starts 6 threads, and in each thread it calls the C++ function once every 300 milliseconds through JNI to query a specific reader from a specific serial port. Based on the integers returned by the C++ function, the Java code determines the current location of each figurine. If it detects a new figurine from a particular sensor, it sends string over the network to the main computer indicating the location and the new toy number. For each tag reader, if no figurine is detected continuously after 6 queries, it concludes that the figurine is not there anymore, and sends a string over the network to the main computer indicating that there is no figurine in that particular location.
On the main computer, the toy location sensor code parses the string and determines the location and the new figurine that was moved to this location, and sends an event to Understanding Module.

The software that interacts with the motion sensor and the door sensor are very similar. The software sets DTR (pin 4) to true. When the switch is closed, DTR is effectively connected to CTS (through the resistor), so CTS registers true. The software checks for a change in state at CTS. In the door sensor software, true means the door is closed, while false means it is open. On the other hand, in the motion sensor software, true means the user is present, and false means user is absent. Both of these programs send an event to the Understanding Module when a change of state is detected.

The Understanding Module resolves the events from different sensor programs. If the event received is a toy location event, it gets the toy number and the toy location from the event, and decides whether each toy is in virtual world or physical world. If currently the portal door is closed and one or two toys are inside the portal, the program would decide that those toys in the portal are in the virtual world, and sends an event to Reaction Module indicating the specific figurines that are in the virtual world. Otherwise, if Understanding Module receives an event about a figurine in the physical world, it sends to the Reaction Module an event indicating the location and figurine.

The Reaction Module is the state machine that determines Sam’s current state. It receives an event from Understanding Module, and decides which state to go to. In the new multiple-toy implementation, once the Reaction Module receives the event about the toys’ existence in the virtual world, it extracts the information about the figurine identities and locations and passes them on to the Generation Module. When the Reaction...
Module receives the event about the figurines’ existence or absence in the physical world while Sam is listening to the user’s stories, it sends animation commands to the animator to generate appropriate eye gaze for Sam. If the event indicates that a figurine has been recently placed at a new location, Sam gazes at the new location. Otherwise, if the event indicates that a figurine previously in the castle is now absent, Sam directs eye gaze toward the user.

The Generation Module generates the specific behaviors for Sam. Specifically, before Sam tells stories, it directs Sam’s eye gaze, allows Sam’s hand(s) to grab the virtual toy(s), and searches through all the stories to find an appropriate story to tell. While Sam is telling stories, it directs the animator to use the animation commands defined in the appropriate files to generate animations. In this new implementation, the Generation Module first determines which toys are in the virtual world based on the information received from the Reaction Module. Before Sam tells stories, Generation Module receives an event from Reaction Module to direct Sam to “get toy.” based on the figurines that are in the virtual world, Generation Module directs Sam’s left and/or right hand to go to the portal, and then shows the appropriate toy in Sam’s hand. When the Generation Module receives an event to direct Sam to tell story, it searches through the array of stories for an appropriate story to tell. First, based on the event from the Reaction Module, it determines the specific room where the user’s last story ended. Then it searches through the list of the stories for an untold story that starts in that specific room. If no such stories exist anymore, the algorithm picks an untold story at random to tell. If all of the available stories have been told to the current user, the algorithm picks any of the stories to tell. During story telling, Sam uses one or both hands to move the toy(s)
around the castle and generates appropriate make-believe behaviors in the virtual world corresponding to the story plot.

5.2.6 Animations

The new animations allow Sam to utilize both toys. Currently each toy is fixed to one of Sam’s hands. For example, the left hand can hold only virtual toy A, while the right hand can hold only virtual toy B. When toy A is placed in the portal, Sam gets the virtual toy A using his right hand (figure 9); when toy B is in the portal, Sam gets the virtual toy B using his right hand (figure 10); finally, if both toys are in the virtual world, Sam uses both hands to play with the toys (figure 11). Thus the figurine that is placed inside the castle determines the animation of the story.

Figure 9. Sam holding virtual toy with right hand
Currently, there are eight new stories that Sam is able to tell, and seven of them involve more than one characters. The stories that involve more than one characters have three sets of animation commands: one set of animation for each toy, and one set involving both toys; on the other hand, the only story that involves only one character has only two sets of animations, including one set of animation involving each toy.
Each animation is synchronized to the corresponding story. Each set of animation commands dictates Sam’s behavior at every millisecond. The animation commands tell the animator the exact time to allow Sam to start speaking, to move an arm, to move the hand, to grab the toy, to look at the user, or to raise eyebrows. When Sam talks about the characters being upstairs, Sam moves the virtual figurines to an upper level room in the virtual castle (figure 12); while Sam exclaims that a child went on top of the roof of the school, Sam moves the virtual figurine above the virtual castle (figure 13); when the story describes the main characters dancing in the living room, Sam moves the two characters up and down and around the room. During story telling, Sam’s gaze follows the figurines most of the time. However, when Sam describes a surprising event or expresses a thought, she focuses her eyes on the user, and sometimes she also moves her eyebrows (figure 14). Sam’s behavior is based on observations from videotapes of children playing with the toys and telling stories. These videos show that when a child tells a story to another child while playing with a toy, she looks at the toy. However, when she raises her voice to tell something very exciting, she looks at the other child. Thus Sam’s behavior is modeled after that of children’s.
Figure 12. Sam moves the toys to a room on upper level

Figure 13. Sam holds the toys above the castle
Figure 14. Sam gazes at the user and raises eyebrows
Chapter 6. Evaluation

Overall, the system worked as expected. When one or two toys are placed in the “magic portal”, Sam correctly recognizes the identity of the toys and acts appropriately by grabbing the correct toys. Sam tells the new stories and plays with the toys according to the identity of the toys in the virtual world. The new stories are told fluently and clearly from the narrator’s perspective, and the animations are well matched with the stories. When the toys are not inside the magic portal, Sam listens attentively by gazing at the location of the most recently moved toy. Figure 15 shows the completed new system.

Figure 15. The new Sam system
There is one minor glitch in the system. Since the toy location sensors sometimes report absence of the toy even when the toy is present, I added a delay in the sensor module software to ensure reliability when detecting the absence of a toy. Specifically, when the sensors detect the location of a toy, the Toy Location Module software immediately reports to Understanding Module about the presence of the toy. However, if the sensor detects absence of a toy, the software queries the reader three more times to confirm that the toy is really absent. This causes a 1 second delay in Sam’s reaction to a disappearing toy. Since the toys’ presence in the virtual world is determined by the presence of the toy in the portal and the closed status of the portal door, the delay does not cause a problem if the user takes at least 1 second to take out a toy from the portal after opening the portal door. However, there is a glitch if the user opens the door, takes out one or two toys, and closes the door immediately within one second. For a brief amount of time, the system mistakenly interprets the toy as still being in the virtual world, thus Sam tries to grab a toy and tell a story. In this case, for a brief moment, a certain toy may appear in the virtual world and the physical world at the same time. One second later, Sam realizes that the toy is gone, and it takes another 1 or 2 seconds before the toy disappears from Sam’s hand. This problem may cause some confusion. However, since the likelihood of a user performing a sequence of actions within one second is small, the system reacts correctly most of the time.

In the future, this problem may be fixed by exploring more reliable sensors. By utilizing better sensors, we can improve the reaction time of Sam and allow Sam to respond more quickly to user’s actions.
Chapter 7. Conclusion and future works

In this thesis, I have examined children’s play patterns and storytelling routines and improved Sam, the virtual companion, to enhance children’s play experience. The new system allows Sam and the child to share multiple figurines, and it allows Sam to tell stories involving multiple characters in different perspectives while acting out the plot with multiple virtual toys. With the new system, children are able to represent multiple characters during storytelling and can take several perspectives while sharing their fantasies. Hence, children’s play experience would be less constrained and more enjoyable.

In the future, Sam and children can share even more toys, if there are a large number of stories in the database such that the stories cover different combinations of characters. The interaction can be further enhanced if Sam and the child can engage in the same pretend play at the same time. While the child tells a story, Sam can play one of the characters and collaborate with the child in the process to make the story more sophisticated and interesting.

There are further improvements that can make Sam an even better companion. For instance, Sam can be more attentive and supportive while the child is telling stories. This would require more sensors to allow Sam to better understand the child’s actions and intentions. For example, accurate detection of the child’s gesture and eye gaze can allow the system to interpret the child’s intention, thus allowing Sam to react more appropriately and act as a more collaborative friend.
Another major improvement can be achieved if Sam were able to partially understand children’s speech. If language recognition can be performed, Sam would be able to engage in better interaction with children. Currently Sam attempts to continue a child’s story by starting a story in the same room where the child finished the last story. If the system were able to spot keywords in the child’s story, Sam would be able to simulate a more smooth continuation of the previous story by matching a story with similar theme.

As Singer & Singer (1990) stated, “Children who are left alone with no encouragement to play, no place to play, and nothing to play with might still play, just because children are inquisitive creatures by nature. But when we add a caring person, a sacred space, and a few props or toys, imagination bursts forth.” The Sam system provides a supportive companion, a shared space, and a couple of figurines to encourage children’s fantasy play and storytelling. In the future, we can further enhance such systems to foster children’s imagination and linguistic skills and to provide them a more memorable and pleasant childhood experience.
References


