An Overview of Computational Co-creative Pretend Play with a Human

Brian Magerko, Justin Permar, Mikhail Jacob, Margeaux Comerford, Justin Smith

Georgia Institute of Technology, Atlanta, USA
{magerko, jpermar, mikhail.jacob, m.e.comerford, jssmith}@gatech.edu

Abstract. This article provides an overview of our ongoing research project to better understand and formally represent the sociocognitive processes employed in sociodramatic pretend play. Pretend play involves the collaborative co-creation and subsequent enactment of a narrative experience using physical objects. This paper describes our empirical study of human pretend play, including preliminary results; the computational representations of those initial findings; and the design of the Co-creative Cognitive Architecture (CoCoA), a set of components that support the design of co-creative intelligent agents.

Keywords: co-creativity · interactive narrative · play · cognition · cognitive architecture

1 Introduction

Although play predates any concept of human culture or society (e.g. animals engage in play as children and adults without any formal cultural context), it is an important part of the human condition within familial and social groups [6]. Play serves to strengthen social ties within groups, increase affect between individuals, and allow meaningful learning and practice at creative problem solving [2]. In other words, *play is a fundamental part of human society*. Therefore, if we aim to build intelligent agents that co-exist in society (real & virtual) with humans, understanding how we can be playful together is an important pursuit for human-centered AI research (i.e., AI research emphasizing interaction between intelligent agents and humans).

In this paper, we focus on sociodramatic pretend play (i.e. *pretend play*) as an exemplar of the co-creative, story-centric processes that are pervasive in play. Pretend play is a social activity that involves constructing and enacting scripts and roles that children draw from life experiences and media sources [4]. Pretend play emphasizes interaction and construction of shared knowledge through negotiation, which makes this domain particularly suited to understanding how computational agents can play with humans.

The crucial research question asked in our work is: *How do we make intelligent agents that can play with people?* It is clear that we play *through* computational systems (e.g., through social networks) and *with* computational systems (e.g., non-player characters in games and entertainment robots). However, play with computational systems occurs today in highly structured and constrained environments and games. Intelligent agents do not yet improvise and negotiate in open-ended activities such as pretend play. That is, agents are currently incapable of playing with humans as peers, where "how" and "what" to play is negotiated in an unstructured environment.

Our current focus is on the sociocognitive capabilities involved in third person pretend play (pretend play where agents control story characters from outside the story world). Our investigation includes both empirical studies and development of a computational architecture, called the Co-creative Cognitive Architecture (CoCoA), which is the focus of this paper. The studies we conduct are designed to elicit a detailed understanding of the sociocognitive processes used during pretend play in sufficient detail to build a computational pretend play agent. Our preliminary analysis suggests that we focus on three aspects of pretend play activities: social interaction required to produce "successful" play (frequent laughter, infrequent pauses, etc.), novel usage of toys (e.g., a monster toy representing the story hero), and negotiation strategies for story co-creation. Additionally, our current work builds on our prior work on improvisation [9], [7] and shared mental models [5]. Our aim is to use CoCoA to build agents that can engage in activities that emphasize social interaction, environmental interaction, improvisation, and creativity. Our longer-term aim is to inform the future development of social agents in our homes, public spaces, workplaces, and virtual worlds. The next section presents our view of the cognitive capabilities required for an agent to engage in third person pretend play.

2 Pretend Play Agent Architecture

Pretend play involves improvisational co-creation of a story that is subsequently enacted by the participants [10]. Our work emphasizes three sociocognitive features of pretend play. The first feature is *establishing a shared mental model* (SMM). A pretend play SMM includes two distinct contexts: a diegetic frame (equivalent to Sawyer's "play frame") and an extradiegetic frame (equivalent to Sawyer's "out of frame" communication and actions) [10]. That is, communication between the participants in the extradiegetic frame ultimately attempts to answer the question "what do you want to play?". The second feature is *creating content for the diegetic frame* [10]. After participants negotiate a shared understanding of a pretend story (e.g., monsters attack) to construct a diegetic frame, they then enact the story. Script theory suggests that participants appropriate and adapt cognitive script content during pretend play [4]. The third feature is *mapping environmental objects to pretend objects*. When acting out events from the diegetic frame, an agent should choose an appropriate environmental object to represent the pretend object. Similarly, when perceiving actions taken by other participants, the agent should constrain the interpretation to those actions appropriate for the mapped pretend object.

These three features of pretend play are the focus of the following CoCoA components: 1) the SMM component, 2) the story content generation component, and 3) the object blending component. CoCoA (see **Fig. 1**) is an abstraction of the architecture that we are developing for our third person pretend play agent in the Soar cognitive architecture [8]. Soar is a symbolic, rule-based theory of cognition. Soar provides an input link for perception, a working memory to store symbolic representations, and a long-term memory to store production rules. The effects of production rules update working memory, and may update the output link to affect the environment (via actuators). CoCoA consists of components for *perception*, *SMM negotiation*, *interpretation*, *object blending, story content generation*, *action selection*, and *enaction*.

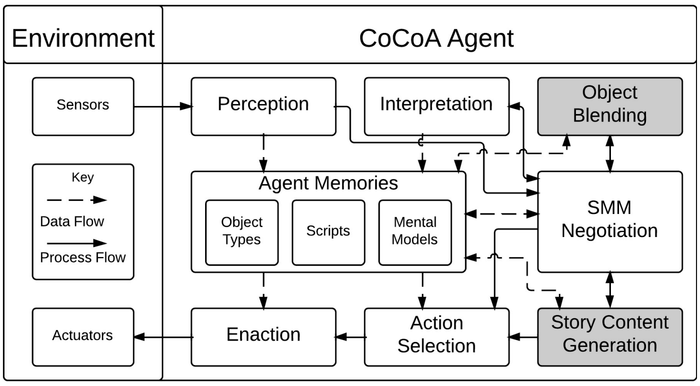

Fig. 1: Co-creative Cognitive Architecture (CoCoA)

Sensor input is a symbolic representation of environmental state, which is primarily object identity and movement over time . For example, a typical input state representation may describe a “Godzilla doll and a set of blocks on the playmat”. Sensor input is processed by the perception component. Thereafter, the SMM negotiation component updates the agent’s SMM to reflect actions taken by other agents and environmental changes. Any inferencing processes the agent performs in order to facilitate sense making are in the interpretation component. The agent uses the story content generator to update story content during SMM negotiation. Similarly, the agent uses the object blending component to map between environmental and pretend objects. When the story content appears to have “settled” through negotiation, the agent subsequently selects and enacts story actions.

Due to our view that SMM construction is a central process of pretend play, our architecture prioritizes decision-making in the SMM negotiation component to pursue the agent’s high-level goal: negotiation of a diegetic frame. The component maintains information about potential *cognitive divergences* (differences between the beliefs or mental models of collaborating agents) and tries to *repair* them, using various repair strategies [5]. Repairs facilitate achieving *cognitive convergence* (the process of correcting differences in beliefs among collaborating agents) and may end in a state of *cognitive consensus* that is characterized by the establishment of a single shared mental model among the collaborating agents. This module is effectively high-level control for our pretend play agent.

We are testing our third person pretend play agent in a virtual environment consisting of a playmat and a diverse set of toys. A human interacts with the virtual environment by using a Kinect to grab, rotate and move objects. A microphone is used to recognize a small set of verbal commands and non-language speech utterances. In the virtual environment, the human and virtual agent play pretend by taking alternating turns. We are also investigating naturalistic turn taking behaviors, the social dynamics in human pretend play, and the application of our work to a real-world context by building an embodied pretend play agent. We are integrating CoCoA with a model of turn taking (i.e., fluent exchange of communication or play turns over bottlenecked resources, such as objects or the speaking floor) called CADENCE (Control Architecture for the Dynamics of Embodied Natural Coordination and Engagement) [3]. This integrated system will enable a humanoid robot to perform naturalistic object play with a human. The next two sub-sections provide detail on the object blending and story content generation components.

2.1 Object Blending

Third person pretend play agents use environmental objects during story co-creation and enactment. This interaction requires mapping between the environmental objects and pretend objects both pro-actively to perform diegetic actions and also reactively to interpret movements of environmental objects. For example, in order to pretend that a tank is firing at soldiers when the environment includes a colored brick and toy soldiers, the agent must map between the environmental and pretend objects and pick an environmental object to represent the tank (i.e., the colored brick). After the mapping is established, the physical qualities of the toy (the colored brick) constrain the agent's subsequent use of the toy – the agent can neither physically climb inside and drive it nor actually fire anything from it. Instead he can move it around with his hands and make sounds as if firing from it. In order to do the mapping process and use physical constraints, we use a process of *object blending* [11] to create objects that contain a blend of properties of the pretend and environmental objects. In the previous example, the colored brick would be blended with a pretend tank in order to create a new blended object with the size and shape of a colored brick but with actions taken by a tank. Note that the actions performed with the toy in the environment therefore have consequences for the possible actions that can be defined in the diegetic frame. For example, when an agent holds the blended "tank-brick" object and makes "boom boom" (shells firing) sounds, the effect of that action is to impact the target (e.g., injure the pretend soldiers being fired at). Object blending is dependent upon knowledge of the properties of various objects present in both the extradiegetic and diegetic frames.

Object blending consists of three sub-processes: 1) *categorization*, 2) *comparison* and 3) *property combination*. Categorization maps a specific instance of an object to the appropriate object category. This allows the agent to retrieve from memory a set of properties attached to the category as well as the sensed perceptual properties of an object instance. Once the agent retrieves the object category and category properties from memory, it proceeds to perform a comparison of that category with other categories using those properties. The goal of comparison is to find the closest matching object. Finally, the agent selectively combines properties of the environmental (colored brick) and pretend (tank) objects in order to create the blended environmental-pretend (tank-brick) object. After the mapping is established, the environmental object is manipulated to perform extradiegetic actions that correspond to actions in the diegetic frame. For example, the agent can "fire the tank's cannon" by 1) moving the brick, 2) playing sounds that resemble tank shots, and 3) knocking down toy soldiers.

2.2 Story Content Generation

The story content generation problem is to create an activity consistent with the agent's shared mental model that is both novel for the agent (p-creative [1]) and interesting for the human participant. As noted above, script theory assumes that content in the diegetic frame is co-created from existing script knowledge of familiar cultural and social activities. In our system, a minimal diegetic frame contains two actors, each with a defined goal, a setting, and a set of actions and objects that can achieve each character's goal. Because the diegetic frame is negotiated, it is unlikely that an agent's existing script knowledge can be used without modification. For example, the human pretend play participant may perform an action that is not in the agent's script, due to a differing script or change of activities. Additionally, object blending may constrain the set of possible objects in the diegetic frame. In these situations, the agent needs to update its SMM and generate story content to match the diegetic frame.

In addition to external influences, our agent has an explicit goal of engaging in creative activities [1]. We can measure novelty qualitatively as a difference from existing script knowledge. A novel activity can differ from a script by a novel setting, actor, goals, actions, or objects. Our planned approach is to generate qualitatively interesting activities using both a source and target activity as inputs.. Our assumption is that the (human) pretend play participant does not want to repeat a previously experienced activity. Consequently, our view is that creating novel activity content is useful (and thus creative) if the activity is both causally coherent and interesting. Identifying specific factors that can be used to generate interesting activities remains future work.

3 Future Work

The computational play project is a work in progress. Our future work includes developing and evaluating CoCoA and evaluating our work.

There are two features of negotiation in third person pretend play not present in our prior work in improvisational theater. The first feature is embodiment during enactment of story content, such as the inability to physically enact specific diegetic actions. The second feature is the possibility of disagreement between pretend play participants. In improvisation, participants are strongly encouraged to accept the offers of other participants (indeed, our SMM negotiation model from our prior work assumed a default policy of accepting offers).

In order to evaluate our work, we will conduct a series of ablative studies to test the effectiveness of individual CoCoA components and also conduct human evaluation of the agent's interaction. For example, we plan to measure the human's *affect* experienced, *fluency* of interaction, *and balance of control* during the play session. We plan to design and use additional evaluation metrics, such as *human playfulness* and *perceived robot creativity*.

The long-term future of this work includes investigating additional features of social activities featuring co-creative, improvisational behavior. We anticipate that behaviors such as social rule construction (within both the diegetic and extradiegetic frames) and affect-centered decision making are highly involved in more complex versions of computational play than are reflected here. Additionally, this work is likely applicable to developing play therapies for autistic children. This work will continue to push on the three fronts of 1) better understanding human play, 2) creating computational representations of our findings, and 3) exploring new play domains within the context of a virtual environment and robotic playmate.

Acknowledgements. This work is supported in part by NSF IIS grant #1320520. Thanks go to the following who have contributed to this research, including Jordan Ashworth, Nicholas Davis, Eric Fruchter, Matthew Guzdial, Shan Li and Alex Zook.

4 References

1. Boden, M.A.: Computer models of creativity. Handbook of creativity. pp. 351 – 372 Cambridge University Press (1999).
2. Caillois, R., Barash, M.: Man, play, and games / Roger Caillois ; translated from the French by Meyer Barash. Urbana : University of Illinois Press, 2001. (2001).
3. Chao, C., Thomaz, A.: Controlling Social Dynamics with a Parametrized Model of Floor Regulation. Journal of Human-Robot Interaction. 2, 1, 4–29.
4. Fromberg, D.P.: A Review of Research on Play. (1999).
5. Fuller, D., Magerko, B.: Shared Mental Models in Improviational Theatre. Proceedings of 8th ACM Conference on Creativity and Cognition, Atlanta, GA (2011).
6. Huizinga, J.: Homo Ludens: A Study of the Play-Element in Culture. Taylor & Francis (2003).
7. Mikhail Jacob et al.: Viewpoints AI. Presented at the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (2013).
8. Lehman, J.F. et al.: A gentle introduction to Soar, an architecture for human cognition. Invitation to cognitive science. 4, 212–249 (1996).
9. Magerko, B. et al.: An Empirical Study of Cognition and Theatrical Improvisation. Proceeding of the Seventh ACM Conference on Creativity and Cognition. pp. 117–126 (2009).
10. Sawyer, R.: Pretend play as improvisation : conversation in the preschool classroom. L. Erlbaum Associates, Mahwah N.J. (1997).
11. Zook, A. et al.: Formally Modeling Pretend Object Play. Proceedings of the 8th ACM Conference on Creativity and Cognition. (2011).