# Bottoms Up: Improvisational Micro-Agents

Brian Magerko
Georgia Institute of Technology
School of Literature, Communication, and Culture
686 Cherry St., Atlanta, GA
magerko@gatech.edu

Casey Fiesler
Georgia Institute of Technology
GVU Center, College of Computing
85 5th Street NW, Atlanta, GA 30332
casey.fiesler@gatech.edu

Allan Baumer
Georgia Institute of Technology
School of Literature, Communication, and Culture
686 Cherry St., Atlanta, GA 30332
allanbaumer0@gmail.com

Daniel Fuller
Georgia Institute of Technology
School of Literature, Communication, and Culture
686 Cherry St., Atlanta, GA 30332
dfuller3@gatech.edu

## ABSTRACT
This paper describes our current approach in implementing computational improvisational *micro-agents*, agents that perform one specific aspect of our findings from the *Digital Improv Project*. This approach is intended to foster bottom-up research to better understand how to build more complex agent behaviors in a theatrical improvisational setting. The *Digital Improv Project* is a multi-year study at the Georgia Institute of Technology focused on studying real life theatrical improvisers with an aim towards better understanding the cognition employed in improvisation at the individual and group level.

## Categories and Subject Descriptors
I.2.1 [Artificial Intelligence: Applications and Expert Systems]; J.5 [Arts and the Humanities]

## General Terms
Algorithms, Design

## Keywords
Improvisation, synthetic characters, cognitive science

## 1. INTRODUCTION
Improvisational agents have been of interest to the interactive narrative community off and on for decades. For example, the Computer-Animated Improvisational Theater (CAIT) was an interactive theater system that allowed children to control avatars in a virtual world in which intelligent animated agents improvised playtime activities such as playing, fighting, and singing [1][2]. The intelligent animated agents reactively reproduced the activities one would expect in an improvised children's play space. The agents selected behaviors reactively by decomposing tasks using a broad but shallow approach. Another example is the Improv system [3], in which virtual animated avatars can be scripted to enact a scenario. The Improv system emphasizes variability at the surface level of the presentation—the exact positioning, movements, and gestures of avatars in a virtual graphical environment by introducing noise [4] to produce natural-looking variability. Hayes-Roth and van Gent combined the Improv system and CAIT to produce a non-interactive scenario about a master and servant that can play out three different ways depending on the setting of personality traits for the master and servant roles [2].

The Mobile Robot Improv troupe at the Carnegie Mellon Robotics Institute [5] focused on creating believable agents that displayed emotionally motivated behavior within the context of a narrative experience. The performance relied on Knight's deconstruction of a dramatic performance [6], which includes defining *hero* and *villain* archetypes, *inner* and *outer obstacles* that prevent the hero from achieving his or her goal, and the *given circumstances* that influence how a character attempts to achieve their goal. These are all examples of previous attempts to model improvisational behavior within a theatrical setting.

These approaches to computational improvisation have assumed that agents are either autonomous (they improvise according to their own goals and beliefs) or semi-autonomous (they can receive direction from another agent, such as a human or story director agent) while improvising. Mateas and Stern have argued against what they called *strongly autonomous* agents, noting that they are difficult to coordinate, and opted for *weakly autonomous* agents in Façade [7]. While the commitment to weakly autonomous agents was a strength of Façade—it allowed for a tight coupling between agent behaviors and story goals—it also forced the agents to use dialogue-based "catch-alls" in an attempt to subtly deal with unexpected user inputs and keep the story moving (e.g., laughing uncomfortably and saying what a "kidder" the user is, then moving on). However, improvisation ("improv") in an interactive narrative might be required of synthetic characters when unexpected situations arise (such as complex world physics that are hard to model or player actions that are not covered in the authored story space [8]) or when relying on dialogue catch-alls is perceived as undesirable.


*INT3 2010*, June 18, Monterey, CA, USA

There are therefore multiple conditions when an interactive agent may require improvisational behavior:

*Story space breached by user*: The user has executed a series of actions that has led to a world state not covered by authored story content. This could mean anything from physically altering the environment or a character (e.g., the canonical example of shooting an important character) to being in an unexpected social situation or conversation.

*Story space breached by environment*: Some series of events in a dynamic environment has led to a world state not covered by authored content.

*Story generation recovery*: An interactive story that has been generated (e.g., by a planner) cannot currently replan given a story state breach. An improvisational agent could keep the story goals in mind while improvising and keep *a story* going, even if not the one explicitly pre-authored. In the case of an educational application, an improvisational agent may be able to keep the desired pedagogical goals in mind while improvising a story, even though the initial dramatic goals can no longer be fulfilled.

*Improvisational theatre*: If authors want to create an improvisational theatre experience in computational fashion (e.g., [5]) it is unclear what approach would be more appropriate than to create improvisational agents for the performance.

It is with these situations in mind that the *Digital Improv Project* has sought to better understand human improvisation with the goal of creating improvisational agents. However, our work on studying people (described in section 2.2) has yielded a large and complex corpus of data. Beyond our immediate goal to determine "what people do" when they improvise, our eventual goal, building an improvisational agent based on this data, is a daunting task given this wealth of data. This paper presents our current approach in the computational modeling of improvisation, which seeks to first understand individual aspects of improvisation through the creation of *micro-agents* that represent singular aspects of improvisation (described in Section 3). The intention behind this methodology is that by developing simpler agents in different environments and situations, we can reach a better understanding of the issues involved in building a more complex improvisational agent. We anticipate this development process to bring up, and hopefully address, questions in knowledge representation, interaction modalities with improvisational agents, and the essential question of how to represent our findings on improvisation in cognition within a computational framework. Section 4 presents our current agent designs, the issues they have brought up, and our potential solutions to those issues.

## 2. COGNITION AND IMPROVISATION

There have been only a handful of studies that have taken a serious look at the cognitive processes involved in human improvisational performance within artistic domains, and it is this existing body of research that we looked to in designing our study in the specific domain of improvisational actors.

### 2.1 Related Work

The majority of this previous work exists within the domain of music. Mendonça and Wallace's study of jazz performers yielded a framework based upon very specific cognitive processes—temporal cognition and creative cognition—and though the study involved duos of musicians, it touched briefly upon the possible significance of collaboration [9]. Seddon's and Reinholdsson's investigations involving groups of jazz musicians focused more on this collaborative aspect, describing the communication methods and social interactions between the improvisers, though Seddon noted that musicians are likely employing subconscious cognitive processes that do not produce observable behavior [10][11].

More computational approaches to the subject include Johnson-Laird's proposal of a principle of algorithmic demands governing improvisation, examining the role of working memory in the task [12]. Other work has focused on the role of knowledge and experience, such as how improvisers draw upon motifs during performance and how skill develops with practice [13]. Ramalho's AI model of an intelligent jazz performer focuses on two specific aspects of musical improvisation: creating a "musical memory" of previous played melodic fragments, and reusing them during live performances [14].

Outside of music, there have also been discussions of the nature of improvisation in areas such as art and dance [15][16], as well as non-artistic domains such as management and engineering [17][18][19]. However, our research is the first large-scale effort to study human cognition within the domain of theatrical improvisation.

### 2.2 An Empirical Study of Theatrical Improvisation

Theatrical improvisation is of particular interest since improvisation in acting has been used as the motivation for research on believable agents [2][3][20]. This work has been based on specific improvisation teachings or concepts (such as character “status” in a scene) without a more granular understanding of what the actors attend to on stage. Therefore, it does not take into account (because it was unknown at the time of the work) the knowledge that goes into improvisers making decisions, how they communicate their decisions to other actors, and how that communication is, in turn, received by others. The *Digital Improv Project* seeks to gain such an understanding with the hope that more robust and useful improvisational agents can be built for use in computer game and interactive narrative worlds.

Our approach to understanding improvisation and cognition has followed a four-phase research plan. The initial phase was to formulate how to best study human improvisers to get the most useful data for building agents. We eschewed any overly fine-grained studies, such as using eye-tracking devices for measuring gaze, and instead opted for an approach that would yield the kind of data most relevant to building agents (e.g., heuristics for decision making, knowledge and goals used, theory of mind, etc.). We decided on a mixed design that incorporated the process of collecting retrospective protocols from actors after they were engaged in an improv scene, followed by an unstructured group interview [21]. This combination allowed us to get the kinds of decision-making knowledge we desired, along with information about group dynamics, misunderstandings, etc.

The second phase was data collection. This involved recruiting members of different troupes across the Atlanta, GA area for involvement in our study. We intentionally selected improvisers that represented novice, intermediate, and expert levels of proficiency so we could more clearly see the differences between expert and non-expert improvisation. 27 improvisers were used in our study from four different local troupes. Each study was conducted in a controlled setting at Georgia Tech's campus. We picked three different improv games to have them play that we determined would yield the best data related to building agents:

- *Party Quirks:* This game involves four players, a party host and three guests. Each of the guests is given a character "quirk" and it is the job of the host to guess these throughout the course of the scene. Our participants' quirks included character traits such as "video game addict," "obsessed with health issues," and "someone who can fly." This game was chosen in order to illuminate group dynamics and how the improvisers communicate knowledge.
- *Film and Theater Styles:* This game is typically for two actors who are given a specific scene and told to perform in different film or theater styles that change throughout the performance. For example, the context of the scene might be "plumbers on a hiking trip" and the actors must switch between "horror film," "cartoon," and "sports drama." This game was chosen to study semantic knowledge and narrative development.
- *Game:* This game has the point simply of creating a scene in accord with some suggestions and constraints. We used three different versions—high, low, and no constraints—in order to show the effect of content constraints on scenes and to illuminate aspects of narrative development.

We concluded with the main data collection phase when we had reached adequate coverage across each game with each skill level.

The third phase of our study is analyzing the large corpus of data we have collected. Our study has generated over seventy hours of performance and interview footage, which has to be cataloged and coded for the various aspects of improvisation and cognitive we are attempting to theorize about and model. This involved an iterative process of generating an initial coding scheme based on our findings, coding example data, and then refining our coding based on the experiences coding the example data. This analysis has led to a better understanding of narrative development in agents [22] and the process of cognitive convergence, which is described in Section 2.2. Our initial analysis has led us to conduct another study of improvisers to fill in the noticeable gaps in our current data set. This gap was created by our reliance on self-report in retrospective protocol collections rather than allowing interviewers to ask pointed questions. The rationale was that we did not at the time have the necessary knowledge of the domain to target our inquiry. Now that we are much more aware of what improvisers actually do, we were able to revisit them with a more structured approach in hand. In the follow-up study, we interviewed improvisers on-site after one of their regular performances.

The fourth and final phase of the *Digital Improv Project* is the creation of synthetic characters that computationally represent our findings. The final result is intended to be a complete representation of our findings (see [23] for a summary of our current findings). However, we contend that immediately jumping into building such a complex agent design without a full understanding of the domain, how to computationally represent that domain, or even the implementation issues involved (e.g. how the agents interact with each other or a user), seems premature. We instead have opted for an initial, bottom-up approach to creating improvisational AI agents, called "improv micro-agents," which is described below in Section 3.

# 3. IMPROV MICRO-AGENTS

This section presents our ongoing work on creating micro-agents to address different aspects of cognition as found in our study of theatrical improvisers. The systems reported are in various degrees of completion, though we expect to be able to report the finished works by Spring 2010.

## 3.1 Character Generation

How to portray a character is one of the most important decisions that an improviser must make when beginning a scene. Although improvisers are often provided with basic information about their characters in a scene (e.g., improvisers are often given content constraints related to character to begin a scene, such as the character's occupation or relationship to another character on stage), they still must choose which trait and behaviors associated with that character to communicate. In other words, improvisers often need to decide how to portray a basic character they have been given or are assuming.

We have discovered different heuristics that improvisers use for making character choices, gleaned from our data gathered from performances and interviews with improvisers. Often, improvisers begin with familiar concepts—following the theory of schemata from cognitive science, in which past experiences are summarized into composite knowledge representations of an aspect of the world [24]. For example, during a game of *Party Quirks*, one of our participants who was given the quirk of a "video game addict" reported his decision to act like a "ridiculous caricature" of that character type based on what he knew to be a stereotype of antisocial, obsessive behavior. He used his own internal "video game addict" schema to pull out the most typical features, essentially creating a prototype (the average values inside a schema) in order to clearly communicate this character type to the other actors and the audience. Another participant, when given the trait that he could fly, also considered his schema, determined the most typical values, and then intentionally took an atypical twist on them, portraying his character as someone who could fly but uncontrollably so. He later reported that he thought this would be more "interesting" than sticking to the "Superman" stereotype of flying.

These two examples illustrate a phenomenon that we saw frequently from improvisers in their character choices. By utilizing stereotypes as a starting point, they are taking advantage of a common organizational scheme for categories of knowledge, in which that knowledge is organized around prototypes made up of a bundle of average values [24]. Therefore, when considering a "bird," one tends to think "has wings," "has beak," "can fly," etc., so that the image that comes to mind more closely resembles a robin than an ostrich. *Prototype effects* come from the tendency of people to judge certain members of categories as being more representative of a category than other members—a robin in the category of "bird," or as in our example, an antisocial, obsessive person in the category of "video game addict" [25]. We saw among the improvisers two common ways of using these most representative examples from a character schema; they would either stick closely to the prototype (often to make their performance more obvious for the other actors) or take a significant step away from it (i.e., giving it a “twist”) in order to make the performance more interesting or humorous. One improv troupe even trains in their classes to build a typical schema for a character and then add a "twist" to make it more interesting. Furthermore, we saw that when improvisers make the decision to incorporate a "twist," they tend to construct them in one of these three basic ways:

*Opposing*. In this case, an improviser takes a typical value or behavior for a schema and replaces it with an opposing idea—such as the man who flies uncontrollably. The prototype for a person who can fly is that of a Superman-type character in control of his actions. The improviser here extracted that trait and turned it on its head so that the character he portrayed was out-of-control and scared of his own abilities.

*Borrowing*. Though similar to opposing, this technique involves a step farther away from the schema by borrowing behaviors from other schema, and thereby allowing improvisers to construct characters more complex than would be possible if relying solely on stereotypes. This phenomenon is the only concept outside of our performance data that we have considered modeling thus far. We have observed an example from a popular improv show where a character was imbued with the traits “mosquito” and “gets drunk on blood” in a performance. His performance was anchored in the “mosquito” concept, but combined that prototype with elements from the “drunk” prototype, including acting silly, getting sick, and flirting with the party host. This behavior points to the cognitive phenomena of *conceptual blending*, which appears to be a strong aspect of concept generation in improvisation [26].

*Caricature*. As in the case of the video game addict, an improviser might take the stereotype to its extreme. Though less of a "twist" than choosing atypical characteristics, a strong exaggeration can be an interesting character choice.

Improvisers can use these techniques to make character choices, with a desire to be more "interesting" determining the degree of divergence from a prototype. This is an important aspect of improvisation, one that relates to both basic cognitive concepts such as knowledge representation, as well as concepts of character development in narrative. We have applied our micro-agent technique to focus on this essential component of character development.

In modeling this behavior, our goal was to create a structure of schemata for our AI improviser while providing it with the freedom to make choices regarding which traits to portray for a given type of character. A feature of prototypes is that they do not necessarily have rigid boundaries; a graded category such as "tall man" instead has *fuzzy* boundaries, with prototype effects resulting from degrees of category membership [27]. Whether than asking whether something *is a* bird, where determination is a binary true or false, we can instead recognize the fuzzy boundaries of categories and ask how similar a particular instantiation of a bird is to our ideal model—a judgment of "goodness" of category membership [25]. Therefore, in determining how to represent these values computationally, we borrowed from the domain of fuzzy logic, which though already used extensively in the field of science and engineering, is gaining a foothold in the field of game development as well [28]. As with prototypes without rigid boundaries, fuzzy logic posits that concepts need not be limited to a binary true or false but can rather have degrees of truth—or degrees of membership in a set. Just as in the real world classes do not have precisely defined memberships (such as the ostrich that cannot fly but is still a bird), the character types in our agent's knowledge structure are represented as fuzzy sets, or classes with continuum grades of membership [29].

To demonstrate the defined problem of how an improviser remains closer to or deviates from a prototype, our application asks the user to provide our AI improviser with two things: a basic character type and a value along a slider of "interestingness." The AI agent therefore knows what type of character it needs to portray and whether it should choose traits based upon a goal of being more transparent (sticking more closely to the prototype) or more interesting (taking a twist on that schema).

Our agent's knowledge is represented as a grid of character classes and attributes (traits or behaviors to portray). There is a fuzzy value for each attribute as it relates to each character type, representing a degree of membership between 0 and 1 (see Table 1). Consider your mental schema for a king or a witch; your prototype of a king likely involves his wearing a crown, whereas a witch cackles. In our knowledge representation, a king has a fuzzy value of 1 for "wears crown" but only .1 for "cackles," whereas a witch has a value of .9 for "cackles." Though a typical king wears a crown, it would be an interesting character choice for that king to cackle—perhaps an evil king. This is an example of the "borrowing" technique for taking a twist on a schema. For the other techniques, an "opposing" choice might have a king refusing to wear a crown at all, and a "caricature" choice could result in a king wearing a comically large crown.

**Table 1. Fuzzy set grid for agent knowledge representation.**

| | **A** | **B** | **C** | **D** |
|---|---|---|---|---|
| | | Wears Crown | Looks in Mirror | Cackles |
| **1** | KING | 1 | .3 | .1 |
| **2** | PRINCESS | .9 | .9 | .05 |
| **3** | JESTER | .1 | .1 | .05 |
| **4** | WITCH | .05 | .8 | .9 |

The portrayal of an attribute (or *defuzzification* in fuzzy logic terms) depends on the function associated with that attribute in terms of how the degree-of-membership (DOM) value maps onto the real world. For instance, a simple step function may be used to dictate that, for “rides broom,” any value above .6 is animated as the agent riding a broom, and for anything below that, the agent has no broom at all.

With respect to the algorithms for the agent's decisions, the borrowing technique is the most complex of the three as it involves combing concepts from multiple prototypes. Once the agent knows both a character type and a value along the interestingness scale, its choices are based on similarity measures, computed for each attribute for each character as compared to the DOM fuzzy value for that attribute and the chosen character. The interestingness value therefore becomes a threshold for determining how dissimilar a given attribute value can be from the chosen character's prototype. For example, given the character of a king to portray, the agent knows that the princess value for "wears crown" (.9) is much closer to that of a king than the jester value (.1) or the witch value (.05). The agent then chooses a trait to borrow from those that have a DOM within a window surrounding the interestingness value; therefore, a performance of a king with very low interestingness might result in borrowing the .9 "wears crown" value whereas a high interestingness might result in a .1 or .05 "wears crown" value. And though the improviser only borrows one attribute, its performance is rounded out by other attributes chosen from that character type's set of fuzzy values, also tied to the interestingness scale. The higher that slider value, the bigger the window of traits that can be chosen. A low interestingness would result in a king only being able to perform behaviors such as "wears crown" (or other typical king behaviors such as riding a horse or sitting on a throne) whereas a high interestingness would result in a king choosing from a wider range, perhaps even getting to "cackles"—but still utilizing the king value for that attribute. This makes intuitive sense, as making interesting character choices opens up a much wider range of possibilities, whereas sticking closer to a stereotype leaves an improviser more constrained.

The opposing and caricature techniques are similar in that they take an attribute with a high DOM associated with the given character type (such as "wears crown" for a king) and alter that value, either lowering it (opposing) or raising it (caricature)—resulting in either a king who does not wear a crown or a king who wears a comically large crown. Additional behaviors are then chosen in the same way as outlined above, taking into account the window of interestingness.

Of course, once the agent makes these choices, the fuzzy values must then be interpreted for agent animations—i.e., what does it mean for a king to cackle at a value of .1? The process of mapping a fuzzy set into crisp values is known as *defuzzification* [30]. For the purposes of these micro-agents, we are using two different methods for translating degree-of-membership values to external actions: (1) linear, in which using procedural animations, the intensity of an action linearly follows the fuzzy value (e.g., a crown gets larger and more bejeweled as the value increases from 0 to 1); and (2) step functions, in which using static animations, the animation selection is based on thresholds (e.g., a value for "cackle" above .7 results in evil cackling a value below .4 results in merry laughter, and a value in the middle results in lackluster chuckling).

This micro-agent, therefore, demonstrates the capacity of improvisational actors to make creative choices within the common constraint of being given a particular type of character to portray. In this model of an isolated phenomenon, the user controls this degree of interestingness; however, our research has provided insights into when improvisers choose to be more transparent or more interesting, and thus in our future work into more well-rounded improvisational agents, the agent itself making this choice will be an important part of the AI.

## 3.2 Cognitive Consensus

When an improviser does not know what another is thinking when working within a scene, *cognitive divergence* occurs [23][31][32]. If an improviser attempts to correct the divergence, they engage in the process of *cognitive convergence*. Cognitive convergence is a multi-step process of attempting to reconcile the two mental states in question (i.e., the mental states of the two or more improvisers or even an improviser and the audience). The process of cognitive convergence takes place in three phases: *observation* (recognition by one agent that a divergence exists), *repair* (that agent trying to change the mental state of themselves or another agent), and *acceptance* (resolution with the repair either succeeding or failing) [33][34]. When the two agents appear to understand each other (i.e., when they are "on the same page"), *cognitive consensus* has been achieved.

This micro-agent represents the process of cognitive convergence on a basic level, by focusing on the aspect of guessing in the game *Party Quirks* (described in Section 2.2). This game is an example of a larger class of improv games which we call *knowledge disparity games*. Knowledge disparity games directly involve the process of cognitive convergence because the crux of these games is that some actors have information about the scene from the onset that is unknown to other actors in the scene. Knowledge disparity games tend to focus explicitly on the process of actors achieving cognitive consensus (i.e., the actors who do not know the privileged information going through the process of guessing/learning it).

Our *Party Quirks* implementation focuses specifically on the process of one agent guessing at another agent's quirk with clues given that are ambiguous in nature. In our implementation, Agent A (the guest) approaches Agent B (the host) and attempts to communicate his *mental model* (i.e., what he is thinking). Agent B attempts to guess the other agent's quirk and therefore goes through the process of cognitive convergence until it fully understands the mental model of Agent B. Bypassing the complexities of natural language processing, this micro-agent approach involves communication via animations that portray quirks. For instance, a "ninja" might disappear and become invisible or a "pirate" might drink a bottle of rum. As stated earlier, how that animation is portrayed (defuzzified) depends on the DOM for the agent in the given attribute. The intent behind using this language abstraction is to explore using non-linguistic symbols (i.e. behaviors) that have multiple possible meanings attached to them in the micro-world. Through the creation of these micro-agents, we are exploring multiple ways of interacting with them from both the user perspective as well as from the agent perspective. We are currently avoiding the introduction of full natural language interaction due to the scope and myriad issues it involves.

During this symbolic discourse, Agent A enacts an animation of varying relevance to its quirk. Agent B reasons about what it thinks the most likely interpretation of what is happening in the scene, and communicates its guess to Agent A. Agent A will either *acknowledge* this as correct, or attempt to *repair* any misunderstandings by enacting a new animation that is more relevant. The process continues until the mental models of A and B are the same. For example, Agent A, who has the quirk of "robot," might give the hint of "performing calculations," which is fairly broad and might relate to other prototypes that have a high DOM for that attribute (a scientist, for instance, might have a DOM of .9). Agent B might guess that Agent A is a "scientist." Agent B then tries to be more specific by giving the hint of "plugging in to recharge." Given the lack of any other prototypes that have high DOM values in both attributes, Agent A then concludes from these two bits of information that Agent B's quirk might be "robot." After this new guess, Agent A acknowledges it as correct and cognitive consensus is achieved.

Using our fuzzy logic framework, this example can easily represent more complex behaviors we see in improv games. For instance, *reverse scaffolding* is in which an improviser intentionally portrays weakly associated attributes for a character in the beginning of a scene and stronger ones as the scene progresses, with the intention of slowly cluing in the others on stage. Our use of fuzzy logic allows us to encapsulate both how improvisers may choose to portray themselves as well as how others may interpret it.

## 3.3 Interesting Conflicts and Reincorporation

The ultimate goal of this third micro-agent project is to develop a partial-order planning system that employs an "interestingness" adversarial search manipulating reincorporation in an evolving virtual environment. Agents within the world can take control of objects and attempt to accomplish various goals. In a traditional adversarial search planning method one agent would simply oppose the other and attempt to prevent the other from achieving their goal. Roberts et al., proposed a narrative model in which each agent would encounter problems in accomplishing their goals which does not necessarily foil their goals, but instead produces a result/interaction other than what they intended. This would "maximize interestingness" and help to develop a satisfying narrative [35].

Our system is based on three narrative methods employed by improvisational theatre performers in our data. They are *offers, yes, and…*, and *reincorporation*. *Offers* are the basic building blocks of improv [22]. They introduce an element (or elements) to the narrative onstage that can potentially be used to help the scene progress (e.g., asking another improviser if they remembered to bring some important object, or an offstage player making a sound like a phone is ringing on stage for an on stage improviser to hear). *Yes, and…* is a method of taking an offer and constructively building off of it—for example, an improviser answering the phone (acceptance) and starting a shocking conversation with their boss, who is firing them for not being at work (addition). *Reincorporation* is when an improviser refers to some object, character, or occurrence from earlier in the scene and introduces it to the scene again (often giving it pertinence or a comedic effect).

We saw these phenomena consistently in our experimental data. An example of an offer (and its acceptance) began with D1 and D2 (names are removed for anonymity) discussing the free trade muffins they were eating and how good it felt to be doing the "right thing." D3 entered the scene with a pantomimed tray of muffins and her head down. D1 rolled his eyes as D3 said, "Mr. Coffeeman, I have more muffins for you from homeland." D3 explained in her retrospective that she wanted to establish the muffins as "anything but fair trade" and introduced herself as a "low-status" character. During the scene, D1 picked up on this and in his interview said, "D3 comes in with this great offer to me that we're ostensibly caring and politically aware, but actually in truth we're subjugating people still." He accepted her offer by rolling his eyes and treating her like an inferior (or low-status character). He further built upon her offer (i.e., *yes, and…*) by describing her as "annoying" to D2 and then talking as if she were not in the room. However, in the climactic moment of the scene, it was revealed that D3 had been putting drugs in the muffins so that she could steal company secrets. At the outset of the scene, D1 had pantomimed raking leaves, which was quickly abandoned in deference to other plot events that arose. Therefore, when he heard what D3 had done, D1 exclaimed, "No wonder I was raking leaves in the break room!" thus reincorporating that idea from earlier in the scene.

In order to model this kind of behavior, we begin with agents who contain a library of plans to achieve certain goals. For example:

A cat is in a tree and two agents want to get that cat out of the tree. Working towards the final goal, one of the agents might work backwards from the final goal of "Get Cat out of Tree." Before this can be done, the agent has to grab the cat. Before that action can be performed, the agent has to be at the same height as the cat. In a plan where the denotation of "-->" means a transition to the next item in the sequence, the plan for the agents would be:

Action: Be at same height as Cat --> Action: Grab cat --> Goal: Get Cat out of Tree.

The interestingness adversarial search would allow difficulties to be generated for interrupting this flow. In the improv scene mentioned earlier, the introduction of a low-status character that overturns the supposed morality of the other characters is something that does not end the scene, but it directly opposes what had previously occurred. For example, completing the first step of being at the same height would be as follows:

Action: Get Object --> Action: Use Object to attain height --> Goal: Be at same height as Cat

The interestingness adversarial search would create disruptions in this plan to introduce conflict into the narrative. The object (such as a ladder) might be too short or broken. In a traditional adversarial search, this foil would try to defeat the agent, preventing it from ever achieving the goal. In this situation it presents a surmountable difficulty to generate a more satisfying narrative. The other agent would then introduce (or offer) a solution to the immediate problem. This could be fixing a broken object or a different methodology (such as trying to lure the cat down or using a pogo stick instead of a ladder).

After a certain number of iterations, the agents' plans will be allowed to succeed in order to end the scene. We plan to set up an algorithm with a higher chance of previously used items or actions being reincorporated as the scene progresses (such as a broken ladder being used to get a fish as food to lure the cat). This would allow elements introduced earlier in the narrative to be linked to later elements, creating a cohesive whole. This should enable us to design computational agents who (at least seem able to) reason about, perform, and adapt narrative structures ad hoc.

# 4. DISCUSSION

As mentioned previously, the efforts discussed in Section 3 are ongoing development efforts to create improvisational micro-agents. We are not at the stage of being able to reflect on the success of this bottom-up approach in relation to building a more sophisticated and complete improvisational agent architecture. However, throughout the design process, several key concepts have emerged that have helped clarify the issues with building such agents. *Knowledge representation*, *communication*, and *interaction* between agents and with an external user are all common issues in building synthetic characters for digital narrative-based environments. However, given the particular real-world domain we are trying to emulate, we have had to reconsider each of these issues and work through how we intend to tackle each during this micro-agent development phase.

## 4.1 Knowledge Representation

When studying improvisational actors, it was no surprise that real world knowledge was an extensive knowledge base that they relied on to introduce concepts into a scene, make decisions about characters, etc. Actors reasoned about and referred to diverse facts about domains such as televisions shows, stereotypical superhero attributes, and how to rake leaves. However, this project is not focused on building a commonsense reasoning database, such as one that encapsulates American cultural knowledge. We have had to consider the problem of how to create agents that reason about a diverse set of knowledge while making the development of micro-agents a tractable problem that is well-scoped to address the core issues of computational improvisation without trying to solve a myriad of AI domains.

Our current solution (shown in Section 3.1) has been to focus on specific aspects of the kinds of knowledge reported: character schemata and the choice of attributes associated with degree of deviation from a prototype [23]. We have iterated multiple times, borrowing from cognitive psychology literature on categorical knowledge [25][27] and the theory of fuzzy sets [29], as well as from our data, on the design of our medieval world story representation. The purpose of this closed world knowledge base is to represent the kinds of knowledge (i.e., character schemata with fuzzy boundaries) that improvisers report employing, thereby focusing on *how* they employ this knowledge rather than *what* is all the knowledge they put to use. Agents can reason about different kinds of characters in the world, "interesting" choices to make regarding character, etc., without having to represent all of human knowledge. This is intended to inform future work both in terms of how to structure knowledge for improvisational agents (whether fictional or real world) and the algorithms employed to use that data.

## 4.2 Communication

We spent considerable time deliberating our options for communication methods between characters. Again, we have no desire at this stage to tackle hard AI problems, such as natural language understanding, so that has directed us to consider simpler alternatives for our micro-agents that still work effectively. We have considered the following basic methods for our agents in 3.2: selection of canned dialogue [5], context-free symbolic language interaction [36] or a simple natural language grammar approach.

We intend to explore multiple options with our micro-agents. For this particular set of agents, we wanted to explore an aspect of improvisation that dealt with cognitive consensus in a knowledge disparity game. Therefore, the communication used has to deal with ambiguous clues/concepts given by one agent. We opted for a symbolic language because it gave us the most flexibility for creating dialogue actions with multiple interpretations without the obfuscation of using canned sentences or grammars. The purpose of micro-agents is to quickly develop agents to test out ideas, so a symbolic language is complex enough to serve our immediate needs. However, this requires that any human observers or interactors must be able to understand or quickly learn this language, which has been pointed out by Crawford as a reasonable onus to place on a human user [36].

## 4.3 Interaction

A question our current work addresses less is "how will agents and/or a user interact with these agents?" We have so far considered letting users input initial information, such as the interestingness measure in 3.1. The work in 3.3 begins to explore how agents can interact with each other in a narrative environment (i.e., reasoning about how to influence the achievement of sub-goals for a narrative goal). Section 3.2 touches on characters engaging in symbolic dialogue through a back-and-forth knowledge disparity game. However, we have yet to reach the point where characters integrate these multiple aspects into a more unified interaction model, either in terms of interacting with each other or with a user. We contest that the limited user interactions are consistent with what we have observed in improvisational theatre. Audience members are rarely brought up on stage to have an integral part in the performance; they are simply not trained to do so, but are rather more commonly employed to give content suggestions for a given scene, which is our current focus in terms of agent/user interactions.

One could imagine mimicking improv games that do have audience members on stage to perform some aspect of the scene, but that has not been considered yet in our current micro-agents. This omission is largely due to the issue of knowledge representation (see Section 4.1). We work under the assumption that the agents in the world have an internally defined closed set of knowledge, which is not shared with a human user. We would have to provide that knowledge base to the user and constrain their interactions to that knowledge domain (or potentially design a game where only minimal knowledge is needed) if we were to "open up" the experience to direct user participation.

# 5. FUTURE WORK

The long-term future of this work is to develop a more unified approach for creating computational improvisational agents. We will need to consider next how to both synthesize this work and extend it. Our micro-agents have already illuminated several important issues with respect to how to create improvisational agents that "improvise how people do" (as described earlier in Section 1). This work has allowed us to clearly see what our next goals should be in terms of the computational side of this project. An obvious need is to further explore how to incorporate discourse, in whatever form, into the experience. Language is such an integral aspect of theatrical improvisation that it must be addressed. We also must decide on how much interaction the user will be allowed to have in various games, as mentioned in Section 4.3.

A final consideration for the future of this work is how to best incorporate it into domains outside of improvisation. There is a clear use for this work in the domain of interactive narrative, but we will need to consider how useful are the techniques we are developing in the larger domains such as computer games, pedagogical agents, and human-robot interaction.

# 6. References


[1] Hayes-Roth, B., Sincoff, E., Brownston, L., Huard, R., and Lent,, B., "Directed Improvisation," *Technical Report KSL-94-61*, Palo Alto, CA: Stanford University.

[2] Hayes-Roth, B. and Van Gent, R. 1996. "Story-Making with Improvisational Puppets and Actors," Palo Alto, CA: Stanford University.

[3] Perlin, K. and Goldberg, A. 1996. "Improv: A System for Scripting Interactive Actors in Virtual Worlds."

[4] Perlin, K. 2002. "Improving noise," In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, San Antonio, Texas, ACM, pp. 681-682.

[5] Bruce, A., Knight, J., Listopad, S., Magerko, B., and Nourbakhsh, I.R. 1999. "Robot Improv: Using Drama to Create Believable Agents," In AAAI Workshop Technical Report WS-99-15 of the 8th Mobile Robot Competition and Exhibition. AAAI Press, Menlo, pp. 4002-4008.

[6] Knight, J., "Characters as Models for Credible Agents," *Working Papers of the AAAI Fall Symposium*, Orlando, FL: 1998.

[7] Mateas, M. and Stern, A. 2002. "A Behavior Language for Story-Based Believable Agents," *IEEE Intelligent Systems*, vol. 17, pp. 39-47.

[8] Magerko, B. 2007. "Evaluating Preemptive Story Direction in the Interactive Drama Architecture," *Journal of Game Development*, vol. 2.

[9] Mendonça, D. and Wallace, W.A. 2004. "Cognition in Jazz Improvisation: An Exploratory Study," Chicago, IL.

[10] Seddon, F.A. 2005. "Modes of Communication During Jazz Improvisation," *British Journal of Music Education*, vol. 22, pp. 47-61.

[11] Gabrielsson, A. 2003. "Music Performance Research at the Millennium," *Psychology of Music*, vol. 31, pp. 221-272.

[12] Johnson-Laird, P. 2002. "How Jazz Musicians Improvise," *Music Perception*, vol. 19, pp. 415-442.

[13] Pressing, J. 1998. "Psychological Constraints on Improvisation," *In the Course of Performance: Studies in the World of Musical Improvisation*, B. Nettl and M. Russell, Eds., University Of Chicago Press, pp. 47-67.

[14] Ramalho, G.L., Rolland, P., and Ganascia, J. "An Artificially Intelligent Jazz Performer," *Journal of New Music Research*, vol. 28, pp. 105-129.

[15] Yokochi, S. and Okada, T. 2005. "Creative Cognitive Process of Art Making: A Field Study of a Traditional Chinese Ink Painter," *Creativity Research Journal*, vol. 17, pp. 241-255.

[16] Carter, C.L. 2000. "Improvisation in Dance," *The Journal of Aesthetics and Art Criticism*, vol. 58, pp. 181-190.

[17] Cunha, M.P.E., Cunha, J.V.D., and Kamoche, K. 1999. "Organizational Improvisation: What, When, How and Why," *International Journal of Management Reviews*, vol. 1, pp. 299-341.

[18] Mendonça, D. and Wallace, W.A. 2007. "A Cognitive Model of Improvisation in Emergency Management," *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, vol. 37, pp. 547-561.

[19] Johnson-Laird, P. 2005. "Flying Bicycles: How the Wright Brothers Invented the Airplane," *Mind & Society*, vol. 4, pp. 27-48.

[20] Harger, B. "Project Improv," *Project Improv*.

[21] Magerko, B. and Riedl, M. "What Happens Next?: Toward an Empirical Investigation of Improvisational Theatre," *on Computational Creativity 2008*, p. 151.

[22] Baumer, A. and Magerko, B. 2009. "Narrative Development in Improvisational Theatre," In the Proceedings of the *Second International Conference on Interactive Digital Storytelling*, pp. 140–151.

[23] Magerko, B., Manzoul, W., Riedl, M., Baumer, A., Fuller, D., Luther, K., and Pearce, C. 2009. "An Empirical Study of Cognition and Theatrical Improvisation," *Proceeding of the Seventh ACM Conference on Creativity and Cognition*, pp. 117–126.

[24] Nuckolls, C.W., 1999. "Cognitive Anthropology," *A Companion to Cognitive Science*, W. Bechtel and G. Graham, Eds., Malen, MA: Wiley-Blackwell, pp. 140-143.

[25] Rosch, E. 1999. "Principles of categorization," *Concepts: core readings*, E. Margolis and S. Laurence, Eds., pp. 189–206.

[26] Fauconnier, G. and Turner, M. 2003. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, Basic Books.

[27] Lakoff, G. 1990. *Women, Fire, and Dangerous Things*, University Of Chicago Press.

[28] Alexander, T. 2002. "An Optimized Fuzzy Logic Architecture for Decision-Making," *AI Game Programming Wisdom*, S. Rabin, Ed., Charles River Media, pp. 367-374.

[29] Zadeh, L.A., 1965. "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338–353.

[30] Leekwijck, W.V. and Kerre, E.E. 1999. "Defuzzification: criteria and classification," *Fuzzy Sets and Systems*, vol. 108, pp. 159-178.

[31] Mohammed, S. and Ringseis, E. 2001. "Cognitive Diversity and Consensus in Group Decision Making: The Role of Inputs, Processes, and Outcomes," *Organizational Behavior and Human Decision Processes*, vol. 85, pp. 310-335.

[32] Salas, E. and Orasanu, J. 1993. "Team Decision Making in Complex Environments," *Decision Making in Action: Models and Methods*, G.A. Klein, J. Orasanu, and R. Calderwood, Eds., Ablex Publishing, pp. 327-345.

[33] Traum, D. 1999. "Computational Models of Grounding in Collaborative Systems," *Working Papers of the AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, American Association for Artificial Intelligence, pp. 124-131.

[34] Clark, H.H. and Schaefer, E.F. 1989"Contributing to Discourse," *Cognitive Science*, vol. 13, pp. 259--294.

[35] Roberts, D.L., Riedl, M., and Isbell, C.L. 2009. "Beyond Adversarial: The Case for Game AI as Storytelling," *Proceedings of the 2009 Conference of the Digital Games Research Association*, London, UK.

[36] Crawford, C. 2004. *Chris Crawford on Interactive Storytelling*, New Riders Games.