Proceedings, The Thirteenth AAAI Conference on
Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17)

Turn-Taking with Improvisational Co-Creative Agents

Lauren Winston and Brian Magerko
Georgia Institute of Technology
Atlanta, Georgia 30332, USA
lwinston@gatech.edu, magerko@gatech.edu

Abstract

Turn-taking is the ability for agents to lead or follow in social interactions. Turn-taking between humans and intelligent agents has been studied in human-robot interaction but has not been applied to improvisational, dance-based interactions. User understanding and experience of turn-taking in an improvisational, dance-based system known as LuminAI was investigated in a preliminary study of 11 participants. The results showed a trend towards users understanding the difference between turn-taking and non-turn-taking versions of LuminAI but reduced user experience in the turn-taking version.

Introduction

Co-creative domains (i.e. collaborative creative domains where humans and/or intelligent agents create as equal collaborators) typically focus on improvisational interactions (Fuller and Magerko 2010). A key feature of improvisation is *turn-taking*, which is an agent's ability to lead or follow other agents in a social interaction. While turn-taking has been studied in domains such as human-robot interaction (Chao and Thomaz 2010), it has yet to be applied to open-ended, improvisational dance-based interactions with a virtual agent. Using dance as medium allows us to learn about the challenges of open-ended, improvisational interactions in turn-taking, and applying turn-taking to co-creative virtual agents can help create more natural interactions in systems where humans and intelligent agents create collaboratively together.

Humans are adept at perceiving leader / follower cues in co-creative human-human interactions and using these perceptions to guide their behavior, such as in conversations where humans can understand who should be talking and listening to avoid talking over one another. It is more difficult to convey these cues and roles in interactions between humans and intelligent agents. These interactions thus appear awkward and unnatural due to confusion over who is leading the interaction. A turn-taking model in a system with open-ended, improvisational interactions such as dance, requires an intelligent agent to interpret a human's action using a variety of signals, re-evaluate whether it should be leading or following, and perform creative and improvisational actions based on its leadership state.

We use prior literature on turn-taking to aid investigation of how a turn-taking model can be applied to a virtual agent engaging in improvisational dance-based interactions as part of an interactive dance installation called *LuminAI*. We aim to explore how the user experience in a co-creative experience like *LuminAI* can be improved with a turn-taking model, the limitations of turn-taking in such a system, and demonstrate how turn-taking can be used to influence human behavior in these interactions.

Related Work

Speech-based turn-taking

Some of the earliest literature on turn-taking stems from conversation analysis and has been used to inform interactions in speech-based systems. Sacks, Schegloff, and Jefferson (1974) describe a framework for conversation analysis, using turn-taking to refer to how actions in multi-agent interactions are ordered. This framework includes conversation features with insight into the turn-taking process, such as a dominating speaker, smooth state transitions, turn-allocation techniques (i.e. a speaker directs a statement to another speaker as a prompt), and variable turn order and length. Though this was intended for speech-based turn-taking, this work can guide analyzation of interaction mediums such as dance. Concepts such as variable turn order and length were applied to our strategy for turn-taking in LuminAI.

Morency, Kok, and Gratch (2008) discuss *backchannel feedback* in speech-based turn-taking. Backchannel feedback refers to the non-speech forms of feedback that occur during a conversation, such as eye gaze, nods, and "uh-huh"s. They tested and trained a sequential prediction model with human-human conversations to generate probabilities that something is a listener backchannel from a variety of speaker features. The features described are speech-based, and we believe that dance backchannel feedback is inherently different, making dance-based turn-taking more challenging to implement. Though our model does not predict whether a dance-based action is backchannel, this work supported our use of a probability-based model in LuminAI to offer flexibility when "backchannel" dance actions do occur.

Action-based turn-taking

We use *action-based turn-taking* to refer to agents engaging in actions other than speech. This is usually either *call-and-response* (Weinberg and Driscoll 2006), where a leader does one action, then the follower does another action, or what we call *sequential,* where the leader performs an action and the follower watches. The leader and follower switch roles when the leader finishes the turn by stopping the interaction.

Several of these interactions have been studied using the robot Simon. In one of these studies, a base condition was compared with a turn-taking condition (Simon's gaze was used as a cue to indicate which leadership state Simon was in) while sorting objects into bins (Chao and Thomaz 2010). A slight increase in turn count with the base condition showed that using turn-taking cues like gaze may reduce discrepancies about who is in which state by more clearly conveying to the human which state the intelligent agent is in.

In a later study, speech, motion, and gaze were all considered as components to signal to the robot which state it was supposed to be in during a version of the game "Simon Says" (Thomaz and Chao 2011). An additional study involving solving Hanoi towers compared a baseline condition to an interrupt condition, which meant that Simon stopped in the middle of an action and yielded his turn if it appeared that the human was beginning a turn (Chao and Thomaz 2012). The results showed increased engagement from human, less "awkwardness", and lower task execution time when using the interrupt condition. This shows that the use of interrupt conditions may lead to improvements in user experience, as a dominating intelligent agent that doesn't yield to humans enough may hinder user experience.

The MIROR Impro system, a musical technology aimed at children, gives an example of action-based turn-taking in an improvisational setting. In this system children can play physical instruments attached to the computer, which responds with a musical piece similar to the child's input. The system plays when a user stops playing and stops when a user plays, so this approach is almost entirely human-driven. The results suggest that though the system was interesting to children, the concept of how turn-taking worked was not understandable without additional instruction. This is an example of an improvisational system without simultaneous interaction between two agents and reinforces the need for cues and feedback for the user to understand the turn-taking (Wallerstedt and Lagerlöf 2011).

Shared mental model construction

Successful turn-taking requires an understanding of teamwork which can be modeled in shared mental models. Fuller and Magerko (2010) describe the idea of shared mental models in an improvisational interaction, defined as the "common framework of knowledge" that agents have about their current state and status. These models can be described in terms of cognitive convergence (when the models that different agents hold match) and cognitive divergence (when the models that different agents hold do not match). In cognitive convergence, there is the notion of an agent observing a divergence in the mental models, attempting to fix the divergence, and re-evaluating whether the mental model is now correct. Cognitive divergence may stem from differences in assumptions. This model was further developed to accommodate for "*co-creative improvisational agents*" (Hodhod and Magerko 2016). Unlike with other types of agents, the mental model that an agent has about itself is created throughout an improvisational interaction. Confidence factor is defined in this context as the strength of an agent's beliefs related to its mental model. Though these models focus on narrative improvisational interactions, these can be extended to account for a looser definition of turn-taking required in improvisational dance interactions.

Dance and rhythmic turn-taking

It is also important to look at music and dance-based interactions as these could possibly differ from speech-based interactions in their fluidity. An example of fluid turn-taking in an improvisational system is given in the form of a human and robot drumming "jam session" (Weinberg and Blosser 2009). A beat detection algorithm is used in a leader-follower model which detects beats from a human drummer. This paper claims that the leader's role in music is vaguer than in speech, and the role in this type of interaction stems from beat changes and tempo. If one changes the current beat, the robot assumes that the human is taking the leadership role. After the human has remained steady for a length of time, the robot will decide to take leadership role, and a human increasing the volume of his music could indicate that the human is now leading the interaction. A leader's role in dance is more similar to the leader role described here than in speech, but some characteristics, like volume aren't applicable.

Another study with a drumming robot used a probabilistic method and observations from human playing to help the robot decide when to begin or end a turn (Kose-Bagci, Dautenhahn, and Nehaniv 2008). The model used in this study uses duration and beat count of the previous turn to determine when to start and stop a robot turn. Three different models were used, but the least preferred model was the one in which the robot seemed to be leading for the majority of the interaction. In addition, this model results in overlaps where the robot and human interrupted each other's turn. This model was used to guide the probabilistic turn-taking method implemented in LuminAI but was not directly applied as beat count is less applicable in dance and because this turn-taking model contains no strategy for "relinquishing" the turn.

Another study observed children's interactions with a dancing Keepon robot (Michalowski, Sabanovic, and Kozima 2007). The robot danced rhythmically to children's physical dances, ignoring the rhythm from the music. Some children danced in the same way as the robot did by bobbing, and some were observed touching the robot in a rhythmic way. When the robot was already in-sync with music, more children began dancing with it than when the robot was out of sync with the music. There was a trend (though not statistically significant) that synchronous robot movements resulted in more rhythmic interactions. This aids our understanding of how rhythm impacts user behavior in dance and child users provide feedback back to the dancing agent.

LuminAI

LuminAI (Figure 1) is an interactive dance installation that was used to research and implement a novel model of dance-based turn-taking (see section *Turn-Taking in LuminAI*). The system (Jacob et al. 2013) tracks user gestures using the Kinect 2 depth sensor and analyzes user gestures within the Soar cognitive architecture (Laird 2012), using case-based reasoning (Aamodt and Plaza 1994) and Viewpoints movement theory (Overlie 2006). Viewpoints is a compositional technique for gestures that has been used for dance creation (Overlie 2006) and actor staging (Bogart and Landau 2005) by systematically analyzing movement along several perspectives or dimensions. The Viewpoints dimensions include space, shape, time, emotion and motion. A virtual character called VAI analyzes and learns from user movements and dances alongside a virtual representation of the user projected onto a screen. VAI can do nothing, mimic, transform the user's motions, use the user's motions as input to inform her motions, or dance "randomly" from her database of learned gestures.

Figure 1: LuminAI

Challenges to turn-taking in LuminAI

Improvisational vs. defined interactions

Various actions are occurring simultaneously in speech, such as backchannel feedback (Morency, Kok, and Gratch 2008), but the dominant action, speech, preferably is given by one agent at a time. This is not the case with fluid interactions, such as dance, where the primary action, dance, is expected to happen simultaneously across both leaders and followers. This makes it more difficult to perceive who is leading or following in an interaction. With defined interactions, agents can look for specific actions to decide whether to be leader or follower. In the simplest case, the leader does an action while the follower observes, so stopping the action signifies that the follower can assume the leader role. With more complex interactions, specific actions act as signals. In dance, one could imagine something such as a specific gesture towards the dance partner signaling the relinquishing of a turn and allowing the other party to take over. However, LuminAI uses improvisational actions, so we cannot rely on defined actions as turn-taking cues.

Fluid interaction

Weinberg and Blosser (2009) studied improvisational, action-based turn-taking in drumming, but in this, the leader and follower still have distinct roles. The leader sets a beat in a drumming interaction that the follower must adhere to. In dance, agents are dancing simultaneously, and there are no distinct actions that always define a leader or follower. Dancers may copy, dance similarly to, or dance entirely differently than their partner. A leader and follower both may do any of these actions at different times in a dance. For instance, a leader may notice a perceived follower becoming disengaged and try a new innovative move to re-engage the follower.

Dance-specific turn-taking cues

The work by Weinberg and Blosser (2009) was used to guide our implementation of fluid, improvisational turn-taking in LuminAI. Though some of the features described in this work can be applied to dance (such as tempo), others (such as increased volume) cannot. The cues used in dance are inherently different from those used in speech, drumming, or action-based interactions. Turn-taking cues in dance-based interactions were determined using informal observations of dancing, which may differ across styles. For example, in ballroom dancing, the leader is set before the dance and remains constant throughout. The leader often uses haptic feedback, such as pressure on the follower's arm, and eye gaze. In informal, improvisational group dances, haptic feedback is less commonly used, but eye gaze and enthusiasm were observed as markers. Leader and follower roles are often less clearly defined. Dancers may mimic one another, dance similarly to one another, or dance entirely differently. To take over or "lead" the interaction, dancers were observed using eye-gaze and more "enthusiastic" motions to re-engage another dancer or affect their dance. "Enthusiasm" involves wider or more up-tempo gestures.

Projection vs. physical robot

Additional barriers to turn-taking in LuminAI stem from VAI being a virtual agent instead of a physical robot. Working with a physical robot allows use of different types of measures within the turn-taking model. In LuminAI, the intelligent agent is virtual, projected onto either a flat screen or onto the walls of a geodesic dome, so VAI is unable to use physical feedback, such as pressure. Because VAI is projected, eye gaze is also an unreliable measure for turn-taking, as the person interacting with the system is going to have different eye gaze patterns when interacting with an agent on a large screen than a physical robot in front of her.

Turn-taking in LuminAI

The previous version of LuminAI has no turn-taking model implemented; when a user dances, VAI mimics the user's dance. When a user gesture is detected, VAI attempts to respond with an innovative response from that gesture. If the user is still for too long or repeats the same gesture for too long, VAI becomes bored and dances with a random response. There is no representation of leader or follower visually or conceptually in the design, but VAI's behavior most closely resembles the follower role since she behaves responsively to the human.

We determined what *following* and *leading* means within the LuminAI system based on related research, informal observations, and the current interaction model in *LuminAI*. Agents in co-creative interactions have a) a leadership state which may be leader, follower, transitionary, or neutral, and b) a perception of the leadership states of other agents in the interaction. These leadership states may switch throughout the interaction as agents receive turn-taking cues from one another. In co-creative interactions, there may be discrepancies between how an agent perceives another agent's leadership state and what that leadership state actually is. As such, it is possible for agents to have conflicting leadership states, such as multiple followers with no leaders or multiple leaders with no followers. The turn-taking agent in LuminAI, referred to as TT-VAI, only has a notion of her own leadership state and is not concerned with human perception of their leadership state.TT-VAI uses different cues and signals to decide whether she should act as a follower or leader in a given moment. Viewpoints predicates of energy, tempo, and size are used as a measure of user enthusiasm.

In turn-taking a follower's actions are mostly directed by the actions of the perceived leader, so followers in LuminAI dance responsively to the leader's dance as an accompaniment. A follower in LuminAI primarily will mimic the leader's motions, but a follower may also decide to dance in a way that is complementary to the leader. Complementary dancing includes transformations of the leader's gestures (such as inverting the dance) or performing gestures that the agent has previously learned as similar to the leader's gestures.

A leader's actions are mostly self-directed. Leaders in LuminAI may mimic the follower from time to time, but a leader primarily will perform gestures from its database of previously learned gestures. As LuminAI doesn't currently utilize audio input, the leading agent may use the follower's gestures to get a feel for the type of dance gesture that is contextually appropriate but also may just dance by choosing a "random" gesture from her database of previously learned gestures without using any input from the follower.

A leader remains a leader until deciding to relinquish the turn to the follower or receiving a cue that the follower wants to assume leadership. There are no signals in improvisational interactions that always indicate someone is relinquishing leadership or becoming leader. Certain signals may indicate that it is likely that someone wants to change leadership state, but none are 100% reliable. To accommodate this, the model for turn-taking in LuminAI implements a probabilistic model for TT-VAI deciding whether or not to change states similar to the one used by Kose-Bagci, Dautenhahn, and Nehaniv (2008). This model generates a set of probabilities for TT-VAI's next state using time in each state, the predicates, and a stillness tracker as reinforcement. A rhythm tracker is additionally used to detect repeated gestural motions, which may indicate lower engagement or enthusiasm.

In addition to the two leadership states of leader and follower, there is a third state used as a sub-state as a way to change behavior without changing state to re-engage the user. For example, if TT-VAI is leading and notices the user becoming disengaged, she may decide to try a different type of gesture to re-engage the user. Through this, TT-VAI calculates probabilities for three states. These probabilities are calculated approximately once per second at the end of an action or at the end of a specified time parameter. These states are 1) remaining in the current leadership state with the same behavior, 2) remaining in the same leadership state while attempting some type of modified behavior to alter the user's behavior, or 3) changing leadership state. Prior work has shown that turn-taking cues are important for user understanding (Wallerstedt and Lagerlöf 2011). VAI's body changes between red and blue to indicate to users that she is leader or follower, respectively.

Changes in predicate values of energy, tempo, and size may signal a change in user engagement, so their changes update the probabilities of TT-VAI remaining in a state. Certain signals increase or decrease the probability that TT-VAI remains leader or follower. For example, if TT-VAI considers herself a "follower", the human becoming still may indicate that TT-VAI can begin leading, but it may also mean that the human is simply pausing for a moment and still wishes to maintain leadership. The longer that a human remains still, the more the probability for TT-VAI to "take over" increases.

Evaluation

Methodology

A preliminary user study was conducted to investigate user experience and understanding of the turn-taking model implemented in LuminAI. 11 participants were recruited for an hour-long user study and were each compensated $20. The majority of participants were college students with varied prior experience with dance and technology.

In a successful turn-taking system, users should have enough understanding to distinguish between leading and following, change behavior depending on the agent leading or following, and be more engaged than in a comparable system without turn-taking. To measure this, participants interacted with two versions of LuminAI: one version that implements turn-taking and one version that does not. Participants were randomly assigned to one of two groups to reduce ordering effects. Group A interacted with the base version first and Group B interacted with the turn-taking version first. To further reduce ordering effects, before beginning either session, participants interacted with both versions of the system during two short warm-up sessions.

After the warm-up sessions, participants interacted with each version of the system for 3 to 7 minutes. After each session, the participant answered a series of Likert-scale questions modeled after questionnaires used in other evaluations of turn-taking experiences with autonomous agents (Cassell and Thorisson 1999, Chao and Thomaz 2010, and Chao and Thomaz 2012). A recording of their dance was reviewed and used to aid discussion about their actions and/or perceptions of VAI's actions. After finishing both sessions, participants filled out a survey which compared the versions, asked which version of the system was preferred, and collected information about their prior experience with similar systems.

Results

Though population size was too small to hold statistical significance, the results show promising trends in user understanding of the system. 8 of 11 participants understood that there was a difference between the two versions through a clear preference for one version or mentioning the differences between the versions during the second video review or questionnaires. Of these, one participant incorrectly interpreted base VAI as leading and TT-VAI as only copying. The other 7 correctly noticed differences between the two versions. Though 1 of these participants also felt that base VAI was more innovative, he correctly observed that the interaction with TT-VAI was more "back-and-forth." 6 identified that the base condition seemed to mimic more than the turn-taking one.

The 3 participants who were unable to distinguish the versions preferred the second version they interacted with, citing familiarity in the questionnaires. The other 8 participants had clear preferences for one of the two versions because of differences in the systems. The participant who noticed a difference in the versions but misinterpreted TT-VAI as always copying preferred TT-VAI because of perceived increased mimicry, while the other participant who felt TT-VAI led less than the base but recognized the turn-taking version as being more "back-and-forth" preferred the base version because it felt more natural and "gave more ideas".

Of the 6 who recognized that the base version mimicked more, 2 preferred the turn-taking version because of less mimicry, saying that TT-VAI "seemed like it was doing more" and "seemed more ready to throw something into the party." The other 4 preferred the base condition because of VAI's increased mimicking.

In the questionnaires, several participants cited increased responsiveness and mimicry as their reasons for preferring the baseline version. Some of the comments given include "[this version] was more responsive to my movement and, thus, I would prefer [it]" and "I preferred [this version] because it mimicked better." Responses to two questions on the session questionnaire indicate increased awkwardness in the turn-taking version, as no participants felt that the base condition was more awkward than the turn-taking condition. Slightly more than half of participants indicated no difference in awkwardness across the two versions. The rest (5) felt that the turn-taking version was more awkward than the base.

In total, 4 participants preferred TT-VAI, one due to misinterpreting TT-VAI as always mimicking, one due to increased familiarity, and 2 due to TT-VAI mimicking less. Despite efforts to reduce ordering effects, participants showed slight bias towards the second version that they interacted with on two comparison questions about VAI's responsiveness and influence. 7 of 11 participants preferred the second version of the system that they interacted with, and 7 of 11 preferred the base condition to the turn-taking condition with no discernable difference between the two groups. Though this trends towards users exhibiting a slight preference for the second version of the system that they interacted with, feedback users gave as to the reasons for their preferences suggests that people may prefer the base condition to the turn-taking condition.

Within the turn-taking version, 5 participants were able to tell a difference between how TT-VAI behaved as a leader and as a follower. 2 participants were able to distinguish that TT-VAI was following when blue and leading while red. An additional participant who noticed a difference in behavior while dancing discerned the correct difference during the video review session. Other participants either were unable to distinguish the colors or had incorrect assumptions about the colors. One thought that the red color indicated that TT-VAI was "thinking," while the blue color indicated that TT-VAI "understood" what the user was doing and was thus following them more closely. Another thought red meant hesitance and blue meant confidence and synchronization.

There was a slight preference for the baseline version in questions which asked about the creative ideation of the dance, how the users influenced VAI's behavior, and VAI's responsiveness. This is expected, as the baseline version of the system is more user-centric so users would have increased ability to impact the dance ideation and influence VAI. Interestingly, in a question asking participants to compare the creative ideation of the two versions, all of Group B felt that the creative ideation was the same or more human-focused in the baseline version, but Group A's responses were more varied. This suggests that the order in which participants interacted with the two versions may have impacted perceptions.

From this, we can gather that since many participants expressed a preference for mimicking, people may prefer the follower mode over the intelligent agent being innovative and "leading" the interaction. Participants expressed a dislike for TT-VAI seemingly ignoring them more while leading the interaction. One example of a comment that showed this was, "VAI was an interesting dance partner to watch but I'm not sure we were interacting." In one question participants were asked to compare VAI to a dance partner who always ignores. TT-VAI Participants tended to rank base VAI higher than TT-VAI, suggesting many participants felt TT-VAI was "ignoring" them more than base VAI.

In the questionnaires 9 participants reported being "very comfortable" interacting with technology, and 8 considered themselves at least "somewhat comfortable" interacting with interactive art. However, most participants considered themselves inexperienced with dance, and responses to comfort dancing in public were varied. There was no discernable connection between a person's experience with dance and how they interacted with and understood the system.

Conclusion

This research has investigated how an intelligent agent engaging in dance-based interactions can use a turn-taking model to improve user experience. There are several ways that this work can be extended for future exploration.

The data collected from the user studies will be used to iterate on the turn-taking model to improve user experience. Prior work has shown that humans dislike when intelligent agents lead the majority of the interaction (Kose-Bagci, Dautenhahn, and Nehaniv 2008) which aligns with the results from our study. We can modify our turn-taking algorithm to bias towards humans being leaders by increasing the amount of time humans are leader relative to the agent and making it easier for humans to claim the leader role. However, the reasons *why* this negative perception of leading agents in co-creative interactions exists require additional research. External factors such as prior experience with dance, interactive art, and intelligent agents may affect how people perceive leading co-creative agents. While this study found no correlation between level of experience with dance and system perception, a larger study with more variety in participant background could help show if different demographics or experience levels impact perception of a leading agent.

The amount of time that participants interact with the system may also be a factor in how leading agents are perceived and how well turn-taking is understood. In this study, all participants did brief warm-up sessions to familiarize themselves with the system, but as these sessions only lasted a few minutes, the system itself was still relatively novel to participants when the recorded sessions began. It's possible that we underestimated the amount of time that it takes for new users to familiarize themselves with LuminAI and that participants were still exploring the bounds of the system when the actual sessions began. If participants were still becoming used to simply interacting with VAI at the time the session began, they may be unable to understand the turn-taking aspects of the system or why the agent's behavior changed when taking leadership. A sudden change in agent behavior could be confusing if participants aren't given adequate time to explore and understand these behavior changes. We can increase the length of the sessions in a future study to see if interaction time has an impact on perception of leading agents. Giving participants much longer periods of interaction with LuminAI before and during the recorded sessions would allow more exploration of the system and might change how participants understand the system and perceive the leading agent.

Another way of improving experience would be training a classifier to use energy, tempo, gesture size, and user stillness to interpret different leadership states automatically instead of with the probabilistic model. This could be further extended to use feedback while dancing to dynamically learn how to behave during each leadership state.

Additional input, such as the direction users face or beat detection from played music, could be used to inform the agent on which state it should be in.

Additional feedback could be shared with users about VAI's leadership state, as the results from this study suggest that colors are not clear at indicating state in dance-based interactions, as many had incorrect assumptions about their purpose or didn't notice the color change. Auditory and additional visual cues, such as a spotlight on the agent VAI perceives as the leader or VAI moving forwards as leader and backwards as follower, may help increase understanding of the turn-taking modes.

Another area for further exploration would be multi-party turn-taking between one human and multiple virtual agents, one agent and multiple humans, or multiple humans and agents simultaneously. This has been explored with conversation (Sacks, Schegloff, and Jefferson 1974), so an investigation into dance-based, multi-party turn-taking could improve our understanding of improvisational turn-taking with virtual agents.

This work describes a turn-taking model for co-creative dance interactions and a preliminary study into how this model is understood and impacts user experience. Implementing turn-taking in LuminAI was a particularly interesting challenge due to its open-ended, improvisational nature. The results show promising trends in user understanding of turn-taking but in reduced user experience. Further work should be done to determine if the trends continue with a larger population size.

Acknowledgments

We would like to thank Mikhail Jacob, Duri Long, Nicholas Davis, and Dor Hananel for their assistance and support of this work.

References

Aamodt, A. and Plaza, E. 1994. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications* 7(1): 39-59.

Bogart, A., and T. Landau. 2006. The Viewpoints Book: A Practical Guide to Viewpoints and Composition. Theatre Communications Group.

Cassell, J., and Thorisson, K. R. 1999. The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence 13: 519–538.

Chao, C. and Thomaz, A.L. 2012. Timing in multimodal turn-taking interactions: Control and analysis using timed petri nets. *Journal of Human-Robot Interaction* 1(1).

Chao, C. and Thomaz, A. L. 2010. Turn Taking for Human-Robot Interaction. In *Proceedings of AAAI fall symposium: dialog with robots.*

Fuller, D. and Magerko, B. 2010. Shared mental models in improvisational performance. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, 15. Monterey, Calif: Association for Computing Machinery.

Hodhod, R. and Magerko, B. 2016. Closing the Cognitive Gap between Humans and Interactive Narrative Agents Using Shared Mental Models. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 135 - 146. Sonoma, Calif.: Association for Computing Machinery.

Jacob, M.; Coisne, G.; Gupta, A.; Sysoev, I.; Verma, G.G.; and Magerko, B. 2013. Viewpoints AI. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference.*

Kose-Bagci, H.; Dautenhahn, K.; and Nehaniv, C.L. 2008. Emergent dynamics of turn-taking interaction in drumming games with a humanoid robot. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication*, 346 - 353. Munich, Germany: Institute of Electrical and Electronics Engineers.

Laird, John E. 2012. The Soar Cognitive Architecture. MIT Press.

Michalowski, M.P.; Sabanovic, S.; and Kozima, H. 2007. A dancing robot for rhythmic social interaction. In *Proceedings of 2nd ACM / IEEE International Conference on Human-Robot Interaction (HRI)*, 89 - 96. Arlington, Virginia, USA: Institute of Electrical and Electronics Engineers.

Morency, L.P.; Kok, I.; and Gratch, J. 2008. Predicting Listener Backchannels: A Probabilistic Multimodal Approach. In *Intelligent Virtual Agents,* 176 - 190. Berlin, Heidelberg: Springer.

Overlie, Mary. 2006. "The Six Viewpoints." Training of the

American Actor: 187–221.

Sacks, D.; Schegloff, E. A.; and Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50: 696-73

Thomaz, A. L. and Chao, C. 2011. Turn-taking based on information flow for fluent human-robot interaction. *AI Magazine* 32(4): 53 - 63.

Wallerstedt, C. and Lagerlöf, P. 2011. Exploring turn-taking in children's interaction with a new music technology. *He Kupo* 2(5):20-31

Weinberg, G. and Blosser, B. 2009. A leader-follower turn-taking model incorporating beat detection in musical human-robot interaction. In *Proceedings of the 4th ACM / IEEE international conference on Human robot interaction*, 227 - 228. La Jolla, Calif.: Association for Computing Machinery.

Weinberg, G. and Driscoll, S. 2006. Robot-human interaction with an anthropomorphic percussionist. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. Association for Computing Machinery.