

Designing for Socially Interactive Systems

Duri Long, Mikhail Jacob, Nicholas Davis, Brian Magerko

Georgia Institute of Technology

Atlanta, Georgia, USA

{duri, mikhail.jacob, ndavis35, magerko}@gatech.edu

ABSTRACT

This paper reports on the design and evaluation of *LuminAI*, a socially interactive art installation in which participants can engage in collaborative movement improvisation with virtual agents and other humans. *LuminAI* was used as a technical probe to study social interaction within interactive art at a local art gala during which over 100 participants interacted with the system. Video and interview data was gathered during the event and analyzed using thematic analysis to develop a taxonomy to guide the design of socially interactive systems involving humans and artificial agents. This taxonomy helped us identify areas where *LuminAI* was successful, where it needs improvement, and conceptual spaces we have yet to explore.

Author Keywords

Artificial intelligence; social interaction; computational creativity; Viewpoints; dance improvisation; interactive installation

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION

Computer colleagues have been developed to facilitate creative collaboration and ideation in human creators [13]. Most of these colleagues are designed to work one-on-one with a human partner to creatively improvise in fields like dance [22, 30], drawing [16], and theatrical improvisation [28]. However, co-creative computer colleagues have not yet been widely employed and studied within large social contexts (i.e. crowd or multi-user environments). In order to explore how to design socially interactive creative systems involving multiple humans and AI agents, we created *LuminAI*, a multi-user interactive dance installation.

Studying creativity within social environments involving both humans and agents raises new questions in terms of the human-centered design of AI-based experiences. For

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

C&C '17, June 27-30, 2017, Singapore, Singapore © 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-4403-6/17/06...\$15.00
<http://dx.doi.org/10.1145/3059454.3059479>

instance, how do we evaluate the creativity of multiple agents in a social environment? How do humans and agents interact and collaborate in a multi-user environment? How does physical space affect social interaction? How do people and agents configure their awareness of each other?

Existing work across several disciplines begins to investigate some of the above questions. Methods have been created for evaluating both creativity support tools [11] and creative engagement [18]. Studies of small group creativity [15, 28, 29, 32, 35] and societal creativity [14, 36] deal with how humans collaborate creatively in social environments. Work at the intersection of the fields of sociology, Computer Supported Collaborative Work (CSCW), architecture, and technology looks at the social affordances of physical spaces [10], social interaction dynamics amongst humans [20], and social communication and collaboration in work environments [4]. Empirical work has also been done to study human-agent interaction dynamics in creative task environments, but the vast majority of these studies focus on interaction between a single human and a single AI agent [8, 15, 22, 30]. This is still an area of research with a lot of room for investigation, and there is a strong need for empirical work investigating human-agent interaction in socially creative environments.

In this paper, we investigate human-agent co-creative movement improvisation in a social dance installation called *LuminAI*. We begin by describing a public installation and evaluation of the *LuminAI* system. We then discuss how the findings from this event informed the development of a detailed taxonomy to guide the design of socially interactive agent-based systems. The taxonomy identifies key design considerations for socially interactive systems, building off of existing work [4] that introduces five high-level themes (*territoriality*, *awareness*, *control*, *interaction*, and *transitions*) that should be considered in the design of social spaces.

THE LUMINAI INSTALLATION

The *LuminAI* installation was created to allow multiple people to dance with each other in collaboration with artificial agents within a geodesic dome. Digital silhouettes of participants and virtual agents are projected onto the walls of the dome. A participant can move his or her body (and consequently his or her silhouette), and the agent will respond with an improvised movement of its own. Multiple agents and humans can meet in the liminal space of the dome and their shadows can dance together. This is in contrast to previous dyadic versions of *LuminAI* (formerly



Figure 1. LuminAI Dome installation: mock-up (left) and live interaction (right). Mock-up credits Jessica Anderson.

Viewpoints AI) where a single human could playfully engage in dance-like movement improvisation with a single virtual character in a two-dimensional shadow theater installation [22]. In these previous versions, any multi-person interaction was purely incidental.

LuminAI is housed within a 23-foot diameter wide, 15-foot tall 3V(B) steel geodesic dome frame with 105 projection-ready translucent textile panels covering the surface. Participants can enter and exit the dome through doorways created by removing textile paneling. They can then dance with each other inside the domed space. An audience outside the dome can view the ongoing shadow party installation as a performance.

The agent architecture controlling the virtual characters uses case-based reasoning [23] to learn movements and patterns of interaction from all participants over time. Viewpoints [5], a widely used theory in the fields of theater and dance that provides a way of thinking about movement and gesture, is used as the system's vocabulary to procedurally represent and reason about movement [21]. Viewpoints guides the system in reasoning about movement along the dimensions of tempo, duration, repetition, kinesthetic response (natural movement timing), spatial relationships, topography (traced floor patterns of movement over time), shape (of movement and body), gesture, and architecture. It also responds improvisationally using procedural response strategies when learned knowledge is not sufficient or appropriate.

The current prototype of the installation consists of three separate interaction spaces, spanning the entire dome's inner surface. The combined video output is projection mapped to the external surface of the dome using MadMapper [26] and displayed on its exterior using three externally mounted projectors. The installation captures participant movements inside the dome using three depth sensors distributed equally around the dome, facing its center. A new prototype currently in development replaces the three separate interaction spaces with a single virtual domed space that receives multiple Kinect inputs and has multiple virtual camera views of that space projected onto the exterior of the dome.

LuminAI was designed as a technical probe to investigate how to convert a dyadic co-creative probe movement-based improvisational experience into a socially interactive experience. In order to create such an experience, a geodesic dome was chosen as the interaction space due to its potential social affordances. We hypothesized that the dome would a) allow for a spectrum of multi-user experiences ranging from dyadic human-agent interaction to a crowd-level dance party; b) create a liminal space that is clearly defined for the co-presence of groups of people and AI agents; c) afford social interaction by being enclosed, forcing people to be in close proximity to each other and shielding participants from outside observers; and d) attract participants to participate via visual spectacle.

EVALUATION METHODOLOGY

We decided to study participant interactions with LuminAI in order to investigate the following research questions: 1) How do we design interactive systems and co-creative AI agents for a multi-user environment?; 2) How do people make sense of AI agents when they are transported into a multi-user environment?; and 3) What aspects of LuminAI help facilitate social interaction? LuminAI was exhibited for a week at the Goat Farm Arts Center in Atlanta, with the opening gala attended primarily by artists and art connoisseurs. The event was six hours long, and during this time, approximately 100 people were recorded on video interacting with the LuminAI installation. All participants were informed that they would be recorded via clearly visible signs placed outside of the installation.

The prototype of LuminAI exhibited was specifically customized for the opening gala. At the event, LuminAI was also being used as a human-AI performance space for a choreographed dance. Therefore, due to the aesthetic requirements of the dance performance, only two-thirds of the textile projection panels were installed, leaving a large cutaway where the audience could see the participants, their digital silhouettes, and the virtual characters. Soon after the start of the opening gala, the visualizations presented to participants were switched from showing only the virtual characters to showing all participants' digital silhouettes as well as the virtual characters (see Figure 2 for an image of the dome with just the agents, no silhouettes). This was

done pragmatically based on observing participant interactions with the system and deciding that it was important for all participants to see some immediate feedback when they entered the interaction space.



Figure 2: The LuminAI installation prior to turning on the silhouettes.

Nearly an hour (57:53) of video data was collected throughout the event. For approximately 13 minutes (12:59) of video data the projected visualization consisted only of the virtual characters. The rest (44:54) of the video data shows digital silhouettes of the participants along with the virtual characters.

During the gala event, we conducted brief, semi-structured interviews with 18 different participants. The semi-structured interviews were conducted based on the following questions: a) What is your initial reaction to the experience you just had? b) How do you think the system works? c) Did you enjoy your interaction? d) What was confusing or frustrating? e) How might the experience be improved? Participants' responses were recorded in field notes that were later analyzed.

Both the field notes and the video data were analyzed using thematic analysis [7]. We identified common themes within the data, and then coded the data for each of these interaction trends. After coding the data and reviewing the literature, we found that many of our observations fit under the umbrella of an existing high-level framework (see *Findings*). However, this existing framework did not go into sufficient detail to explain some of our observations. As a result, we used the results from our analysis to create a new, more detailed taxonomy for analyzing social interaction in human-agent systems (see Figure 3).

FINDINGS

Based on a synthesis of literature across disciplines including sociology, CSCW, architecture, and cognitive science, as well as observations from the public installation of LuminAI, we present a taxonomy (see Figure 3) of conceptual spaces that should be considered when designing agent-based systems that are focused on facilitating social interaction. At the highest level, this taxonomy is organized according to the *Tacit* framework:

five design themes that “provide a way of structuring the issues that designers of social spaces need to deal with” [4]. We chose this framework because the high-level results from our data-driven analysis corresponded with these five themes (territoriality, awareness, control, interaction, and transitions), but the themes as originally presented in the literature were not specified in enough detail to thoroughly answer our questions about social interaction in LuminAI. In addition, the *Tacit* framework does not address how each of these themes might be affected by introducing artificial agents into the social space. As a result, we have expanded and specified the framework in more detail based on a synthesis of the literature as well as our own empirical observations. This section provides an overview of the literature along with empirical data and detail how both contributed to the resulting taxonomy.

Territoriality

The first theme, territoriality, “concerns spaces and the relationships between people and spaces” [4]. Based on relevant aspects of territoriality discussed in the literature, our taxonomy divides territoriality into two conceptual spaces: *territory* and *conflict*.

Territory

Background

Territory deals with the relationship between physical space and interaction dynamics. Humans have “spatial literacy” and are able to make sense of social situations and interactions due to people’s spatial relationships to each other [10]. As a result, different spatial constraints afford certain types of interactions [10]. For example, in dance clubs, different practices are acceptable in different spaces [27]. It may be normal to dance on the dance floor, but this type of interaction is less acceptable at the bar [27].

The physical barriers between different interaction spaces can serve as *involvement shields* or “barriers to perception” [20] that allow individuals to go “off stage” or “out of play” [27]. For example, the barrier between the dance floor and the bar allows people to move “backstage” and retreat from active physical participation [27]. In addition to defining the barriers between territories, the liminal “regions and seams where spaces come together” can also facilitate certain types of social interactions and transitions [4]. For instance, an individual standing in the space between the dance floor and bar may interact differently than they would in the center of the dance floor or sitting at the bar.

Observations

In the case of LuminAI, three different areas of interaction were observed: the physical space inside the dome, the digital projection space on the dome surface, and the area outside of the dome. People outside of the dome were typically not interacting directly with the system, and spent their time either watching participants or engaging with other non-related activities. People physically inside the dome were typically engaged with the system, and their

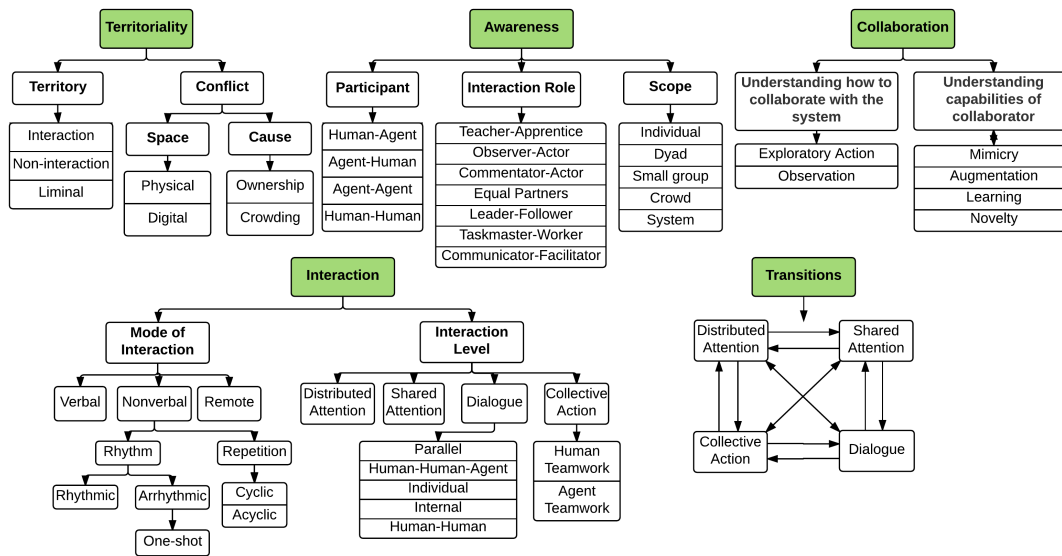


Figure 3: Synthesized and Expanded Tacit Taxonomy

interactions with the system involved both physical movement on the dance floor and virtual interaction between their shadow and the agent.

Reflections

In our taxonomy, these territories are described by three different types: interaction (the space where interaction with the system is occurring), non-interaction (the space where interaction with the system is not occurring), and liminal (the space between interaction and non-interaction spaces). Some spaces can take on two of these roles; for instance, the dome projection surface serves both as a liminal space separating the dance floor from the outside and as an interaction space where human silhouettes and virtual dancers dance.

Conflict

Background

The second aspect of territoriality in our taxonomy deals with *conflicts* of space that arise and how they are addressed. Conflicts can be positive in that they force users to come together and engage in conflict management, but they can also contribute to a negative user experience [31], so it is important to consider them when designing a system. Conflicts can occur in two different *spaces*: *physical* and *digital* [31]. Conflicts can also result from numerous different *causes*. One cause noted in the literature was the invasion of personal space [31]; we call this an *ownership* issue.

Observations

The ownership conflict type did not occur frequently in the LuminAI installation. There were four instances in which one participant stepped in front of another participant as they were dancing with an agent, but this was never upsetting enough to spur conflict management or dialogue.

It was also observed that conflicts in digital space occurred due to *crowding*, which resulted in participant confusion. Four participants mentioned in their interviews that they found it difficult to distinguish which agent was responding to their movements, and three additional participants noted that they were able to see their own silhouette, but did not notice the virtual agent. It is clear from the video that early in the event (when there were fewer people and the silhouettes were turned off), there was sufficient digital space. However, later in the evening, when the silhouettes were turned on and more people arrived, the digital space became crowded. Silhouettes and agents began to overlap, and this made it difficult to distinguish between the two.

Conflicts due to crowding in physical space were less of an issue in LuminAI. Towards the end of the event, when many people were present in the dome, the installation became more like a crowded dance floor than an interactive movement improvisation experience. Some participants still attempted to interact with the agents/silhouettes despite the crowded space, but others adapted and danced with each other when they were not able to approach an agent. Everyone appeared to be dancing fluidly, and no one halted his or her interaction to engage in conflict management.

Reflections

The lack of ownership conflicts could be because dancing is an ephemeral interaction. In contrast to public digital workspaces where participants create a lasting product in their personal space, personal contributions in LuminAI disappear just moments after they occur. Possible solutions to crowding in digital space include reducing visual clutter by turning off the silhouettes, creating a visual mapping from the agent to the participant with whom it is interacting, and introducing a limit on the number of participants that are tracked by the system at one time.

Awareness

The second theme focuses on awareness of others and their activities. Our taxonomy divides awareness into three aspects: *participant*, *interaction role*, and *scope*. These deal with who is aware of whom, how people configure their awareness of each other, and the extent of one's awareness.

Participant

Background

The participant aspect describes who is aware of whom in a given space. Literature in the field focuses almost exclusively on awareness amongst humans (*human-human*).

Observations

During LuminAI, it was observed that additional types of awareness affected the interaction experience. In particular, awareness issues involving virtual agents (*human-agent*, *agent-human*, and *agent-agent*) were relevant. Seven different participants in the video recordings circled around the dome or switched sides of the dome while dancing, possibly in an attempt to get their agent to dance around the dome with them. However, since the initial prototype of LuminAI involved three individual agents stitched together using projection mapping, each agent could only dance within a defined sector of the dome. This discrepancy caused some confusion; for example, one participant hopped from one side of the dome to the other and waited to see if his agent would move with him; when it did not, he had to switch and dance with another agent.

Reflections

These observations led to the conclusion that when artificial agents are introduced into a social interface, new types of awareness need to be considered. All of these scopes of awareness also apply beyond a dyadic scope. For instance, human-agent awareness could involve three humans and one agent who are mutually aware of each other.

Prior to developing this framework for analyzing socially interactive systems, we had not considered the possibility of agent-agent awareness. Implementing functionality that allows agents to be spatially aware of each other would solve some of the digital space conflicts that were identified by participants (see *Conflict*). Agents could even move beyond spatial awareness and dance with each other; this could potentially serve to attract participant attention when no humans are dancing in the dome.

The installation would also benefit from improved agent-human awareness. Discontinuities like the ones observed could cause less eager participants to disengage, and they could also inhibit participant understanding of the agent's ability to learn from past action sequences. In future installations, agents could be made aware of the three-dimensional space and be able to track and follow their partners as they move around the dome.

Interaction Role

Background

Awareness of others can occur in many different forms

depending on the roles that are adopted by individuals in the group. This is what the *interaction role* aspect addresses. When interacting with multi-user displays, people often do not participate as *equal partners* and instead configure their awareness of others in terms of relationships like *teacher-apprentice*, *observer-actor*, or *commentator-actor* [31]. Computers can also take on some of these roles in the creative process, in addition to assisting in a more traditional manner as a *taskmaster* or *communication facilitator* [24].

Observations

These roles are not always dyadic; they can also appear within larger social groups. For example, in one group of four that interacted with LuminAI, three of the participants were actively dancing with each other and the agents (adopting the actor role), and one participant stood back and watched (adopting the observer role). These interaction roles are not necessarily static; in this same group, each of the three actor participants stepped back at some point and briefly adopted the role of observer, and the observer participant physically interacted with the system twice.

In addition to the six roles previously identified in the literature, an additional interaction relationship was observed. The *leader-follower* relationship occurred both between humans and between humans and agents. This is distinct from the teacher-apprentice role in that the leader is not always consciously trying to impart knowledge, as a teacher would.

Different interaction roles were adopted amongst participants who were familiar with each other and amongst strangers. Participants who entered together in a group were more likely to adopt the equal partner or leader-follower relationship. Strangers typically did not interact with each other or adopted the observer-actor relationship.

Reflections

Existing work further characterizes the leader-follower behavior that was observed in both human-human and human-agent interaction [1, 12, 38]. The LuminAI agents are currently designed to act as followers unless the human ceases interaction and needs encouragement. Ongoing research is exploring how the agent might alternate naturalistically between leader and follower roles.

Scope

Background

Finally, *scope* deals with the extent of one's awareness of others. Small physical spaces like the dome provide an affordance for large scopes of awareness because it is easy to see everyone in the space [4]. Even so, managing awareness of many people at one time can be difficult or overwhelming for humans, although it has been found that in certain goal-directed social activities, humans are capable of *holoptism*, or "the ability for any interacting co-agent to perceive the dynamics of the whole interactive system" (*system awareness*) [6].

Observations

We observed that when humans are socially dancing, they tend to focus on coordinating their movements with one other individual (*dyad*) or a small group of co-located partners (*small group*). In the case of LuminAI, these partners can be either human dancers or virtual agents. Currently, the artificial dancers in LuminAI are only aware of the movements of the person closest to the Kinect sensor associated with that agent.

Reflections

Something we had not considered prior to conducting this analysis is that artificial agents can have a larger range of possible scopes of awareness than the typical human, depending on how they are programmed. An artificial agent might only be capable of awareness of itself or of a single person (*individual*), restricting its scope of awareness below that of most humans. Agents might also have the capability to perform superhuman feats, such as being aware of the movements of everyone on a large dance floor (*crowd*). This is a different problem from holoptism in that it requires the agent to be aware of the individual motions of each dancer in the crowd rather than the dynamics of the dance floor system as a whole.

Future work could investigate how it would affect the experience if the agents could factor in the movements of everyone dancing within a certain radius (like a human might do) or even the movements of an entire crowd. Highlighting the shadow(s) of participants whose movements the agent is currently aware of could emphasize these changing scopes.

Control/Collaboration

The third theme, control, deals with who has power over the technology and how the technology is manipulated. This changes when an AI agent is introduced as part of the system, as the assumption that a human is controlling the technology is no longer necessarily true. If the technology plays an equal role in the control of the human-agent interaction, the control scheme shifts to *collaboration*. This collaborative dynamic raises several additional questions. How do the agent and human negotiate control of the interaction? How does the human make sense of the collaborating agent's actions (and vice versa)?

Background

The conceptual space here deals with how to collaborate with the system and understand the capabilities of the system. Users have to experiment with the system to gradually make sense of its functionality [17, 21]. These experimental interactions help users build a mental model of how the system works, including a conceptual space of possible interactions (i.e. the system's affordances). Once users establish a preliminary model of how the agent works, they can begin creatively collaborating with it. This creative collaboration involves mutually coupled interactions during which both parties are influenced by the actions of their partner. This process of co-creating meaning through

interaction is referred to as participatory sense making in the cognitive science literature [17, 19].

Based on this concept, our taxonomy divides collaboration into *understanding how to collaborate with the system* and *understanding the system's capabilities*. Understanding how to collaborate with the system deals with making sense of the current interaction. *Exploratory action* and *observation* are both part of this process of coordinating collaboration through real-time feedback [17].

Understanding the system's capabilities deals with making sense of the system's functionality. This category is divided into four sub-categories that reflect different levels of understanding: understanding that the system can *mimic* your actions, understanding that the system can *augment* your actions, understanding that the system can *learn* from your actions, and understanding that the system can come up with *novel* actions. These four levels of understanding are based on the LuminAI system capabilities, which were inspired by strategies for improvisation in jazz music [32].

Observations

People did not in general have difficulty understanding how to collaborate with the public LuminAI installation. None of the interviewed participants indicated any difficulty learning how to physically interact with the installation. In addition, only 14 people entered the dome and did not physically interact with the system at all. This non-interaction could have been caused by confusion, but it could also be the result of disinterest, social anxiety, or outside factors such as social obligations or time constraints. While participants understood how to interact with the system, interview data indicates that most if not all participants did not fully understand the capabilities of their artificial collaborator. Three participants did not understand that they were interacting with an artificial agent. Of the participants that did recognize that they were interacting with an AI agent, few recognized that the agent had the ability to mimic (5), augment (4), learn (1), or perform novel actions (0). This underestimation of the capabilities of an AI agent is known as the TALE-SPIN effect [40].

Participants offered several reasons why they failed to understand the capabilities of their collaborator. Five participants noted that the previously mentioned conflicts in digital space made the agents' actions unclear (see *Conflicts*). Seven felt that issues in agent resolution and definition impaired their recognition. Two others stated that they were having too much fun to care about what they were collaborating with.

Despite the failure to understand the capabilities of their collaborators, there was clear evidence in both the video and the interviews that participants engaged in the participatory sense making process. Users engaged in exploratory action by pointing or waving at the agents. One participant said in their interview, "We are all playing with our body to see what the system does." Another participant

was recorded waving her arms at the agent and then turning to her friend and asking “Did anything happen? Is it working now?” The subsequent moment of recognition when she saw the agent responding and built a mental model of the situation was clearly visible. Another participant mentioned this type of “moment of recognition” in their interview. Many (11) participants formed hypotheses about what the system was doing as they were interacting with it and discussed these in their interviews. Theories ranged from the idea that the silhouette was supposed to “grab” or “pet” the agent to the notion that the color of the agent was related to the color of the participant’s clothing.

Reflections

It is possible that the agent’s scope of awareness inhibited understanding. For example, in one interaction, there were three humans interacting with one agent, and it was unclear which human the agent was learning from and responding to. Future iterations of this project may benefit from considering ways to augment the installation in order to better communicate agent capabilities to the participant without reducing the enjoyable aspects of the experience.

Interaction

The fourth theme, interaction, deals both with how humans interact with each other and how they interact with the system. This section of our taxonomy is divided into *mode of interaction* and *interaction level*. Mode of interaction deals with the way participants engage with the system and each other, and interaction level deals with the level of engagement of the participant.

Mode of Interaction

Observations

Three different modes of interaction with the system were observed: *verbal* (e.g. dialogue), *nonverbal* (e.g. physical movement), and *remote* (e.g. taking a photo of the interface and sharing it on Facebook). Verbal interactions occurred exclusively in human-human interactions (since the agents/silhouettes were unable to speak). While the audio we recorded did not capture dialogue for the most part (the dance music overpowered participant voices), it was visually apparent when group members were speaking to each other. Participants frequently engaged in dialogue before an interaction and during pauses in interaction.

Nonverbal interactions occurred both between humans and between humans and agents. These types of interactions can be categorized according to *rhythm* and *repetition*. Both *rhythmic* and *arrhythmic* interactions were observed, as well as *cyclic* and *acyclic* interactions.

Rhythmic interactions consist of strong gestures with a defined periodicity. Arrhythmic interactions can be both cyclic and acyclic, but they do not have a defined periodicity. *One-shot* interactions are a particular type of arrhythmic interaction that involves simple non-repeated physical gestures such as waving an arm once, kicking a leg out, or jumping.

Most participants who entered the dome tried to interact with the agent initially using a one-shot gesture (14 participants did not interact at all and 10 transitioned immediately into fluid rhythmic/arrhythmic dance). It is worth noting that one-shot gestures were not common in human-human interactions, but very common in human-agent interactions. This could be because one-shot gestures are seen as an initial means of testing or making sense of the system, in line with the idea of exploratory action and observation as key elements of the collaboration or participatory sense-making process [17].

Finally, remote interactions involved humans taking photos or videos of both the artificial agents and of other participants interacting with the artificial agents. In the video data, nine different people took photos or videos of their friends interacting with the system. This type of performance and recording allows remotely located participants to participate in the interaction.

Reflections

Certain modes of interaction that we observed are specific to movement improvisation (e.g. rhythm). However, designing to facilitate multi-modal interaction is important for any socially interactive system, and considering how each mode manifests itself in a particular system is also important to understanding interaction dynamics.

Interaction Level

Background

Interaction levels deal with how humans interact with each other and the system differently depending on their level of engagement and/or focus. Different frameworks for analyzing interaction levels have been presented in the literature. Several of these are presented below.

Ludvigsen develops a framework that identifies four levels of interaction that occur with social interfaces, based on Goffman’s sociological research: distributed attention, shared focus, dialogue, and collective action [26]. *Distributed attention* refers to a group of people who are sharing the same space, but who are not focused on any shared interface or activity. *Shared focus* is when a group of people is focusing on a single activity/interface within a space. *Dialogue* occurs when people engage in a shared activity, and *collective action* occurs when people (*human teamwork*) or agents (*agent teamwork*) work together to achieve a shared goal.

Brignull and Rogers define a framework for analyzing public interface interactions that is similar to Ludvigsen’s. Their framework consists of three different activity spaces: *peripheral*, *focal*, and *direct* [9]. These activity spaces can be mapped to Ludvigsen’s activity levels of distributed attention, shared focus, and dialogue, respectively. Since Ludvigsen’s framework has the additional level of collective action, this section of our taxonomy is based on Ludvigsen’s framework. See *Transitions* for observations of and reflections on these different activity levels.

Types of Dialogue

Observations

Once users transitioned from shared focus to dialogue, there were four different types of dialogue that were observed in the video recording: *parallel*, *human-human-agent*, *individual*, *internal*, and *human-human*. Any interaction in which there were multiple people dancing with agents or silhouettes but not engaging with each other was considered to be parallel dialogue. This was the most common type of interaction, a finding that is consistent with previous public interface installations [31].

Human-human-agent interaction was also very common, occurring in 15 different groups. These types of interaction occurred when two or more humans coordinated their dance moves both with each other and with one or more virtual agents and/or silhouettes. The frequency of this type of interaction reinforces the potential value of expanding the scope of agent awareness from one participant to a small group of participants.

Individual dialogue was infrequent, but it did occur. Any interaction between a person and an agent when the dome was otherwise empty was considered to be individual dialogue. In the video recording, there were four instances of individual dialogue with an average duration of 31 seconds. Three of these interactions involved participants who had previously interacted with the installation (in the presence of other participants) and later returned on their own. Two instances involved the same individual.

Another type of interaction observed was when a person appeared to be dancing by him or herself without interacting with an agent or another human (despite the presence of both agents and other humans in the dome). This type of dialogue is classified as *internal* in the taxonomy. Like individual dialogue, internal dialogue was infrequent but it did occur. Four different participants were observed dancing by themselves while surrounded by other humans and agents. These participants did not appear to coordinate their actions with anyone else in the dome.

Human-human interaction, or interaction between two or more humans without engaging with an agent, occurred when the dome became too crowded for everyone to dance with an agent. More surprisingly, it also occurred with some groups when there was plenty of digital and physical space. For instance, one couple did a ballroom dance together while they were alone in the dome, during which they paid no attention to the agents or silhouettes. There were five different instances of this type of interaction, suggesting that the dome itself provides an affordance for dancing with or without the presence of an AI agent.

Reflections

The occurrence of both individual and internal dialogue is particularly surprising because dancing without human companions at a party is often considered to be socially unacceptable. The presence of individual dialogue in

particular is contrary to previous findings that people hesitate to initiate interaction with public interfaces if others are not already engaging with them [9, 31]. This suggests that the dome, unlike other public interfaces, allows and even encourages people to engage in what might otherwise be considered *occult interaction*, or social interaction that is considered unnatural and unpleasant [20]. This idea is reinforced by the interview data; one participant said that the installation “made dancing less awkward” and another stated that dancing with the agent “made me want to get crazy and do things I wouldn’t normally do”.

Transitions

The final theme, transitions, can refer to transitions between spaces, between physical and digital environments, or between levels of interaction. Since the territoriality theme also deals with transitions between physical and digital spaces, we focus this section of our taxonomy on transitions between interaction levels. The first part of this section will deal with how to facilitate transitions between interaction levels, and the second part discusses transitions that were observed in the public installation of LuminAI, focusing on three of the twelve transitions represented in the taxonomy (*Distributed Attention* → *Shared Focus*, *Shared Focus* → *Dialogue*, and *Dialogue* → *Collective Action*, with an additional discussion of *backwards transitions*). We focus only on these three both to avoid recounting redundant observations and because these transitions are the most relevant to designing to facilitate social interaction.

Designing to facilitate transitions

Background

The term *situational interaction mobility* is used to “describe...the change in level of social interaction in the framework, and how well a service, product or installation supports this change in engagement” [25]. Design choices can create both barriers to and facilitators of situational interaction mobility. As a result, it is important to consider how to design to encourage upwards transitions in the level of participant engagement.

Barriers to interaction can be both physical and mental. Lack of physical access points can inhibit transitions from physical to digital or transitions into the interaction space [33]. Having a large empty space around an interface can also intimidate people and prevent them from approaching [9]. Social embarrassment is another key factor that prevents people from interacting with public interfaces [37]. People’s hesitation to interact publicly can be attributed to the fear of participating in occult interaction [20].

Facilitators of transitions are also important to take into consideration. *Attractors* are aspects of a system that inspire participants to initiate interaction [18], and they can also facilitate transitions between interaction levels. One such attractor is the “honey-pot effect”, which is an increase in the number of people standing in the vicinity of the interactive public display [9]. This effect creates a “sociable buzz” in the area and increases social interaction [9].

Seeing people interact with a system not only encourages the transition from distributed attention to shared focus, but it can also encourage passersby to interact with the system [31]. Other attractors include visuals that are viewable from a distance and lightweight, clearly visible interaction dynamics [9]. These generate curiosity, reduce intimidation, and encourage transitions from both distributed attention to shared focus and shared focus to dialogue [9].

Distributed Attention → Shared Focus Observations

The first interaction level transition occurs when participants shift their focus to a shared object or system (in our case, the LuminAI installation). Unfortunately, the LuminAI video data only captured interactions inside and at the edge of the dome; therefore, there was not a lot of information about indicators of distributed attention or the transition from distributed attention to shared focus. Some limited video data of participants outside of the dome was captured from one camera angle; from this, indicators of the transition from distributed attention to shared focus appear to include *physically approaching* the installation, *pausing to watch* the installation, or performing a *gesture of interest* towards the installation, but more rigorous observation and analysis is needed to confirm this.

Reflections

Several different aspects of the installation could have attracted attention and helped to facilitate the transition from distributed attention to shared focus. The otherworldly or awe-inspiring nature of the dome may have drawn people to look more closely at the installation, a concept which has its basis in literature on museum exhibit design [2]. In addition, projections visible on the outside of the dome may have served as an attractor, in line with the design principle that large visual displays attract participants [9]. Finally, seeing people interact with the exhibit encouraged others to stop and watch. Far more people paused outside to look at the installation when others were inside dancing (21) than when the dome was empty (2). This finding is in line with previous research on public interfaces [9, 31].

Shared Focus → Dialogue Background

In order to transition from shared focus to dialogue, participants can either *socially negotiate* their approach by queuing up or waiting their turn [31] or *directly initiate* interaction.

Observations

Most users initiated interaction with LuminAI via a direct physical gesture (i.e. waving arms or dancing), and did not appear to socially negotiate their approach (i.e. *queuing up*, *turn-taking*).

Due to the partially enclosed nature of the dome, curious attendees were forced to move physically closer to and sometimes even enter the dome in order to see what was going on. Once inside of the dome, attendees quickly

transitioned from shared focus to dialogue with the system. As mentioned earlier, only 14 people who entered the dome did not physically interact with the system at all.

The social aspects of the installation also encouraged transitions from shared focus to dialogue. Dancing with peers (rather than experts) reduced social embarrassment; one participant liked that everyone dancing was “at the same baseline” and explained, “It’s not about judging or whether we know how to dance, just about playing with body movements...that’s how dancing should be”. Social encouragement facilitated transitions as well. In the video, seven users verbally or physically encouraged their friends to interact with the system. In addition, four different individuals/small groups were recorded interacting with the installation after watching another individual/small group interacting first. The actual number of participants who engaged in this type of activity may be larger since the camera only recorded actions inside of the dome and people could have been observing from outside before joining in interaction. For example, at one point, two participants entered the dome; one immediately began engaging with the agent, but the other participant was more hesitant. The second participant felt more comfortable and began to interact after watching the first participant dance. In another interaction, an adult and a child entered the dome, and the adult illustrated how to interact with the system by pointing and dancing with the agent. After observing for a while, the child joined in the interaction.

Reflections

The direct initiation of interaction that was observed could be a result of turning on the silhouettes in addition to the agents; once this was done, there was a digital figure that responded to every person, suggesting that anyone could participate in the installation without waiting for an available agent. This could also be due to the “occasion” or social construct of dance floors, where people are used to dancing all at once without taking turns [20].

The data, examined through the lens of territoriality, supports our original hypothesis that the dome (in contrast with the original 2D installation) provides an affordance for social interaction. Interaction and non-interaction territories are clearly defined, with the dome itself serving as the seam or liminal space between the two. Making the liminal space a physical barrier creates an involvement shield between participants and nonparticipants, reducing the risk of social embarrassment [20]. This hypothesis is also supported by interview data; one participant said that they felt “protected in the bubble”; another mentioned that the dome made them feel “enclosed and protected” and a third participant said that “the dome shape really invited me to dance”.

The clearly defined interaction territory also allowed participants to perform in a stage-like environment without a high barrier of entry or risk of social embarrassment. This enabled participants to engage in performative play and creatively explore space and bodily gesture. When

interviewed, one participant said, “We were not sure if we were allowed to be in the dome. We were waiting at first, and then people followed us in.” This indicates that the dome delineates a stage-like space.

Dialogue → Collective Action

Background

The final transition is between dialogue and collective action. This transition involves moving from undirected social interaction to a goal-oriented activity such as playing a game or managing a conflict [31]. Such a goal can either be *explicitly stated* (via a verbal conversation) or *emerge* out of collective interaction. The idea of emergent goal-creation is an example of *collaborative emergence*, a term used to describe phenomena that “result from the collective activity of social groups” [34]. The outstanding question is how to identify when a goal has emerged.

Observations

There were only a few scattered instances in which users engaged in collective action with other participants in order to manipulate LuminAI, and as a result, further study is needed to identify clear behavioral markers indicating when an undirected activity has transformed into goal-oriented collective action. Two LuminAI participants described trying to play games with the agent in their interviews (i.e. trying to encapsulate the agent within the silhouette), but it was unclear whether or not these games involved teamwork with other humans. During three different interactions in the video recording, couples or friends joined hands or hugged each other in order to get their silhouettes to merge into one, and then experimented with how they could collaboratively move their joined silhouette.

While some users did reach the collective interaction level, this was a rare occurrence, and most participants remained in the dialogue stage throughout the course of their interaction with the system. Even when users achieved collective action, it was only reached incidentally; there is nothing built into the system that encourages users to transition from dialogue to collective action.

Reflections

Collective action is the most socially engaging level of interaction, so designing to encourage participants to reach this final level could be a valuable improvement [25]. It is also worth considering the possibility of collective action between agents or between humans and agents.

Backwards Transitions

Background

It is also natural for participants to transition from a higher level to a lower level of interaction as they cease interacting with the installation for various reasons. Previous work has found that social factors such as an “obligation to leave” play a large role in causing disengagement with public interfaces [3].

Observations

In the video data, several different causes for backwards

transitions emerged including *boredom*, *confusion*, *pressure from outside forces* (e.g. another participant, time obligations), *lack of physical space*, and the *desire to discuss* the interaction with a friend or observe others interacting. Certain behavioral markers were observed during backwards transitions in the video. Episodes of teamwork (e.g. two users working together to get their silhouettes to merge give up and instead dance in parallel with their agents) trigger transitions from collective action back to dialogue. This could manifest itself in several ways, including the end of a verbal conversation between participants, or the end of a coordinated physical “conversation” between participants or agents (observed in video). A transition from dialogue to shared focus was indicated by the cessation of interaction with the system (e.g. a participant steps back to observe the system). Finally, a participant turning away from the system or shifting his or her focus to another object (e.g. a participant walks away to get a drink) indicated a transition from shared focus to distributed attention.

Reflections

Designers should concern themselves not with eliminating backwards transitions, but with extending the period of time a participant spends at a particular interaction level before transitioning down.

CONCLUSION

In conclusion, we developed LuminAI, a socially interactive installation in which humans are able to improvise dance movements in collaboration with each other and virtual agents that are projected onto the liminal space of a geodesic dome. This installation was presented to the public at a local arts event, and interview and video data were collected during the event. This data was then analyzed and this analysis, in combination with a synthesis of interdisciplinary literature, was used to create a taxonomy to guide the design of socially interactive human-agent systems. This taxonomy helped us to identify areas where the installation was successful (e.g. facilitating transitions between different interaction levels), where it needs improvement (e.g. conflicts in digital space), and conceptual spaces we have yet to explore (e.g. agent awareness and collective action). In this paper, we have contributed a detailed framework for analyzing socially interactive installations involving both humans and artificial agents in addition to developing an art installation that explores the social fabric of the dance floor and human-agent interactions.

ACKNOWLEDGMENTS

We would like to thank T. Lang Dance Company, Jessica Anderson, Lauren Winston, Ju-Hwan Lim, Sasimon Viriyayuthakorn, Lisa Li, the Ivan Allen College of Liberal Arts (Georgia Tech), and The Goat Farm Arts Center for their efforts and support of this work.

REFERENCES

1. Baris Akgun, Kalesha Bullard, Vivian Chu and Andrea L. Thomaz. 2014. An HRI Approach to Learning from Demonstration. In *AAAI Fall Symposium Series*, 10-12.
2. Leslie Bedford. 2014. *The Art of Museum Exhibitions: How Story and Imagination Create Aesthetic Experiences*. Left Coast Press, Inc.
3. Ben Bengler and Nick Bryan-Kinns. 2015. "I could play here for hours.." (thinks the visitor and leaves): Why People Disengage from Public Interactives. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C '15)*, 177-180. <http://dx.doi.org/10.1145/2757226.2764548>
4. David Benyon. 2014. *Spaces of Interaction, Places for Experience*. Morgan & Claypool.
5. Anne Bogart and Tina Landau. 2003. *The Viewpoints Book: A Practical Guide to Viewpoints and Composition*. Theatre Communications Group.
6. Jérôme Bourbousson and Marina Fortes-Bourbousson. 2016. How Do Co-Agents Actively Regulate Their Collective Behavior States? *Frontiers in Psychology* 7: 1732.
7. Richard E. Boyatzis. 1998. *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage Publications.
8. Mason Bretan and Gil Weinberg. 2014. Chronicles of a Robotic Musical Companion. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 315-318.
9. Harry Brignull and Yvonne Rogers. 2003. Enticing people to interact with large public displays in public spaces. In *Proceedings of INTERACT '03*, 17-24.
10. William Buxton. 2009. Mediaspace—Meaningspace—Meetingspace. In *Media Space 20+ Years of Mediated Life*, Steve Harrison (ed.). Springer, London, UK, 217-231.
11. Erin A. Carroll and Celine Latulipe. 2009. The creativity support index. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*, 4009-4014. <https://doi.org/10.1145/1520340.1520609>
12. Crystal Chao and Andrea L. Thomaz. 2011. Timing in Multimodal Turn-Taking Interactions: Control and Analysis Using Timed Petri Nets. *Journal of Human-Robot Interaction* 1, 1: 1-16.
13. Kate Compton and Michael Mateas. 2015. Casual Creators. In *Proceedings of the Sixth International Conference on Computational Creativity*, 228-235.
14. Mihaly Csikszentmihalyi. 2014. *The Systems Model of Creativity: The Collected Works of Mihaly Csikszentmihalyi*. Springer.
15. Nicholas Davis, Margeaux Comerford, Mikhail Jacob, Chih-Pin Hsiao, and Brian Magerko. 2015. An Enactive Characterization of Pretend Play. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C '15)*, 275-284. <http://dx.doi.org/10.1145/2757226.2757254>
16. Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Lisa Li, and Brian Magerko. 2016. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI '16)*, 196-207. <http://dx.doi.org/10.1145/2856767.2856795>
17. Hanne De Jaegher and Ezequiel Di Paolo. 2007. Participatory sense-making. *Phenomenology and the cognitive sciences* 6, 4: 485-507.
18. Ernest Edmonds, Lizzie Muller, and Matthew Connell. 2006. On creative engagement. *Visual Communication* 5, 3: 307-322.
19. Thomas Fuchs and Hanne De Jaegher. 2009. Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences* 8, 4: 465-486.
20. Erving Goffman. 1963. *Behaviour in Public Places: Notes on the Social Organisation of Gatherings*. The Free Press.
21. Mikhail Jacob, Alexander Zook, and Brian Magerko. 2013. Viewpoints AI: Procedurally Representing and Reasoning about Gestures. In *Proceedings of DiGRA 2013: DeFragging Game Studies*.
22. Mikhail Jacob, Gaëtan Coisne, Akshay Gupta, Ivan Sysoev, Gaurav Gav Verma, and Brian Magerko. 2013. Viewpoints AI. In *Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 16-22.
23. Janet Kolodner. 2014. *Case-based reasoning*. Morgan Kaufmann Publishers, Inc.
24. Todd Lubart. 2005. How can computers be partners in the creative process: Classification and commentary on the Special Issue. *International Journal of Human-Computer Studies* 63, 4-5: 365-369.
25. Martin Ludvigsen. 2005. Designing for Social Use in Public Places - a Conceptual Framework of Social Interaction. In *Proceedings of Designing Pleasurable Products and Interfaces Conference (DPPI '05)*.
26. MadMapper. 2017. MadMapper: The video Mapping Software made by Artists for Artists. Retrieved January 12, 2017 from <http://www.madmapper.com/>.
27. Ben Malbon. 1999. *Clubbing: Dancing, Ecstasy, Vitality*. Psychology Press.
28. Brian Magerko, Waleed Manzoul, Mark Riedl, Allan Baumer, Daniel Fuller, Kurt Luther, and Celia Pearce. 2009. An empirical study of cognition and theatrical improvisation.

In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, 117-126.
<http://doi.org/10.1145/1640233.1640253>

29. David J. Mendonca and William AI Wallace. 2007. A Cognitive Model of Improvisation in Emergency Management. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 37, 4: 547-561.
30. Marek P. Michalowski, Selma Sabanovik, and Hideki Kozima. 2007. A Dancing Robot for Rhythmic Social Interaction. In *Proceedings of the Second ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 89-96. <http://doi.org/10.1145/1228716.1228729>
31. Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. 2008. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 1285-1294.
<http://dx.doi.org/10.1145/1357054.1357255>
32. Jeff Pressing. 1988. Improvisation: methods and models. In *Generative processes in music*, John Sloboda (ed.). Oxford University Press, Canada, 129-178.
33. Yvonne Rogers. 2006. Moving on From Weiser's Vision of Calm Computing: Engaging Ubicomp Experiences. In *UbiComp 2006: Ubiquitous Computing*, 404-421.
https://doi.org/10.1007/11853565_24
34. Keith R. Sawyer. 1999. The emergence of creativity. In *Philosophical Psychology* 12, 4: 447-469.
35. Keith R. Sawyer. 2003. *Group creativity: music, theater, collaboration*. L. Erlbaum Associates.
36. Rob Saunders and Oliver Bown. 2015. Computational Social Creativity. *Artificial Life: Special Issue on Art, Creativity and Culture* 21, 3: 366-378.
37. Robyn Taylor, Pierre Boulanger, Patrick Olivier, and Jayne Wallace. 2009. Exploring participatory performance to inform the design of collaborative public interfaces. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*, 3721-3726. <http://doi.acm.org/10.1145/1520340.1520561>
38. Andrea L. Thomaz and Crystal Chao. Turn Taking Based on Information Flow for Fluent Human-Robot Interaction. *AI Magazine* 32, 4: 53-63.
39. Noah Wardrip-Fruin. 2009. *Expressive Processing: Digital fictions, computer games, and software studies*. The MIT Press.