Interaction-based Authoring for Scalable Co-creative Agents

Mikhail Jacob, Brian Magerko
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA USA
{mikhail.jacob, magerko}@gatech.edu

Abstract

This article presents a novel approach to authoring co-creative systems - called interaction-based authoring - that combines ideas from case-based learning and imitative learning, while emphasizing its use in open-ended co-creative application domains. This work suggests an alternative to manually authoring knowledge for computationally creative agents that relies on user interaction "in the wild" as opposed to high-effort manual authoring beforehand. The Viewpoints AI installation is described as an instantiation of the interaction-based authoring approach. Finally, the interaction-based authoring approach is evaluated within the Viewpoints AI installation and the results are discussed guiding development and further evaluation in the future.

Introduction

Within the computational creativity community, our research has focused on domains that are open-ended, artistically performative, improvisational, and co-creative between human and AI agent. co-creative AI agents that can succeed in these kinds of domains tend to be large-scale and knowledge-rich since they have to collaborate creatively on an equal footing with humans. Therefore, one of the key bottlenecks for developing co-creative agents has been the knowledge-authoring bottleneck. According to Csinger et al. (1994), the difficulty, cost, or delay in acquisition of expert instantial knowledge followed by its subsequent structuring and storage so as to enable efficient future utilization is often referred to as *the knowledge-authoring bottleneck*. In fact, the knowledge-authoring bottleneck has historically been a significant problem for the intelligent agent community in general and the computational creativity community in particular.

Many solutions have been proposed in the past to mitigate the problem. Case-based reasoning (CBR) approaches and machine learning approaches have utilized online case acquisition and data mining from corpora as fundamental methods for dealing with the knowledge-authoring bottleneck. Data mining approaches have faced a general lack of corpora for instantial or behavioral content within improvisational performative domains. Traditional CBR systems while learning from experience still require instantial content to be authored in the form of an initial case library. Learning by demonstration or observation can avoid these pitfalls, but traditionally require explicit training or teaching phases before they can be used in the final task.

Within the games research community procedural content generation (PCG) research has focused on developing algorithms to generate the instantial content that was once manually authored by expert designers. This has seen success with the development of procedurality-centric games such as Spore and Galactic Arms Race (Hecker et al. 2008; Hastings et al. 2009). However, PCG systems have yet to focus on generating behavioral content that is flexible enough to work in open-ended improvisational domains.

In contrast to the previous authoring approaches mentioned, this article describes a hybrid knowledge-authoring paradigm that combines case-based learning with learning by observation / imitative learning – called *interaction-based authoring*. Interaction-based authoring aims to i) minimize the authoring bottleneck while ii) ensuring that the subjective experience of interacting with the system is high quality and that iii) the computer collaborator supports equal creative agency (the extent to which a creative collaborator can take decisions, make choices, and affect co-creation). It proceeds to demonstrate the interaction-based authoring paradigm within an improvisational interactive art installation called *Viewpoints AI* (Jacob et al. 2013a) after comparing the installation to related work in the field. A brief updated system description is provided (c.f. Jacob et al. 2013b). Finally, the paper details an initial attempt to evaluate the interaction-based authoring approach instantiated within the Viewpoints AI installation and discusses the results as a guide for iteratively developing / refining the installation.

Interaction-Based Authoring

Interaction-based authoring is a hybrid approach to authoring instantial knowledge and control knowledge for co-creative interactive systems, combining case-based learning with imitative learning. While using an interaction-based authoring approach learning occurs over the lifetime of the full performance and not during an explicit training or teaching phase. This is done to boost participant motivation and engagement encouraging prolonged interaction with the agent thereby facilitating greater knowledge acquisition.

There are three main aspects to interaction-based authoring. First, case-based learning is used to index and store agent experiences in a reusable manner that can be utilized to drive future behavior or responses in general. Cases can be stored as input–output pairs (from the agent's perspective) with a process to map between inputs and outputs in order to use them interchangeably.

Second, an imitative learning / learning by observation system (Tomasello 2000) that can model the way a human partner responds to the agent's actions is utilized in order to interact with other partners in other interaction contexts. If the new partner's input action (from the new interaction context) is similar enough to an input action it has learnt a model for in the past, it can use that to select an output action. The agent takes the interactor's role in that case and responds, as they (presumably) would have.

Finally, an open-ended co-creative improvisational domain in which to situate the agent is required so that the participant or interactor is engaged and therefore motivated to teach the system for an extended period of time. The open-ended nature of the domain encourages exploration of the interaction space, increasing the coverage of the learning algorithms for future interactions. The co-creative and improvisational aspects of the domain emphasize the egalitarian nature of creative decision-making. They also encourage the user to further explore novel regions of the interaction space in the event that the system makes a 'poor' choice of response, thinking of it as an interesting offer that they hadn't considered rather than a mistake. The interaction-based authoring approach has been instantiated in an interactive improvisational human–AI art installation called *Viewpoints AI*. A description of the installation follows a brief account of related work.

Related Work

Technology has been contemporarily used to augment performances and art installations (Reidsma et al. 2006; Latulipe 2011; MacDonald et al. 2015). These pieces use performance technology as an integral part of their overall aesthetics and content of the artwork. However, these technologies have been subservient to human performers, with shallow knowledge, and / or a lack of clear collaboration between the machines and humans on stage.

Combining research in arts, AI, cognitive psychology and philosophy, the field of computational creativity has focused on many different creative domains (c.f. Boden 2003; c.f. Colton 2012). However, most traditional computationally creative systems assemble pre-authored content in novel combinations, without attempting to solve the knowledge-authoring bottleneck, leading to small systems with limited scope. In addition, many in the past have ignored creative collaboration or co-creativity focusing on systems that do not involve humans except as consumers or evaluators of the creative artifact or process.

Computationally co-creative systems on the other hand collaborate with humans in order to participate meaningfully in the creative process or outcome. Much work has been done on co-creative agents in the music improvisation domain (Thom 2000; Hsu 2008; Hoffman and Weinberg 2010). The Digital Improv Project virtual agents that could perform theatrical improvisation (O'Neill et al. 2011) and the Computational Play Project virtual agents that could play pretend with people using toys (Magerko et al. 2014) are examples of co-creative systems in other domains. Both however, required extensive pre-authored instantial content to produce improvisational behavior. The Digital Apprentice (a virtual collaborator for abstract visual / sketch art creation; Davis et al. 2014) is a co-creative system that closely resembles an instantiation of the interaction-based authoring approach.

Figure 1: The Viewpoints AI Installation

The Viewpoints AI Installation

The Viewpoints AI installation is a participatory interactive installation where a human interactor and a virtual agent – named *VAI* – collaborate to improvise movement-based performance pieces together in real-time. The installation (see Figure 1) is composed of a large translucent muslin projection screen that has a human-sized manifestation of VAI projected onto it from the front and the interactor's shadow cast onto it from the rear. This enables an occlusion-free juxtaposition of the interactor's shadow onto the projected virtual agent when their positions overlap. While the installation is highly participatory in nature and the experience of improvising is intrinsically tied to it, an audience can also watch the unfolding performance from the front of the installation.

Participants interact with the virtual agent behind the muslin screen while a Microsoft Kinect depth camera senses and records their movements. Recorded movements are analyzed systematically using a formal version of the Viewpoints framework, as described by Bogart and Landau (2005). Viewpoints is used in theatrical movement and the staging of scenes to focus on the physicality of action and analyze performance in terms of the physical Viewpoints of time (tempo, duration, kinesthetic response, and repetition) and space (shape, gesture, spatial relationship, topography, and architecture), as well as the vocal Viewpoints (pitch, dynamics, acceleration, silence, and timbre). The participant's movements are interpreted through a subset of the Viewpoints framework and are then responded to by the agent. The formalization of Viewpoints is thus used as a framework to represent and reason about movement.

The Viewpoints AI installation uses contrasting visual elements of light and shadow to showcase how the human participant and the virtual agent arrive at this liminal interaction space from two very different worlds. Visually, VAI is a glowing anthropomorphic character composed of a playful cloud of fireflies. The participant's crisp shadowed form is transported to the ephemeral 2D space between the two worlds through the medium of shadow theatre.

System Description

The Viewpoints AI installation is powered by an agent architecture that is conceptually composed of three modules – *perception*, *reasoning*, and *action*. Earlier versions of the system are described in Jacob et al. (2013a; 2013b). The following sections describe the agent architecture briefly, going into more detail where necessary to illustrate updated aspects of the system.

Perception

The Viewpoints AI agent architecture receives input from the depth camera as a frame of joint positions in continuous 3D space at a certain frame rate to get "joint space" gestures. It then discretizes the joint space gestures and derives additional information about them in real-time using a formalization of the Viewpoints framework to get discrete "predicate space" gestures. These two types of gestures are then sent along to the reasoning module.

Parsing Viewpoints Predicates The Viewpoints predicates that have been formalized to date make up a subset of the physical Viewpoints, including tempo, duration, and repetition, as well as parts of spatial relationship, topography, shape, and gesture. The current version of the installation has a general purpose machine learning toolkit (Hall et al. 2009) integrated within the agent architecture that classifies Viewpoints predicates using classifiers trained using supervised learning on expert movement-practitioner / dancer data. Adding new predicates to the system is as straightforward as training new classifiers with more data demonstrating or exemplifying that particular attribute or aspect of the Viewpoints framework. Emotional content of the performance portrayed through gestures are also classified through this supervised learning process.

The current movement analysis pipeline employs modular feature detectors for motion-based features (eg. vertical knee velocity, tangential knee acceleration, etc.) of the joint space gestures. These are then used to feed classifiers (with the specific classification algorithm chosen empirically according to classification performance). Training data for classification is obtained by collaborating with expert local movement-practitioners and dancers.

Turn-taking Model Turn-taking refers here to the process of naturalistically timing the use of the shared performance space so as to coordinate each other's (potentially overlapping or simultaneous) movements. This can be decomposed into the problems of how to best time the agent's movement turns coordinating with the interactor and how to segment a user movement turn or gesture. The first problem is solved by the interaction convention that the agent moves whenever the interactor does, either mirroring them (when they move arrhythmically) or improvising an original response to their movements (when they perform rhythmic repeated movements). The second problem is discussed below.

In the current version of the installation the agent tries to detect a beat to the interactor's movement (helped by playing dance music during the interaction) and segments their gestures using the detected beat. It does this by creating a set of 1D motion vector-based local beat detectors for each moving joint. These report possible joint-level candidate beats by looking at the half period of the joint motion. When candidate beats are repeated multiple times, they are confirmed and reported to a global tracker. The global tracker then chooses a candidate local beat as the global beat, which is then used to segment the movement at the start and end of the beat duration. Additional trimming of the segment is done so that the start and end are the same.

Reasoning

Segmented gestures in both joint and predicate space are sent to the reasoning module for the agent to determine an appropriate response gesture. Joint space gestures are then stored in a gesture library in exemplar clusters, each cluster having a universally unique identifier number (UUID). These clusters are produced through an approximate gesture recognition algorithm using a content vector of aggregated versions of the same set of motion-based features used earlier in Viewpoints predicate classification. This is done in order to find patterns in interactions and cluster similar gestures together. It is a simplification of the hard problem of online matching in real-time of an input gesture to one (or potentially none) out of a potentially unbounded set of gestures without prior training of any sort. The corresponding predicate space gesture is then sent to a Soar agent (Laird 2012) for further processing in order to choose a response gesture. This case-based learning is a key mechanism within the Viewpoints AI installation that helps it instantiate interaction-based authoring.

Response Strategies The Soar agent has a set of strategies for selecting responses to the input gesture that are then output to the action module. These strategies are selected amongst using pragmatic and aesthetic rules for agent behavior. The response strategies were chosen using an analogy to methods that jazz improvisers use to respond to offers from fellow musicians. For example, repetition is important for establishing a motif, signaling understanding or acceptance of a communicative intention, signaling which performer is being lead by another, etc.

The most important response strategy, which forms the lynchpin of the interaction-based authoring approach, is the application of observationally learned input–response gesture pairs. The agent observes the collaborator's response to its action and builds an association with parameters to control its application. The use of observed action response association leverages the collaborator's more advanced reasoning faculties in order to respond to some other interactor in another context.

For example, when the agent learns to associate "waving" gesture inputs to "bowing" gesture responses by watching the collaborator execute "bowing" responses to its own "waving" gestures, it can respond using the learned association of "waving" and "bowing" gestures when a new interactor "waves" at the agent. Of course in this example, "wave" and "bow" gestures are actually clusters of gestures with corresponding IDs to which the actual input gestures match approximately (as mentioned earlier) – no semantics of the words "wave" or "bow" are implied to be understood by the agent.

A key assumption that the input response association is based on is that the interactor's response is always related to the previous gesture from the agent and that there is always some reason behind it. Both of these could well be false, if the interactor gets bored and tries something completely new for example. However, associations that are seen more often are given positive reinforcement helping to weed out weaker associations. This role-taking process forms the key mechanism for learning by observation and imitative learning within the Viewpoints AI installation that helps it instantiate interaction-based authoring.

Another response strategy is the selection of emotional reflex reactions to emotional content portrayed in the gestures. There is an "emotional algebra" authored in the system that responds according to a commonsensical set of rules (e.g. responding to angry input gestures with angry or fearful responses). This emotional algebra is rigid and uncomplicated yet enables a simple short-circuit reflex response system to quickly respond to portrayed emotionally salient content within gestures.

An important response strategy is for the system to mimic the interactor's input gesture back to them. Mimicry / repetition is important in facilitating smoother interactions between people (Behrends et al. 2012). In contrast, a (trivial) response strategy involves performing no response at all, though this promotes a sense of uncertainty and is thus discouraged unless as a last resort.

Another response strategy is for the agent to consider an existing gesture and transform it. This creates a variation of that gesture using dimensions or aspects of the Viewpoints framework (eg. faster in tempo, smoother in movement, adding repetitions, etc.). In addition, the system can use acontextual functional transforms to add variety in the enacted form of the gesture, such as reversing the direction of movement, changing the limb in which movement takes place, etc. See Jacob et al. (2013a; 2013b) for more detail.

A final response mode is for the agent to consider past experiences from its episodic memory and choose a similar gesture to bring into the new interaction context. This is achieved using Soar's episodic memory partial graph matching capabilities in order to approximately match the Viewpoints predicates of gestures and / or the direct predicate space representation of their movements from other interaction contexts to the current interaction context. This is valuable to inject novelty into the current interaction context. It uses the lower dimensionality (and higher level of abstraction) of the Viewpoints predicate space to pick a gesture that is roughly midway on a scale of novelty (from completely identical to absolutely novel). This episodic retrieval process is a key mechanism for case-based learning within the Viewpoints AI installation that helps it instantiate interaction-based authoring. Viewpoints predicates form the index vocabulary for the case-based retrieval. It should also be noted that this particular response strategy introduces novelty to the creative experience, balancing the predictability of other response strategies such as the application of observationally learned patterns.

Action

The action module receives both predicate and joint space gestures from the reasoning module and proceeds to create the suitably transformed and rendered virtual agent embodiment procedurally. The Viewpoints predicates associated with the gesture being performed directly affect the visualization (e.g. the energy of the agent's movements control the colors of the agent). The visual embodiment of VAI is an anthropomorphic figure with a body composed of glowing particles that keep to the bounds of the figure while flying around probabilistically. In the current version of the installation, the agent also has a region around the chest of a corresponding interactor that glows with a diffuse red colour in time to a rhythm if the agent has detected the user moving to a beat. This has the visual effect of a glowing heart beat that rises and falls with the interactor's movements. This was also designed to serve as a subtle form of coordination between the two collaborators. For more detail see Jacob et al. (2013a; 2013b).

Interaction-Based Authoring Beyond the Viewpoints AI Installation

The Viewpoints AI installation instantiates the interaction-based authoring approach to acquire knowledge from interactors while attempting to provide a high quality subjective experience to the interactors and support their creative agency. It does this through knowledge acquisition of two kinds. Firstly the case-based learning component stores all gesture content it has seen or experienced in episodic memory. Secondly it learns how to use these gestures to respond to people by learning interaction patterns or pairs of gestures from observing people and then imitating their actions in a novel context. Finally the installation is situated in a co-creative performative domain so that there is a low bar for meaningful collaboration as well as to encourage exploration of the interaction space due to player engagement and acquire more knowledge as a result. The approach differs from others by attempting to provide a full-fledged co-creative experience right from the outset without requiring explicit training or teaching phases.

The approach can be extended beyond the movement-improvisation domain to increase the scalability of other co-creative agents as well. The instantial gesture content that the system learns using case-based learning can be generalized to other types of response content, for example strokes on a canvas or notes played on a synthesizer. Imitation learning in turn can also be used to learn more general response control knowledge. For example, the system could learn patterns of strokes on a canvas or sequences of notes. Currently the Viewpoints AI installation only does a first order pairwise learning of gestures, however that could be extended to higher order sequences of patterns.

Evaluation Methodology

The following sections describe an initial effort to evaluate the success of the installation in addressing three main research questions. **RQ1**: Can the interaction-based authoring approach minimize the authoring bottleneck? **RQ2**: Can usage of the interaction-based authoring approach create high quality subjective experiences using the system? **RQ3**: Can systems built with the interaction-based authoring approach support collaboration with equal creative agency (the extent to which a creative collaborator can take meaningful decisions, make meaningful choices, and affect the co-creative process or product)?

RQ1 was evaluated with formal analysis of *authorial leverage* (Chen et al. 2009) as an initial attempt. More detailed, pragmatic testing is required next. For the analysis, three cases were compared: 1) a purely mirroring version of the installation where the agent would only mirror the movements of the interactor but not respond in any other way, 2) a version of the installation with a pre-authored tree of 'plot points' (pairs of input gestures and agent responses) of arbitrary length and branching factor, and 3) the full Viewpoints AI interactive art installation.

The RQ2 and RQ3 were evaluated using empirical quantitative and qualitative methods in a pilot study (sample size of 10). For the empirical evaluation, three different experimental conditions were used. Condition 1 had only mirroring of interactor movement as our baseline for comparison. Condition 2 had mirroring of interactor movement along with random movement responses, selected from a library of prerecorded movements, when the participant was performing rhythmic repeated movements. Finally, condition 3 had the full response capabilities of the agent available to respond whenever the interactor was making rhythmic repeated movements. The order of the experimental conditions was also randomized each time. In each case, participants interacted with the experiment for 3 minutes, filled out two surveys administered online, and then debriefed with a semi-structured interview. RQ2 was evaluated using a set of validated survey instruments measuring *system usability*, *flow*, and *enjoyment* of the installation (Brooke 1996; Jackson et al. 2008; Vorderer et al. 2004). RQ3 was evaluated using a set of validated survey instruments measuring the *creativity support index* (CSI) and *effectance* of the installation (Cherry and Latulipe 2014; Klimmt et al. 2007). The individual scales (excluding the CSI) were administered online as part of the IRIS Measurement Toolkit (Vermeulen et al. 2010). The CSI had responses on a 7 point Likert scale while the IRIS Measurement Toolkit used a 5-point Likert scale.

Results

Formal Analysis

The interaction-based authoring approach was designed primarily to address the knowledge-authoring bottleneck. Therefore the results of the formal analysis directly estimate how much of an improvement is achieved using this approach for acquiring knowledge within a co-creative agent in the movement improvisation domain. For the three experimental conditions described earlier (as with most existing literature in the field) only the variability was used as a factor for quality of the user experience. Therefore authorial leverage was calculated as the ratio of the number of unique experiences (variability) to the number of authorial inputs involved in creating the system. In addition, a few assumptions were made during the calculation. 1) In order to compare the Viewpoints AI installation variants to existing interactive narrative literature, the notion of plot points was loosened to represent sequences of human – agent movements or gestures. 2) The authorial inputs were considered to be the sum of the number of gestures that were authored prior to the start of the calculation in addition to any manually authored transition rules between them or between interactor gestures and agent responses.

For the first condition evaluating the purely mirroring agent, it was assumed that both interactor and agent movement responses were occurring simultaneously. Thus a plot point would represent the interactor gesture and the same agent gesture performed simultaneously. Therefore the same sequence of $\mathbf{N}$ interactor gestures input to the system would always return the same sequence of $\mathbf{N}$ gestures back as responses. The authorial leverage is thus nearly infinite since there is almost no prior manual authoring of instantial content (authorial inputs near zero).

For the second condition with the pre-authored branching tree of input gesture and agent response pairs, a tree of average branching factor b and depth d would have at most a total of $\mathbf{(b \times (b^d)-1) \,/\, (b-1)}$ nodes or loose analogs to plot points. Also, such a tree would have at most $\mathbf{(b^d)}$ linear paths through it from root to leaf node representing unique experiences. Therefore, the authorial leverage is roughly $\mathbf{(b^d) \times (b-1) / (b \times (b^d)-1)}$. This function has an asymptotic upper bound of $\mathbf{1}$ given any $\mathbf{b}$ or $\mathbf{d}$.

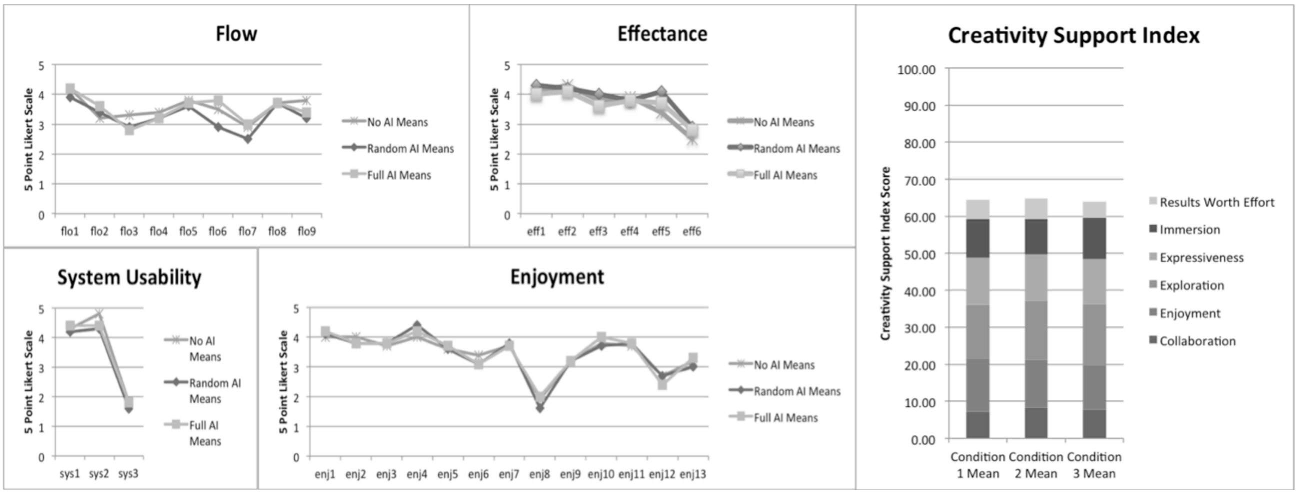

Figure 2: Results of Pilot Empirical Study. For descriptions of specific questions refer provided citations.

Finally, the third condition with the full installation active has the capability to select responses dynamically based on the reasoning processes mentioned earlier. For the first condition, the only possible response was to mirror the interactor's input gesture simultaneously. In the full installation that capability is present (though mimicking not mirroring) in addition to various other responses possible. Therefore the number of unique responses possible for a specific input gesture can only be higher.

For the same **N** input gestures as in the first condition, the number of possible agent responses would be **RN**, where **R** is the total number of unique responses to any one input gesture given a set of response strategies (rather than just mirroring). In the worst case this is **1** and **RN** reduces to **N** possible unique agent responses. In the best case, this becomes $\mathbf{\Sigma R_iN}$ where $\mathbf{R_i}$ represents the total number of unique responses to one input gesture from the $\mathbf{i^{th}}$ response strategy. Each of the response strategies is analyzed below.

For the "no response" response strategy, there is only ever one agent response. For the "repeat input gesture" response strategy, $\mathbf{R_iN}$ is **N** since each input gesture returns the same gesture as the agent output. For the "transform input gesture" response strategy, $\mathbf{R_iN}$ becomes $\mathbf{2^{(V+F)}N}$, where **V** and **F** are the number of Viewpoints dimensions and functional transformations that the agent can use to transform the input gesture into an agent response. In this case $\mathbf{2^{(V+F)}}$ represents the cardinality of the power set containing **(V+F)** elements. For the "emotion algebra" response strategy, the number of emotionally appropriate gestures available to respond with is dependent on the past history of gestures learned by the agent. For a history of **H** gestures with h appropriate gestures, that amounts to **hN** possible responses. In the worst case, this reduces to repeating the input gesture and $\mathbf{R_iN}$ reduces to **N**. This is justified by equating the emotional mirroring taking place with emotional contagion (Hatfield et al. 1994). In the best case, the entire history of gestures has the appropriate emotional content and $\mathbf{R_iN}$ becomes **HN**. For the "novel response from episodic memory" response strategy, the exact magnitude of $\mathbf{R_iN}$ is difficult to estimate for the best case since it is completely dependent on the past ordering of learned gestures and received input gestures. However, the lower bound for $\mathbf{R_iN}$ is **N** since in the worst case, if no gesture is found that is similar to the input gesture, the input gesture is repeated as the agent output. Finally, for the "learned interaction patterns" response strategy, given a set of **b** learned responses on average for the right hand side of each of N input gestures, the $\mathbf{R_iN}$ would be **bN**.

Therefore **RN** or $\mathbf{\Sigma R_iN}$ for all the **i** response strategies in the Viewpoints AI installation becomes $\mathbf{1 + N + 2^{(V+F)}N + N + N + bN}$ or $\mathbf{(1 / N + 3 + 2^{(V+F)} + b) \times N}$ in the worst case. This becomes $\mathbf{(1 + N + 2^{(V+F)}N + HN + \geq N + bN)}$ or $\mathbf{(1 / N + \geq 2 + 2^{(V+F)} + H + b) \times N}$ in the best case. Thus, the number of unique experiences possible with the full installation is much higher than in the first condition. The amount of authorial input is equally minimal in the full Viewpoints AI system. Therefore, since the authorial leverage for the first condition is very large, the authorial leverage for the full Viewpoints AI system is even larger. In addition, if complexity were a factor in our calculation of authorial leverage, it is visibly clear that the full installation has significantly higher complexity in its decision-making and in the user experience offered than the mirroring version of the system.

Pilot Empirical Study

The aggregated results for both the IRIS Measurement Toolkit and Creativity Support Index are presented in Figure 2. The system usability, flow, and enjoyment scales were used to evaluate the system in terms of its ability to produce high quality experiences for the user. The effectance and creativity support index scales were used to evaluate the ability of the system to co-create alongside the participant with equal creative agency. The results show that each of the experimental conditions did well, though no statistically significant results could be obtained between the different conditions potentially due to the small sample size (sample size of 10). However, regardless of the apparent lack of difference between the conditions, the survey ratings for the third condition show clearly that usage of the interaction-based authoring approach instantiated within the Viewpoints AI installation can indeed both create high quality subjective experiences for participants interacting with the installation as well as support collaboration with equal creative agency.

The semi-structured interviews were used to guide future development and contained questions regarding feedback about the experience, goals that users had while interacting with the system, what they liked / disliked about the installation, etc. The feedback was overwhelmingly positive, with particular emphasis on the aesthetics of the VAI's visual representation, freedom of creative expression felt by participants, amount of fun had by users, and sheer "cool factor" of the installation. Some of the negative feedback suggested that more was required to show that the agent was actually doing something other than mimicking the user. In addition, potentially indicating a miscommunication of the design goals for the installation, it became clear that some users felt like they should have been able to control the agent's actions to a greater degree. The goals of the users varied depending on how many times they had interacted with it and how inhibited they were. The goals generally went from exploring the boundaries of the system, to trying to get the agent to do certain reactions / responses, to trying to do novel interactions with the system that hadn't been tried before.

Discussion

The results given above help answer the three questions used to evaluate the interaction-based authoring approach instantiated within the Viewpoints AI installation. Using the interaction-based authoring approach led to a significantly higher authorial leverage (the ratio of variability of the experience, or more generally the quality of the system, to the amount of authorial input) than any pre-authored or pure mirroring version of the installation. The pilot study showed that the interaction-based authoring approach also led to high quality experiences, as judged by the system usability, flow, and enjoyment metrics administered. In addition, the study revealed that the interaction-based authoring approach was able to support collaboration with equal creative agency using the effectance and creativity support index metrics. However, it did not show a significant difference in ratings between the three experimental conditions for any of the survey metrics.

The lack of significant different between ratings for the different experimental conditions could be because of a number of reasons. Firstly, the study was conducted using a very small sample size. However, given that the ratings were so similar for all three, it is also possible that users had difficulty distinguishing between the different conditions in terms of the metrics used. Secondly, in terms of the evaluation, users were blind to the nature of the experimental condition as well as blind to the processes occurring within the virtual agent. According to Colton (2008), the process and the product are equally important to influence evaluation of creativity within the system. Therefore Turing-test style approaches to evaluation are found lacking. This seems particularly true when the creative domain is improvisation where participants evaluate the improvisational experience / process.

The results (especially from the semi-structured interviews) suggest that in order to improve the differential ratings of the full Viewpoints AI installation to the other conditions, the system's actions and outputs should be more noticeably different to highlight the system's original efforts better. This points to the requirement for a more full featured list of implemented transforms (both Viewpoints and functional transforms as well as gestural combination). In addition, video analysis showed that novice users had a hard time triggering the system's rhythmic repeated movement gesture segmentation mechanic. Thus current efforts focus on replacing the existing gesture segmentation algorithm with a more naturalistic automated gesture segmentation algorithm from Kahol et al. (2004). Finally, the experimental design is being refined to make the framing more explicit and will be scaled up.

Conclusion

In conclusion, this paper introduced a hybrid approach to knowledge authoring for co-creative systems called interaction-based authoring. The approach incorporates ideas from case-based learning and imitative learning, while emphasizing incorporation into open-ended co-creative application domains. This paper then presented an instantiation of the interaction-based authoring approach within the Viewpoints AI installation. The installation was then evaluated in terms of the extent to which it mitigated the knowledge-authoring bottleneck, produced high quality subjective experiences, and supported equal creative agency. Finally, the results of the evaluation were discussed in terms of guiding the future iterative development and evaluation of the installation.

References

Behrends, A., Müller, S., & Dziobek, I. 2012. Moving in and out of synchrony: A concept for a new intervention fostering empathy through interactional movement and dance. *Arts in Psychotherapy*, *39*(2), 107–116.

Boden, M. A. 2003. *The creative mind: myths and mechanisms*. Psychology Press.

Bogart, A., & Landau, T. 2005. *The viewpoints book: a practical guide to viewpoints and composition*. New York: Theatre Communications Group.

Brooke, J. 2013. SUS : A Retrospective. *Journal of Usability Studies*, *8*(2), 29–40.

Chen, S., Nelson, M. j, Sullivan, A., & Mateas, M. 2009. Evaluating the Authorial Leverage of Drama Management. In *Proceedings of the AAAI 2009 Spring Symposium on Interactive Narrative Technologies II* (pp. 20–23).

Cherry, E., & Latulipe, C. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support

Index. *ACM Transactions on Computer-Human Interaction*, *21*, 1–25.

Colton, S. 2008. Creativity Versus the Perception of Creativity in Computational Systems. In *Proceedings of AAAI Spring Symposium on Creative Systems* (pp. 14–20).

Colton, S., & Wiggins, G. A. 2012. Computational creativity: The final frontier? In L. De Raedt, C. Bessiere, & D. Dubois (Eds.), *20th European Conference on Artificial Intelligence:* (Vol. 242, pp. 21–26).

Csikszentmihalyi, M. 1997. *Flow and the Psychology of Discovery and Invention*. New York: HarperPerennial.

Csinger, A., Booth, K. S., & Poole, D. 1994. AI meets authoring: User models for untelligent multimedia. *Artificial Intelligence Review*, *8*, 447–468.

Davis, N., Popova, Y., Sysoev, I., Hsiao, C.-P., Zhang, D., & Magerko, B. 2014. Building Artistic Computer Colleagues with an Enactive Model of Creativity. In *Proceedings of the Fifth International Conference on Computational Creativity (ICCC 2014)*. Ljubljana, Slovenia.

Hall, M., National, H., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. 2009. The WEKA Data Mining Software : An Update. *SIGKDD Explorations*, *11*(1), 10–18.

Hastings, E. J., Guha, R. K., & Stanley, K. O. 2009. Automatic content generation in the galactic arms race video game. *IEEE Transactions on Computational Intelligence and AI in Games*, *1*(4), 245–263.

Hatfield, E., Cacioppo, J., & Rapson, R. 1994. *Emotional contagion*. *Current Directions in Psychological Science* (Vol. 2, pp. 96–99). Cambridge University Press.

Hecker, C., Raabe, B., Enslow, R. W., DeWeese, J., Maynard, J., & van Prooijen, K. 2008. Real-time motion retargeting to highly varied user-created morphologies. *ACM Transactions on Graphics (TOG)*, *27*(3).

Hoffman, G., & Weinberg, G. 2010. Gesture-based human-robot jazz improvisation. In *Proceedings - IEEE International Conference on Robotics and Automation* (pp. 582–587).

Hsu, W. 2008. Two approaches for interaction management in timbre-aware improvisation systems. In *Proceedings of the International Computer Music Conference (ICMC)*. Belfast

Jackson, S. a, Martin, A. J., & Eklund, R. C. 2008. Long and short measures of flow: the construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport & Exercise Psychology*, *30*, 561–587.

Jacob, M., Coisne, G., Gupta, A., Sysoev, I., Verma, G. G., & Magerko, B. 2013b. Viewpoints AI. In *Proceedings of the Ninth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*. Boston, MA.

Jacob, M., Zook, A., & Magerko, B. 2013a. Viewpoints AI: Procedurally Representing and Reasoning about Gestures. In *Proceedings of the Digital Games Research Association (DiGRA)*. Atlanta, GA.

Kahol, K., Tripathi, P., & Panchanathan, S. 2004. Automated gesture segmentation from dance sequences. In *Proceedings - Sixth IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 883–888).

Klimmt, C., Hartmann, T., & Frey, A. 2007. Effectance and control as determinants of video game enjoyment. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *10*(6), 845–847.

Laird, J. 2012. *The Soar cognitive architecture*. MIT Press.

Latulipe, C., Charlotte, U. N. C., Huskey, S., Beasley, R., & Nifong, N. 2011. SoundPainter. In *Proceedings of the 8th ACM conference on Creativity and cognition* (pp. 439–440).

MacDonald, L., Brosz, J., Nacenta, M. a., & Carpendale, S. 2015. Designing the Unexpected: Endlessly Fascinating Interaction for Interactive Installations. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14* (pp. 41–48).

Magerko, B., Permar, J., Jacob, M., Comerford, M., & Smith, J. 2014. An Overview of Computational Co-creative Pretend Play with a Human. In *Proceedings of First Workshop on Playful Virtual Characters at the Fourteenth Annual Conference on Intelligent Virtual Agents*. Boston, MA.

O'Neill, B., Piplica, A., Fuller, D., & Magerko, B. 2011. A knowledge-based framework for the collaborative improvisation of scene introductions. In *Proceedings of the 4th International Conference on Interactive Digital Storytelling* (pp. 85–96). Vancouver, Canada.

Reidsma, D., van Welbergen, H., Poppe, R., Bos, P., & Nijholt, A. 2006. Towards Bi-directional Dancing Interaction. In *Entertainment Computing - ICEC 2006* (pp. 1–12).

Thom, B. 2000. BoB: an interactive improvisational music companion. In *Proceedings of the fourth international conference on Autonomous agents* (pp. 309-316). ACM.

Tomasello, M. 2000. *The Cultural Origins of Human Cognition*. Harvard University Press.

Vermeulen, I. E., Roth, C., Vorderer, P., & Klimmt, C. 2010. Measuring user responses to interactive stories: Towards a standardized assessment tool. In *Interactive Storytelling* (pp. 38–43).

Vorderer, P., Klimmt, C., & Ritterfeld, U. 2004. Enjoyment: At the Heart of Media Entertainment. *Communication Theory*, *14*, 388–408.