MASSE: A System for Music Action Selection Through State Evaluation

Prashanth Thattai Ravikumar
Department of Communications and New Media
National University of Singapore
prashanth.thattai@u.nus.edu

Lonce Wyse
Department of Communications and New Media
National University of Singapore
cnmwll@nus.edu.sg

Abstract

This paper concerns the development of a music co-creation system that engenders creative partnerships in musical interaction. Towards this goal, specific design guidelines are identified for improving the musicians' sense of creative partnership with the system through behaviors that communicate musical intent, alternate musical roles based on musical stability, and negotiate transitions in musical characteristics based on togetherness in interaction. The design guidelines motivate the development of an agent architecture for interactive improvisation. The architecture enables an agent to select music response behaviors based on the stability and togetherness sensed during interaction. An agent improvisation system is developed that uses the architecture to make decisions during improvised percussion duets. The agent negotiates transitions between musical roles and musical characteristics by maintaining different degrees of metrical coherence and similarity with the rhythms played by the musician. By negotiating transitions with a musician, the system presumably has an impact on the musicians sense of creative musical partnership.

Introduction

The central question that is of interest here in this work is - how do we design systems that are creative musical partners to human musicians during live interaction? Free improvising musicians identify three functions that are important to engender a sense of creative partnership (or *co-experience*), namely - achieving communication, alternating musical roles, and negotiating transitions (Ravikumar, McGee, and Wyse 2018). In this paper, we discuss the design of a music co-creation system that performs the above mentioned functions to engender a sense of co-experience during real-time interaction with an improvising musician.

In order to study the impact of system design on co-experiences, we review work on music generation systems that are designed to perform alongside human musicians. The subset of music generation systems that are of interest here utilize sound as the primary medium of interaction with the musician. In the rest of the paper, these are termed as music response systems. Other systems have been developed that use note-based representations (Pachet 2003; Assayag et al. 2006) for direct engagement, and extra-musical (François, Chew, and Thurmond 2007; Cicconet, Bretan, and Weinberg 2012), and notational devices (Rowe 2004; Nika and Chemillier 2012) for making decisions along with the musicians. In group free improvisation situations that are of interest in this work, musicians often free themselves from the restrictions of note based interactions and focus on interacting through the sounds that are produced by the performers. During their interaction, musicians also report that extra-musical communication such as eye contact and physical gesturing is often a distraction from their interaction through sound, and prefer to play along by listening to the sounds of other musicians (Ravikumar, McGee, and Wyse 2018). The rest of the paper is organized around the issues and challenges in engendering a sense of co-experience with music response systems.

In the following section, specific examples of music response systems are presented to highlight the impact of designing specific system components on the musicians' sense of co-experience with the system.

Music response systems

Related work has focused on developing music response systems that interact using musical actions, alternate between lead and follower roles, and make decisions to negotiate transitions in the music. Reports are gathered from musicians' interactions with various systems and the impact of each component on the musicians' sense of co-experience is analyzed. Following the analysis, design guidelines are proposed to improve each component of the system.

Interaction through musical actions

Music response systems interact through musical actions that respond, in varying degrees, to the musical characteristics identified in the musicians' input. We review two music response systems that use different action descriptions. In the first system, musical actions are designed to produce musical material that is related to the sound made by the musician (Brown, Gifford, and Voltz 2016). In the second work, system uses musical actions that are intended to communicate a symbolic response (e.g., agreement, disagreement) to the musician along with the content (Murray-Rust and Smaill 2011). In this work, sonic interaction with a music response system is improved through musical actions that communicate the systems' intent to the musician.

Brown and his colleagues developed a music improvisation system that triggers musical actions to interact with the musician (Brown, Gifford, and Voltz 2016). The system contains a collection of six interaction behaviors - *repeat, imitate, shadow, initiate, silence, and turn taking* that it uses to respond to a musician. During the performance, the system triggers the behaviors at random. In a study that involved duet interactions with the system, the authors analyzed the impact of the system on musicians' co-experiences. Musicians who played with the system reported that they felt most engaged with the system during the duet interaction. However, they also felt a lack of creative initiative on the part of the computational co-creator.

The second system uses symbolic musical actions that are associated with communicative meaning (e.g., agree, disagree) to interact with the musician (Murray-Rust and Smaill 2011). During the performance, the system selects a musical action using a two step procedure. First, it listens to interpret the relation between the musical material played by the system and the musician and represents them as symbolic actions. Then, it consults a corpus to retrieve the continuation of the current action. Musicians interacted with two different system configurations, one that constructed musical actions through retrieval mechanism, and one that always imitates the musician. While musicians reported that their sense of interactivity was marginally higher with the system configuration that constructed musical actions through retrieval from the corpus, there were no significant differences to interactivity across the different conditions.

In summary, musical actions that are used in these music response systems are effective in engendering responsive interaction (e.g., initiating music), but they are restricted in their ability to communicate musical intent. Canonne and Garnier (2015) observe that musicians use strategies that combine their actions with the actions of other musicians. These strategies involve stabilizing their playing, waiting to see whether the other musician is going along, playing along, or densifying the musical texture. Musicians use these strategies to clearly communicate their intent.

> *In order to improve the ability of the system to impact co-experiences, we propose the design of music response behaviors that combine with the actions of the musicians to communicate musical intent. This is achieved by designing behaviors that stabilize through repetition, wait to let the musician play, play along with the current musical characteristics, and densify the musical texture of the improvisation.*

Alternating roles

Some music response systems maintain or change their response behavior in unpredictable ways in order to introduce changes in musical roles. In particular, musicians are often unaware of the exact time at which the system introduces changes in its behavior. In order to respond to behaviors that are unpredictable, musicians change their degree of attention and responsiveness to the system's actions, and consequently their role with respect to the system. Here, we look at three systems that use different processes to achieve this effect. The first system changes its behavior by changing the musical source that it listens to and uses for response generation (Lewis 2000). Another system changes behavior by altering the variability in its responses with respect to the musicians' input and inbuilt internal goals (Young, Bown, and others 2010). The third system alters its behavior through changes in its internal state (Donnarumma 2017). While the above mentioned systems bring about a change in the musical roles, the roles change because the musician adapts to the changes introduced by the system.

The Voyager system is notable for being recognized by several authors and musicians as an equal co-improviser (Lewis 2000). During the performance, the system changes behavior by changing the musical sources that it listens to and the number of musical channels that it uses for output. These changes are triggered through chance processes that musicians will not be able to predict beforehand. Following the change, the system resumes its usual responsive behavior that involves adapting to the timbre and textural elements to the input from the changed source. The responsive behavior of the system engenders a feeling that the system has a personality of its own. However, the unpredictability in the system's behavior often makes it difficult for the musician to follow the changes.

The second system (Young, Bown, and others 2010) introduces changes based on the musicians' input and an internal goal state. Young, Bown, and others designed a system that improvises by clapping rhythmic patterns along with a musician. The system responds by generating variations on the last rhythm played by the musician, and selecting the variations that are most similar to a target rhythm. In certain configurations when musicians played with low variability, they were able to interpret the changes that were introduced by the system. But the system's behavior was found to be erratic and difficult to play along with when musicians' varied their rhythm patterns. As a result, musicians found it difficult to engage with the unpredictable behaviors exhibited by the system.

The third variety of systems introduce variability in their behavior through autonomous state changes. One such system designed by Donnarumma (2017) gets input from the performers' body gestures and generates musical responses based on its dynamically evolving state. The performer restricts the scope of actions in order to interpret the changes made by the system. With these restrictions in place, the performer is able to interpret the behaviors initiated by the system, such as a crescendo, as a part of musical movement. While imposing restrictions appears to be effective in working along with the unpredictable nature of the system, performing with these constraints requires considerable training on the part of the human performer.

In summary, these three musical systems behave unpredictably in order to bring about changes in musical roles, but the changes in the roles are a result of the musician adapting to the system's unpredictability. In human improvised performances, co-improvising musicians make decisions to alternate roles based on the degree of relative stability that they maintain with respect to the other musicians during the interaction (Canonne and Garnier 2012). When musicians sense that the group has been playing with too few changes, they introduce instability by diversifying the musical material. When musicians notice that there are too many changes, they introduce stability in the music (e.g., through repetition).

> *Improvements to the design of systems such as those described above require the development of modules that introduce changes in roles by sensing the relative stability of the systems' behavior with respect to behavior of the musician in the interaction.*

Negotiating transitions

Music response systems negotiate transitions along with the musician by changing their musical configuration over a period of time. One such system changes its musical configuration through incremental adjustments to the music generation parameters (Albert 2013). Another system responds to the differences sensed in low level parameters with rapid changes in the musical variables (Bown 2018). As an improvement in design, the system makes decisions to change its musical characteristics by sensing its state of togetherness with the musician. The specific design guideline for developing this aspect of a music response system involves the development of a listening component that monitors the degree of togetherness in its actions with respect to the musician.

Let us consider the first system that introduces changes to its musical configuration independently of the musician (Albert 2013). The system contains a collection of music generation parameters it varies during the interaction. In one of the experimental configuration of the system, the systems introduces gradual and incremental changes to the music generation parameters in order to change the musical characteristics. Musicians who played with the system reported that the changes introduced by the system were responsive to their playing. However, musicians also felt a "latency between the human's action and system's adoption of that action" (when IMP chooses to adopt that action) (Albert 2013).

Bown's (2018) system, on the other hand, is designed to be very responsive to the slightest changes introduced by the musicians. The systems' immediate responsiveness to the low level parameters and their non-linear coupling with higher level musical behaviors introduces rapid changes in the music. While it is clear that the system adopts to the changes introduced by other performers, musicians face a different challenge with the system. During the interaction, the system introduces rapid changes that are difficult to follow. In these moments, the musicians feel the need for the system to slow down, and to "lock in to the moment".

Existing musical systems perform transitions with an improvising musician, but are unable to negotiate them in a manner that feels interactive. Interaction during transitions is an important characteristic of collective free improvisation as there is a high degree of correlation between the decisions made by the musicians about the occurrence of the transition points (Canonne and Garnier 2015). Wilson and MacDonald (2016) found that during the moments of transitions, human musicians make decisions to maintain or change their actions by evaluating their togetherness with co-improvising musicians (Wilson and MacDonald 2016). In particular, human musicians seem to evaluate the degree of togetherness through the homogeneity of the combined musical output. When musicians sense that they have been playing musical material that is homogeneous for a long time, they perform strategies that diversify the musical texture. Similarly, when musicians sense that they are playing material that is very complementary to each other, they adopt strategies that bring the musical material closer to each other.

> *Improvements to the design of systems such as those described above require the development of modules that negotiate transitions based on sensing the togetherness in the musical output that is jointly produced by all the musicians.*

In summary, there are two aspects of related work that offer scope for improvement based on the strategies used by free improvising musicians. The ability of the system to respond to musician's actions is improved through the design of musical actions that combine with the actions of other musicians to communicate musical intent. The decision making strategies that are used by the system are improved by evaluating the relative stability and togetherness of the musical output to make decisions. Next, the different design recommendations that have emerged from the literature described above are integrated within a new agent architecture.

Agent architecture

The agent's architecture is divided into two parts: 1) a reactive part, and 2) a decision making part. The reactive part contains music response behaviors that agent uses to generate music from input. The decision making part stores the sequences that it listens to, and evaluates based on stability and togetherness. During interaction, the agent generates music by choosing musical actions that combine with the musical sequences played by the musician. The agent also makes decisions to maintain or change its selected actions by evaluating the degree of togetherness, and stability during its interaction with the musician.

Musical action

The agent uses musical actions to generate sequences that are responsive to the input from the musician. Musical actions are transformation functions that combine two or more musical sequences to produce a third sequence that is partially similar and partially different from the input sequences. Each musical action that is used in the agent requires two musical sequences as input - a sequence that is internally generated by the system, and a sequence that is obtained from a musician or from the environment. New sequences are generated by combining the musical characteristics of the internal sequence with the characteristics of input.

Action descriptions are central to the reactive part of the system labeled as region (*a*) in Figure 1. The reactive part of the system is responsible for the generating musical material that is responsive to the real-time input from the musician.

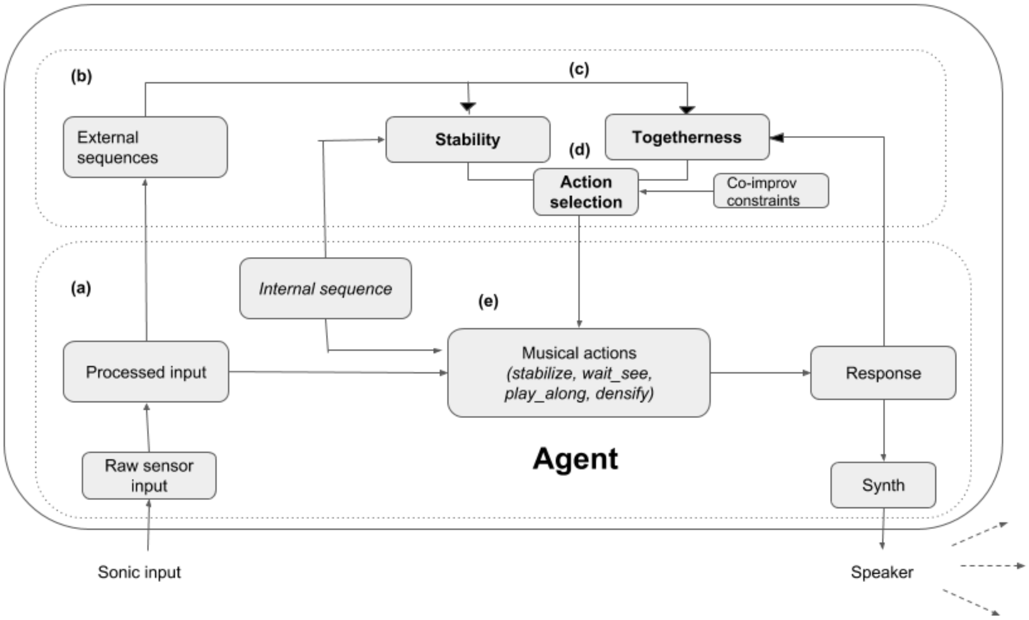

Figure 1: Parts of the agent architecture: *a) Reactive part, b) State and decision making, c) Scoring states, d) Selecting new action, e) Generating musical material*

The reactive part contains modules for music generation and corresponds to the region (*e*) of Figure 1. Once an action is selected, the two sequences - an internal sequence and an input sequence are combined to generate new musical material.

Strategy	Description
Stabilize	Repeat source pattern
Wait and see	Play only when musician plays
Play along	Imitate the characteristics of musician
Densification	Introduce characteristics different from self and past

Table 1: Music response behaviors

During the interaction, the agent uses the four actions (listed in Table 1) to generate musical material that is responsive to the input from the musician. The agent uses the generated musical material to stabilize the music, wait to see where the musician is taking the music, play along with the musician, or densify the playing with respect to the musician. The descriptions of the four actions are presented in Table 1.

The musical actions are implemented through operations that combine musical characteristics from the internal sequence and the input sequence. As an example, one possible set of operations that combine the intensity levels of the internal sequence and the input sequence is described here. In order to produce new sequences that maintain stability in the music, the agent assigns an intensity level that is equal to the mean intensity level of the internal sequence and corresponding input from the musician. To implement the wait and see strategy, the system selects events from the internal sequence only when the intensity of the input exceeds a threshold. In all the other conditions, the agent remains silent or selects an intensity value of 0. To play along with the musician, the agent decreases the intensity of its output sequence when it senses that the intensity of input is high and increases the intensity when the input is low. In order to densify the musical texture, the system increases the intensity of its output when it senses that the input is high, and maintains its intensity when the input is low.

The agent performs the above mentioned operations at the rate that is equal to the rate at which it receives inputs (every 20 ms). Every input that is received from the musician is interpreted, and combined with an internal event to produce a new musical event. The ability to react to musical changes at a short time scales keeps the agent responsive to sudden changes in the input during interaction.

The remaining parts of the agent architecture (region (*b*) of Figure 1) use state to make decisions and select actions and are described here.

State

In order to make decisions, the agent evaluates the improvisation by keeping track of two metrics of its interaction with the musician, namely, togetherness and stability. This information is represented as numerical values that correspond to the degree of togetherness and the degree of stability. The numerical values are computed using the features of the musical sequences that the agent has listened to and played in the past.

The agent represents the information about togetherness and stability through numerical values between 0 and 1. A high degree (1) of togetherness corresponds to a situation in which musicians play musical material that is very similar to other musicians, and a high degree of stability (1) corresponds to a situation in which the agent plays musical material that maintains musical coherence with the musician. The values assigned to togetherness and stability are stored in state variables that the agent uses for decision making.

The procedure for assigning a numerical score to togetherness and stability is explained through notions of internal and external component of the state (region (*c*) in Figure 1). The internal component of the state keeps track of the musical material that is generated by the agent (e.g., internal sequences, responses). The external state stores the musical material that the agent receives as input. It is important to note that unlike the processes of response generation, state is computed at a timescale on the order of 5 seconds. During the period of listening, the agent pairs the input from the musician and the internal sequence to form musical sequences. Then, it computes the differences in the musical sequences that are stored in the internal and the external state to compute the degree of togetherness and stability. The different steps that the agent follows to assign numerical scores to the state are described in detail below.

Scoring togetherness The agent scores the degree of togetherness by scoring the differences between the responses generated by the agent and the external sequences stored in the external state. The following procedure is used to score the differences that the agent senses between the musical material.

- The agent continuously obtains new sonic input from the environment in terms of musical features (e.g., intensity values, spectral changes). The musical features are directly obtained from the sensor and stored as musical sequences. The agent generates an internal sequence that it combines with the input sequence.
- The musical features are concatenated to form musical sequences that are stored in memory. These sequences are stored in the external state variable.
- While the agent is listening to input, it is also generating new musical material using the process described in region (*e*). The agent stores the generated responses in an internal state variable.
- In order to compute togetherness, the agent compares the values of the musical sequences in the external state with the responses stored in the internal state and assigns a numeric value based on the magnitude of the differences.
- The numeric value of differences is converted to ratio that corresponds to the number of differences per unit of time (e.g., beat)
- Finally, the numeric value of differences is subtracted from 1 to obtain the degree of togetherness.

$$degree_{togetherness} = 1 - ratio_{difference} \qquad (1)$$

The highest value of degree of togetherness is 1 and corresponds to a situation when musicians are playing musical material that are completely homogeneous with respect to their changes in intensity and timbre values. This signifies the state of togetherness in playing. The lowest value of togetherness is 0, and corresponds to a situation when musicians are playing musical material that is complementary or contrasting with respect to the changes in intensity and timbre values. This signifies a state of indifference between the agent and the human. Values that lie between 0 and 1 correspond to different degrees of togetherness.

Scoring stability The agent assigns a numeric value to the stability in the interaction based on the correlation between the musical sequences expected by the agent and the actual musical sequences played by the musician. The first 3 steps in the procedure for scoring stability are the same as the steps for scoring togetherness.

- Step (1), (2), and (3) from scoring togetherness.
- The agent computes the correlations between the musical features that are stored in the internal and external state.
- The correlation values are converted to a range between 0 and 1 to correspond to the degree of stability.

  The highest value of degree of stability is 1 and corresponds to a situation when the agent finds a strong correlation between the events that it expected to occur and the events that the musician played in the performance. The lowest value of stability is 0, and corresponds to a situation when the agent finds a weak correlation between the events that it expected to occur and the events that the musician played in the performance. Values that lie between 0 and 1 correspond to different degrees of stability in the interaction.

Scoring co-improvisation state The values that were assigned to the degree of togetherness and the degree of stability are added to assign a numerical value to the state of co-improvisation.

$$coimprov_{score} = degree_{togetherness} + degree_{stability} \qquad (2)$$

Next, the musical actions and the state values are integrated via a mechanism that selects musical actions based on state information.

Action selection

Using the co-improvisation score, the agent makes decisions to maintain or change its actions during live interaction (region (*d*) of Figure 1). First, the agent assigns a numerical score to the co-improvisation state by computing the degree of togetherness and degree of stability. Then, it compares the co-improvisation score against a target score to check whether the current state satisfies the internally specified constraints of co-improvisation. When the agent determines that the state score does not meet constraints, it selects a new state to move to, and selects one of the available musical actions to reflect its decision to change. Given that the agent does not know the sequences that the musician will play, it makes decisions to maintain or change actions based on the conditional assumption that musicians will repeat their action.

Decision boundaries Decisions boundaries are constructed in order to enable the agent to make decisions to maintain or change its state. The decisions boundaries divide the state space into ranges that correspond to different regions of togetherness and stability with another musician. Periods of high stability and high togetherness are characterized by values that are between 0.75 and 1. Values between 0 to 0.25 correspond to regions of playing in which musicians play with low stability (high variability) and low togetherness. Values between 0.25 and 0.75 correspond to the periods of playing in which musicians balance stability with togetherness in order to play along with other musicians. Co-improvisation values that lie in the range of 0.25 to 0.75 are ideal for the agent to play along with the musician.

Decision to maintain/change state The numerical co-improvisation score that is computed in the previous steps is compared with the decision boundary values to make decisions to maintain or change the state. The agent maintains its current state when it senses that the state score is within the acceptable bounds. The agent changes its state when the state score lies outside the acceptable bounds. The agent implements the change by selecting a new action from the available list of actions.

Selecting new action In order select a new action, the agent scores the stability values for each of the alternate actions. The stability scores are computed using the processes described in previous steps, and combined with the degree of togetherness score to generate a new co-improvisation score. The new co-improvisation score is checked with decision boundary values to determine whether the alternate actions that are available to the agent are viable. The agent selects one of the alternate actions that is produces a score closest to the desired co-improvisation state. In the event that the agent finds that no actions bring the co-improvisation score within bounds, the agent selects one of the alternate actions at random. The selected action is combined with the musical input to produce musical material.

Rhythmic improvisation agent

The architecture is used to develop a rhythmic improvisation system that selects musical actions to perform rhythmic duets. The various components of the agent architecture that are developed for interactive rhythmic improvisation are described here.

Numeric scoring functions The agent improvisation system that is developed in this work computes numerical measures of stability and togetherness using the notions from Gifford and Brown (2006) as well as Cao, Lotstein, and Johnson-Laird (2014).

Gifford and Brown implemented an interactive rhythm improvisation system that generated rhythmic accompaniment that varies in its degree of coherence with respect to a rhythmic input (Gifford and Brown 2006). Rhythmic coherence was numerically measured through the correlations between the meter, velocity, pitch, and the duration of events in two rhythm patterns (Gifford and Brown 2006). The lower the rhythmic coherence between the rhythms, the less stability that musician feels while playing with the agent. The higher the rhythmic coherence between the rhythms, the more stability that musician feels while playing along with the agent.

Cao, Lotstein, and Johnson-Laird (2006) propose a numeric metric of rhythmic similarity based on the similarities in metrical organization of rhythm patterns. The agent that is developed in this work assigns a numeric value to togetherness based on the similarities in three metrical units - syncopated notes, notes on the musical beat, and other events (e.g., rests) - that it finds with the musician.

Rhythm representation In order to assign numerical values to stability and togetherness using the notions from Gifford and Brown as well as Cao, Lotstein, and Johnson-Laird, the agent uses a three part representation of rhythm patterns. Each rhythm pattern is specified through three parameters: 1) binary sequences of hits and rests for 16 hits, 2) intensities that correspond to each hit in the binary sequence, 3) the timbre that is played at each hit. Each hit in the source rhythm has a duration of 1/8th beats at the tempo of 60 beats per minute.

Musical actions Using the rhythm representation mentioned above, the improvisation agent chooses actions that alter the intensity and the metrical division of a musical unit (e.g., a beat). Alterations to intensity and metrical subdivisions changes the density and syncopation of the rhythm, and subsequently affects the metrical stability and rhythmic similarity values. Each musical action generates an internal sequence with metrical subdivisions for each beat (e.g., single hits, eight notes). In order to *wait and see* where the musician is going, the agent generates internal sequences with a musical meter that has a lower metrical subdivision compared to the previous meters played by the musician and the agent. To *play along* with the musician, the agent generates internal sequences that have the same number of metrical subdivision as the musician. The *stabilization* action repeats sequences with the same metrical division that the agent is currently playing with. In order to *densify*, the agent generates internal sequences with musical meters that have a higher number of metrical subdivisions compared to the previous meters played by the musician and the agent.

Computing numerical values After each action, the agent computes the numerical scores of similarity and stability to compute the co-improvisation score. Each 1/8th division of the beat is scored based on whether the agent and the musicians registered a hit or left a rest at that beat division. If musicians performed a hit in the place of a rest left by the agent on a 1/8th beat, or vice versa, it counted as a single unit of difference. Togetherness is the ratio of number of differences in 16 1/8th notes. The degree of togetherness is computed as *1-ratio of togetherness*. In order to measure metrical stability, the agent finds the correlations between velocity and duration of hits played by the musician and the agent. Timbre values are not used in stability evaluation. Differences in timbres are experimentally controlled and do not contribute to metrical stability. Based on the numerical scores that are computed after each action, the agent computes the score for the co-improvisation state. The sum of togetherness and stability is normalized to 1 and corresponds to the co-improvisation score.

Selecting actions The agent compares its co-improvisation score with the boundary values to maintain or change its musical action. When the agent determines that the co-improvisation score is higher or lower than the range, it changes its behavior. When the agent determines that the co-improvisation score is within the range, it maintains its behavior. The boundary values were experimentally set to 0.25 and 0.75 as they produced to the periods of interaction in which agent balanced stability with togetherness.

Performing with the system An improvisation situation in which the agent negotiates transitions along with the musician is described here. Through an unpredictable change, the agent negotiates a change in musical texture and brings the co-improvisation value within bounds. At the beginning of the interaction, the agent selects a musical action and repeats it. Once the musician begins to play, the agent senses an increase in the degree of togetherness and stability in the performance. The agent maintains its action until the value of the co-improvisation state is within the range specified by the decision boundaries. When the agent determines that the value assigned to the state is outside the range, it chooses a new musical action that either introduces a greater (*densify*) or lesser (*wait and see*) metrical subdivision of the beat with respect to the musician. By maintaining and changing the metrical subdivisions in the beat, the agent introduces changes in the musical texture while keeping the co-improvisation score within bounds. Videos of performances with the agent improvisation system are available at `https://prashanthiccc19.wordpress.com`.

Discussion

In this paper, state-of-the-art music response systems were evaluated to derive two design guidelines that were used to develop an agent architecture for music co-creation. A rhythmic improvisation agent was developed based on the architecture that plays along with musicians. During the interaction, the agent negotiates transitions between regions that vary in the degree of metrical coherence and similarity to the rhythms played by the musician.

An agent was developed from the design guidelines that generates music by combining its musical actions with the actions of the musician. While earlier systems select specific actions that communicate an intent to the musician (Murray-Rust and Smaill 2011), the agent that is developed in this work combines its actions with the actions of the musician to communicate its intent. It remains to be studied whether an agent that communicates its intent by combining its musical actions with the musician's actions improves the musicians' co-experiences with the system.

During moments of change, the agent makes decisions to maintain or change its response behavior by monitoring the changes in the co-improvisation state. With earlier systems, musicians have been unable to negotiate these moments in a manner that feels interactive. The particular procedure for decison making using co-improvisation state and the decision boundaries allows us to adjust the latency in the systems' responses to a range that feels interactive.

In order to compute the co-improvisation state, the agent monitors the togetherness and stability of its actions with respect to the musician. The operationalization of co-improvisation states through togetherness and stability allows humans to explain the decisions made by the system at an intuitive level of abstraction. A review by Karimi (2018) identified the need for improving the explainabilty of the decisions made by the co-creative systems that are otherwise non-interpretable. The ability to make decisions that are interpretable is also a first step towards improving the explainability of co-creative systems.

Finally, the rhythmic improvisation system that was implemented using the architecture measured stability through metrical coherence, and togetherness using rhythmic similarity. During the performance, the system negotiates transitions between different regions that vary in the degree of metrical coherence and similarity to the rhythms played by the musician. During the preliminary trials with the system, it was observed that the systems' behavior did not consistently correspond to the musicians' sense of changes in metrical coherence and similarity. A possible explanation is that the particular method of additively combining the scores of metrical coherence and similarity may not be an accurate measure of human judgement. Improvements to the system design will use metrics that have a closer perceptual correspondence with musicians' notions of togetherness and stability.

Conclusion

This paper concerns the design of a music co-creation system that engenders a sense of creative partnership. Towards this goal, specific design guidelines are identified to improve individual system components. The various components are integrated within an architecture to develop an agent that interacts by communicating musical intent, alters behaviors through monitoring the stability in interaction, and negotiates transitions by sensing togetherness in musical characteristics. In rhythmic improvisation situations, the agent improvisation system negotiates transitions in roles and musical characteristics through trade-offs between stability and togetherness. Future work will study the ability of the agent improvisation system to impact co-experiences.

Acknowledgments

I would like to thank Dirk Stromberg and Isaiah Koh for their insightful discussions.

References

Albert, J. 2013. Interactive musical partner: A demonstration of musical personality settings for influencing the behavior of an interactive musical generation system. In *In Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Assayag, G.; Bloch, G.; Chemillier, M.; Cont, A.; and Dubnov, S. 2006. Omax brothers: a dynamic yopology of agents for improvization learning. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, 125–132. ACM.

Bown, O. 2018. Performer interaction and expectation with live algorithms: Experiences with zamyatin. *Digital Creativity* 29(1):37–50.

Brown, A. R.; Gifford, T.; and Voltz, B. 2016. Stimulating creative partnerships in human-agent musical interaction. *Computers in Entertainment (CIE)* 14(2):5.

Canonne, C., and Garnier, N. B. 2012. Cognition and segmentation in collective free improvisation: An exploratory study. In *Proceedings of the 12th International Conference on Music Perception and Cognition and 8th Triennal Conference of the European Society for the Cognitive Sciences of Music*, 197–204.

Canonne, C., and Garnier, N. 2015. Individual decisions and perceived form in collective free improvisation. *Journal of New Music Research* 44(2):145–167.

Cao, E.; Lotstein, M.; and Johnson-Laird, P. N. 2014. Similarity and families of musical rhythms. *Music Perception: An Interdisciplinary Journal* 31(5):444–469.

Cicconet, M.; Bretan, M.; and Weinberg, G. 2012. Visual cues-based anticipation for percussionist-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI'12)*, 117–118. Boston, Massachusetts, USA: ACM.

Donnarumma, M. 2017. Beyond the cyborg: Performance, attunement and autonomous computation. *International Journal of Performance Arts and Digital Media* 1–15.

François, A. R.; Chew, E.; and Thurmond, D. 2007. Mimi-a musical improvisation system that provides visual feedback to the performer. Technical Report 07-889, University of Southern California Computer Science Department.

Gifford, T. M., and Brown, A. R. 2006. The ambidrum: Automated rhythmic improvisation. In *Proceedings of the Australasian Computer Music Conference*, 44–49. Adelaide, Australia: Australasian Computer Music Association.

Karimi, P.; Grace, K.; Maher, M. L.; and Davis, N. 2018. Evaluating creativity in computational co-creative systems. *arXiv preprint arXiv:1807.09886*.

Lewis, G. E. 2000. Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal* 10(nil):33–39.

Murray-Rust, D., and Smaill, A. 2011. Towards a model of musical interaction and communication. *Artificial Intelligence* 175(9):1697–1721.

Nika, J., and Chemillier, M. 2012. Improtek: integrating harmonic controls into improvisation in the filiation of omax. In *Proceedings of the International Computer Music Conference (ICMC'12)*, 180–187.

Pachet, F. 2003. The continuator: Musical interaction with style. *Journal of New Music Research* 32(3):333–341.

Ravikumar, P. T.; McGee, K.; and Wyse, L. 2018. Back to the experiences: empirically grounding the development of musical co-creative partners in co-experiences. In *6th International Workshop on Musical Metacreation. 9th International Conference on Computational Creativity, ICCC*, 1–7.

Rowe, R. 2004. *Machine musicianship*. MIT press.

Wilson, G. B., and MacDonald, R. A. 2016. Musical choices during group free improvisation: a qualitative psychological investigation. *Psychology of Music* 44(5):1029–1043.

Young, M. W.; Bown, O.; et al. 2010. Clap-along: A negotiation strategy for creative musical interaction with computational systems. In *Proceedings of the International Conference on Computational Creativity 2010*, 215–222. Libson, Portugal: Departament of Informatics Engineering University of Coimbra.