Imitation Game: Real-time Decision-making in an Interactive Composition for Human and Robotic Percussionist

Artemi – Maria Gioti
Institute of Electronic Music and Acoustics (IEM)
University of Music and Performing Arts Graz, Austria
gioti@iem.at

ABSTRACT

This paper describes an interactive composition for human and robotic percussionist exploring decision-making processes in the context of composed interaction scenarios. The composition is based on a dynamic form, shaped by decisions made by the musician and the robotic percussionist in real-time. Using a Neural Network trained to recognize different instruments and playing techniques, the robotic percussionist makes long-term decisions based on metrics of musical contrast. Similarly, the musician interprets a non-linear score, which enables him/her to interact with the robotic percussionist in real-time. The paper describes various components of the system, including the auditory processing and decision-making stage, and introduces a framework for artistic experimentation borrowing evaluation methods from human-computer improvisation.

1. INTRODUCTION

Musical robotics is a fast expanding research field, covering a wide range of musical instruments, from percussion to string and wind instruments [1], as well as interaction paradigms: from interactive human-computer music systems to laptop orchestras [2].

This growing interest for acoustic – or, more accurately, mechanically produced – sound can be attributed to both sound- and interaction-driven design and compositional choices. The complexity of acoustic sound [3], the expressive potential of physical actions [3] and the role of visual communication in anticipating and coordinating performers' actions, as well as establishing cause-effect (i.e. action-sound) relationships [4] are some of the most commonly cited advantages of musical robotics over electronically (re)produced sound.

Research in musical robotics has expanded to encompass a large variety of applications, from *robotic musical instruments*, played by human musicians or triggered by predetermined sequences, to *anthropomorphic musical robots*, designed to imitate (physical) human actions, and *perceptual robots* [3, 5]. The last category refers to autonomous musical robots, able to "perceive" and interact with their sonic environment, suggesting an overlap with the field of interactive music systems.

The work described in this paper falls under the latter category, incorporating both hardware components and software agency. *Imitation game* is an interactive composition for human and robotic percussionist based on a dynamic form, which is shaped by decisions made by both the musician and the robotic percussionist in real-time. The robotic percussionist interacts with the human based exclusively on machine listening, particularly a feed-forward Neural Network trained to recognize different instruments and playing techniques. Decisions are made both on a meso and macro time scale, based on metrics of rhythmic, timbral and dynamic contrast.[1]

2. DECISION-MAKING IN INTERACTIVE MUSIC SYSTEMS

Most interactive music systems – whether hardware of software-based – incorporate one or more interaction "modes" [3] or "modules" [6], which are responsible for different agent behaviors and, therefore, sonic affordances. In the case of interactive robotic percussionists these modes can differ with respect to rhythmic material, interaction timing (e.g. synchronous vs. asynchronous action) and/or the sensory processing and decision-making processes involved in them.

For example, *Haile* [3] is a perceptual robot equipped with six different interaction modes, some of which are synchronous and some sequential. These modes are not selected in real-time, but are activated in predetermined sequences in compositions written by its creators [3]. Real-time decision-making processes are employed mainly on the "phrase" level: the robotic percussionist calculates the stability of an input rhythm and then chooses from a database of rhythms based on similarity metrics and a target stability value [5]. Another perceptual robot, *Shimon* [6], is based on three "interaction modules", which are described as segments with a fixed or condition-dependent duration, while the *CIM software* [7] is based on a model of duet interaction centered around six different types of musical activity: "imitate", "initiate", "loop", "restate", "shadow" and "silence". In the current version of the software, activity selection is random, a shortcoming that the authors are planning to address in next iterations of the software [7].

[1] A video documentation of the piece is available in the following link: http://www.artemigioti.com/demos/Imitation_game.html

2.1 Generating Meaningful Responses

The integration of different interaction modes and complex decision-making processes in the above mentioned systems is indicative of an interaction design oriented towards a "conversational" [8] – i.e. *reciprocal* – model of interaction, rather than one based on cause-effect relationships. Emmerson [9] distinguishes between "causing" a reaction and "provoking" a response – particularly a "meaningful response" – using the example of two musicians improvising in a call-and-response fashion as a model for the second. However, as Emmerson [9] points out, "meaningful" is a musical judgement.

In interactive musical robotics, musical meaning is – not unjustifiably – linked to "higher-level percepts" and "subjective concepts" [4]. What Weinberg [4] refers to as "higher-level percepts" are musical meta-parameters (e.g. metrics of rhythmic stability, melodic similarity etc), which are used to describe the meso and macro time scale, rather than the sound event level. Meaning is, therefore, not only subjective but also context-dependent. Furthermore, these higher-level percepts are in most cases specific to the instrumentation, the musical idiom (e.g. jazz vs. free improvisation) and even the compositional idea.

2.2 Can the computer say "no"?

Another key distinction between a reciprocal interaction based on decision-making processes and a mere input-output mapping is that of *intention*, as well as that of *negotiation of different intentions* between actors. Or, as Emmerson [9] puts it: can the computer say "no, thanks"?

A behavior that is strictly reactive and not pro-active falls under causality, rather than interactivity. A meaningful response does not mean just following, but also leading, even ignoring or rejecting your co-player's actions – behaviors often incorporated in the decision-making stage of interactive music systems [4, 10].

3. IMITATION GAME

Notions of musical intention and meaning – particularly a meaning that is constructed through context (i.e. on a meso and macro time scale, rather than on a sound event level) – are some of the central concepts explored in *Imitation game*. This meaning is not universal, but composition-specific and constructed – *composed* – based on the composer's subjective criteria.

Auditory processing in *Imitation game* therefore extends beyond the sound event level (instrument and playing technique recognition), to the phrase level (calculating metrics of musical contrast) and form level (monitoring the evolution of contrast metrics as a function of time). Similarly, decision-making extends beyond the selection of single actions to the initiation of various "interaction scenarios" [11], in which the agent assumes different roles, (e.g. following and leading). The auditory processing, decision-making and action stage of the robotic percussionist in *Imitation game* are described in detail in the following sections.

3.1 Auditory Processing

The auditory processing stage of the robotic percussionist is based on a feed-forward Neural Network (NN) trained to recognize different instruments (cymbals, bongos and cowbells) and playing techniques (strokes, scraping and bowing). In order to train the NN, several examples of each class were recorded using a large number of different mallets and various microphones, to ensure variability in the data set and prevent overfitting. The recorded examples were analyzed using a window of 2048 samples and 50% hop size (sampling rate: 44100 Hz) and divided into three sets: a training set (60% of the data set), a cross-validation and a test set (each 20% of the data set). The final set of features used for machine learning was selected through an iterative process of training and testing and consists of the following features: onset, spectral centroid, spectral spread, spectral slope, spectral flatness, spectral roll-off and Mel Frequency Cepstral Coefficients (MFCCs).

In order to solve this classification problem several methods were tested, including breaking the task into smaller classification problems (e.g. using one NN for instrument recognition and another one for playing technique recognition). Both multiclass classification (assigning a single label to each sample) and multi-label classification (assigning several labels to a single sample, e.g. an "instrument" and a "playing technique" label) performed equally well on a balanced training set (i.e. a training set in which none of the classes are significantly over- or underrepresented). Eventually, multiclass classification was preferred over multi-label classification due to its practical advantages (faster training and lower computational cost in run-time).

In its final form, the NN consisted of one hidden layer (consisting of the same number of units as the input layer) and 11 output units corresponding to the following classes/labels: "bongo, stroke", "cymbal, stroke", "cowbell, stroke", "bongo, scraping", "cymbal, scraping", "cowbell, scraping", "cymbal, bowing", "cowbell, bowing", "cymbal, resonance", "cowbell, resonance" and "background noise". Background noise was added as a separate class in order to integrate noise gating in the classification task. The activation function used was the logistic sigmoid.

The accuracy of the NN on the test set reached 85%, with one of the main weaknesses of the algorithm being the low accuracy of the onset detection algorithm[2] on cymbal strokes, presumably due to the characteristic envelope shape of the instrument (slow attack). Finally, a confidence threshold was introduced to filter out some false predictions and improve the overall accuracy of the algorithm.

[2] "Onsets" SuperCollider UGen [12] using the rectified complex deviation onset detection function [13].

3.2 Decision-making

The decision-making stage of the robotic percussionist processes data collected in the auditory processing stage and chooses among three different interaction scenarios:

(1) *repeat* (play the exact same material),

(2) *imitate* (play similar material) and

(3) *initiate* (introduce new material).

The terms "imitate" and "initiate" were borrowed from Brown et al.'s previous work [7] and adapted to describe specific interaction scenarios used in the composition. Particularly, "imitate" is used to refer to the generation of similar material, using high-level percepts such as rhythmic contrast as similarity measures, rather than the reuse of material within a short time frame [7].

It's been suggested that, "musical changes" [14], particularly "non-arbitrary" ones [15], are key to designing meaningful musical interactions. That is presumably because the ability of an interactive music system to propose changes (e.g. introduce new sound material) is indicative of a high level of music understanding, as well as a high level of autonomy. In line with that view, interaction scenarios in *Imitation game* are not selected randomly by the robotic percussionist, but based on metrics of *rhythmic, timbral* and *dynamic* contrast, which are calculated as follows:

• **Rhythmic contrast:** standard deviation of (detected) Inter-Onset Intervals (IOIs).

• **Timbral contrast:** standard deviation of the (detected) timbre probability distribution (where timbre x is treated as a random variable that can take 8 possible values: "bongo, stroke", "cymbal, stroke", etc., excluding resonances and background noise).

• **Dynamic contrast:** standard deviation of the (detected) dynamics probability distribution (where "dynamic" x can take 3 possible values: p, mp/mf and f).

These contrast metrics are calculated on a phrase basis and their values are stored in arrays, allowing the robotic percussionist to make decisions based on their evolution in time. Specifically, if the estimated rhythmic contrast has been constant (i.e. around the same value), or monotonic (i.e. constantly increasing or constantly decreasing) for the last few phrases, the robotic percussionist is less likely to play similar material ("imitate") and more likely to introduce new, contrasting material ("initiate").

From the three interaction scenarios mentioned above, "imitate" and "initiate" follow the call and response paradigm, while "repeat" is the only scenario entailing synchronous action (i.e. both the human and the robotic percussionist playing simultaneously). In this scenario, auditory processing and decision-making are based on short- rather than long-term memory functions. Instead of calculating contrast metrics and generating responses on a phrase level, the robotic percussionist interacts with the musician on a sound event level, freely repeating some of the actions performed by the musician. The initialization conditions for this scenario are not dependent on contrast metrics, but a record of past scenarios, kept to ensure that it is not repeated too often.

The musician can alternate among the same scenarios as the robotic percussionist, while "navigating" a non-linear score that consists of both descriptive and prescriptive notation. The composed fragments/phrases used in the "imitate" and "initiate" scenarios are organized in three concentric rectangles according to pre-calculated contrast metrics as follows:

• *From the center outwards:* in order of decreasing rhythmic contrast,

• *From the center upwards:* in order of decreasing timbral contrast, with strokes being the predominant playing technique,

• *From the center downwards:* in order of decreasing timbral contrast, with scraping being the predominant playing technique.

This "topological" organization of the sound material facilitates real-time decision-making and interaction, allowing the musician to adapt to the robotic percussionist's actions (Fig. 1).

The material used in the "repeat" scenario is less thoroughly notated: instead of playing composed musical phrases, the musician is instructed to improvise on a set of notated actions with variable or open instrumentation and duration.

The "repeat" scenario has two variations depending on who is "leading" the improvisation: the musician or the robotic percussionist. In the former case, the musician can improvise freely, while in the latter, he/she is instructed to "repeat" the actions of the robotic percussionist ad libitum (i.e. freely).

The beginning and end of the piece are fixed and based on two differentiated instances of the "repeat" scenario in which the musician is leading and the robotic percussionist is following. Concretely, the beginning of the piece is based on a direct mapping of the input amplitude (human bowing a cowbell) and has the character of an instrumental interaction, rather than an interaction with an autonomous agent. The ending sequence of the piece, which is initiated by the robotic percussionist, is based on a repetition of detected strokes (onsets) initially with a variable delay, which is progressively reduced until only the latency of the onset detection algorithm and the actuation mechanism remains.

3.3 Action

The responses generated by the robotic percussionist are based on pre-composed sequences of onset times. The specific actions (instrument and playing technique) to be performed and their durations are chosen on the fly based on the current scenario (i.e. according to whether the current response is an "imitation" or an "initiation", the robotic percussionist might choose the same or different actions than those performed by the human).

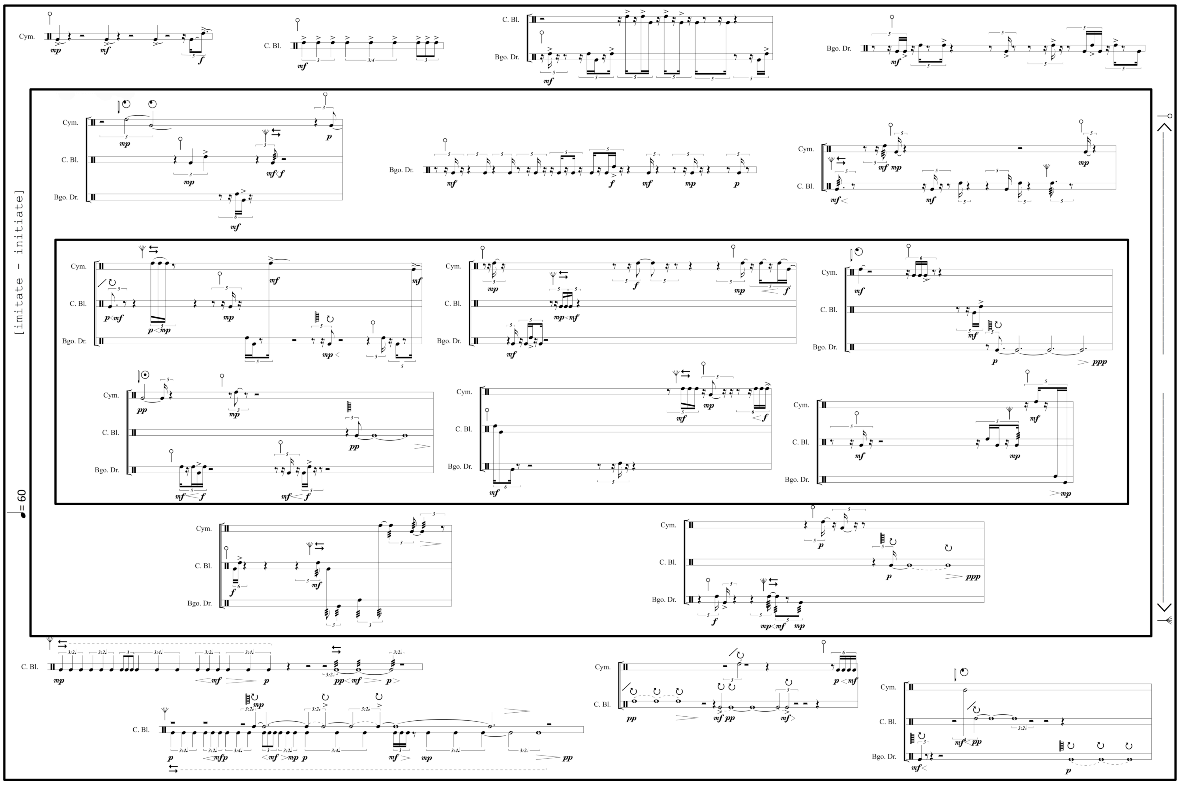

Figure 1. Score excerpt: "imitate" - "initiate".

The actions employed by the robotic percussionist include strokes and scraping and are implemented through the use of servo-motors, controlled by an Arduino UNO micro-controller, and two permanent magnets suspended over one of the cymbals and set into motion by two computer-controlled electromagnets, which are placed directly underneath the cymbal (Fig. 2).

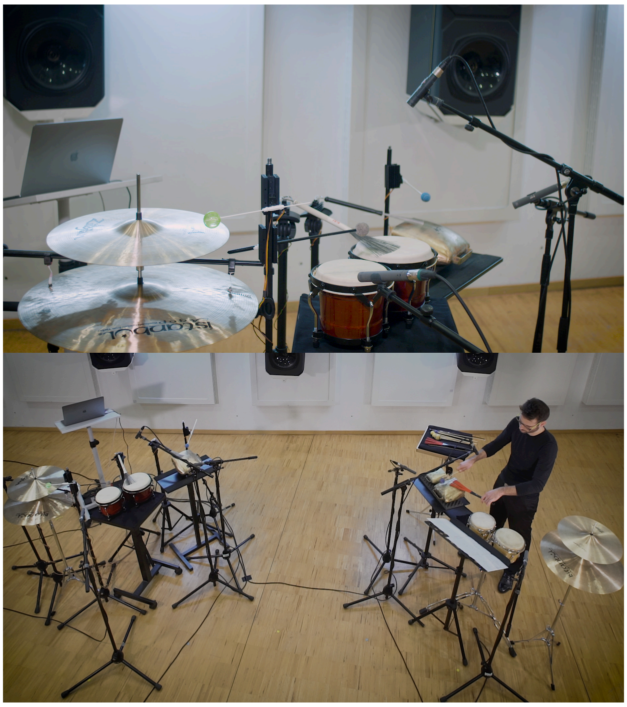

Figure 2. Robotic percussionist: instrument setup.

4. SYSTEM AUTONOMY AND RESPONSIVENESS

Decision-making in *Imitation game* is centered around two seemingly contradictory – if not mutually exclusive – agent attributes:

- *Responsiveness* or “reactivity” [16]: the agent’s ability to act in response to its environment, including human actions, and
- *Autonomy*: the agent’s ability to act independently of human actions.

Admittedly, balancing responsiveness and autonomy is a key factor and, at the same time, a major challenge in designing meaningful sonic human-computer interactions [7]. A high degree of responsiveness coupled with a low degree of autonomy is associated with linear input-output mappings – and therefore cause-effect relationships – rather than complex decision-making processes. For example, a musician would not always respond to the same sound stimulus in the same way. His/her response to it – or lack thereof – would be the result of a decision informed by the overall sonic context, rather than an independent response to the stimulus per se. Conversely, high autonomy and low responsiveness are suggestive of erratic, rather than intelligent behavior. Balancing agent responsiveness and autonomy is therefore key to designing intelligent behaviors – or at least behaviors perceived as such.

4.1 “Imitate”: Establishing System Responsiveness

In *Imitation game*, system responsiveness is established through the “imitate” mode. The robotic percussionist’s ability to play similar material to that played by its human counterpart (e.g. by choosing similar rhythms, instruments and playing techniques) suggests that the agent is not only collecting auditory information, but also interpreting it in a musically meaningful way (i.e. understanding human/musical concepts such as instrument and playing technique categories), while confirming that the agent is in fact responding to the human percussionist and not acting at random.

4.2 “Initiate”: Establishing System Autonomy

Along with responsiveness, the system also displays a high degree of autonomy, demonstrated mainly in the “initiate” scenario. Following a complex decision-making process based on an aesthetic evaluation of musical contrast, the robotic percussionist might choose to stir the interaction in a different direction, by introducing new sound material.

4.3 “Repeat”: Increasing Musical Tension

The “repeat” scenario deviates from the other two scenarios regarding both the level of compositional control (improvised vs. notated material) and the timing of the interaction (synchronous vs. asynchronous action).

This contrasting relationship is a source of *musical tension* due to both the higher density of sound events (2 “voices” instead of one) and the accelerated response time, resulting from decisions being made on a sound event rather than a phrase basis.

The co-existence of interaction scenarios in which the robotic percussionist rejects (“initiate”), confirms (”imitate”) and even mirrors (“repeat”) human actions implies an *instrument-agent continuum* – rather than a dichotomy – in which system responsiveness and autonomy are alternately established and questioned.

5. NAÏVE REHEARSALS AS A FRAMEWORK FOR ARTISTIC EXPERIMENTATION

Evaluation is becoming a topic of increasing importance for human-computer improvisation systems, with evaluation frameworks often being borrowed from other disciplines – mainly Human Computer Interaction (HCI). Linson et al. [17] argue that qualitative evaluation by experts is the most appropriate evaluation method for freely improvising interactive computer music systems and a preliminary literature review reveals that it is indeed the most commonly used method.

Brown et al. [7] adopt an iterative design process based on evaluation by expert musicians, during which they collect both quantitative and qualitative data in the form of open-ended feedback. Hsu and Sosnick [18] focus on usability, interaction and “musical results”, combining expert evaluation (“naïve” and “informed” rehearsals, as well as questionnaires) with audience surveys. In Weinberg and Driscoll’s user study [3], (expert) users are asked to interact with a robotic percussionist, participate in a “perceptual experiment” and answer a questionnaire.

While in the case of human-computer improvisation systems, these evaluation methods seem to provide interesting insight, by helping identify and subsequently address possible weaknesses, the question of their applicability to compositions is undoubtedly a complex one.

For instance, usability and interaction – both important aspects in HCI evaluation frameworks – may be irrelevant and even undesirable in the context of a specific composition. For example, in Mark Applebaum’s *Aphasia* [19] the performer (“singer”) is asked to synchronize highly detailed hand gestures to an audio tape. Since there are no sensors involved, the synchronization is left entirely to the performer’s ability to execute the score as accurately as possible. This creates a carefully composed illusion of interaction, which leaves the audience wondering whether the performance was in fact based on some kind of sensor technology. In this example, there is essentially no interaction – at least not in an HCI sense. In fact, evaluating parameters such as usability and interaction would contradict the very concept of the composition.

To complicate things further, the evaluation of human-computer improvisation systems often includes aesthetic components [3, 7, 18]. Applied to a composition, this approach could lead to a paradox, as it is based on the implicit assumption that the objective of the composition at hand is to satisfy listener preferences. But, what about the case of a musical work that does not aim to satisfy preferences, but rather question them and establish new aesthetics?

These reservations aside, some of the evaluation methods mentioned earlier can be a useful tool when used in the context of *creative experimentation* instead of a formal evaluation. In the case of interactive compositions, in particular, balancing compositional control and real-time decision-making remains a significant challenge and one that can only be addressed through extensive experimentation in collaboration with the musician(s).

In the development of the composition described in this paper, “naïve rehearsals” [18] were used as a framework for artistic experimentation throughout the creative process. In these sessions, percussionist Manuel Alcaraz Clemente was asked to improvise with the robotic percussionist, without being given any prior information on how it would respond to his actions. The purpose of these experiments was to observe and identify unintended emergent behaviors and interaction affordances. Some of these behaviors were undesirable, in which case a revision of the score and/or software would be considered necessary. In other cases, “hidden” interaction components would emerge which, though initially unintended, were considered as musically interesting and were later integrated in the composition. At this point, it’s important to clarify that what constitutes an “undesirable” or a “musically interesting” interaction component was determined by the composer and not the user/performer, since the purpose of these sessions was not an evaluation of the composition, but rather *aesthetic experimentation as part of the compositional process*.

Initially, these experimentation sessions started with a naïve rehearsal, followed by a semi-structured interview in which the musician was asked to describe his experience and the way in which the system responded to his actions in each scenario. Following this short interview, the musician was asked to fill-in a questionnaire regarding the degree of controllability, responsiveness and autonomy of the system, the degree of influence that the generated responses had on his actions, the timing (synchronous vs. asynchronous responses) and time-scale of the interaction (i.e. whether the responses were based on short or long-term changes), as well as the specific parameters of the human input to which the system was responding. Finally, the musician was asked to fill-in a similar questionnaire after participating in an informed rehearsal.

Later in the experimentation process, the format of these sessions was modified and centered around a naïve rehearsal, with observation and a semi-structured interview serving as the main data collection methods. The questionnaires and informed rehearsals were eventually abandoned, due to the limited scope of the data collected in them. Specifically, the naïve rehearsals seemed to provide a more suitable framework for experimentation in comparison to the informed rehearsals, in which the musician's actions seemed to be restricted by the capabilities of the system, instead of exploring its limits. Similarly, the interview encouraged open-ended feedback, providing useful insight that extended beyond the scope of the questionnaire.

These experimentation sessions fed back into the compositional process, fostering creative ideation. For instance, the "repeat" scenario emerged from a naïve rehearsal during which the musician mistakenly thought that the robotic percussionist was repeating his actions one by one. This misinterpretation of the robotic percussionist's actions resulted in an interesting counterpoint between the human and the robotic percussionist, which was later integrated in the composition as a separate interaction scenario.

6. CONCLUSION

The composition described in this paper assumes an anthropomorphic model of agency, which is reflected in all three stages of the robotic percussionist. Concretely, the decision-making stage is based on aesthetically-driven decisions incorporating subjective high-level percepts, while the action stage involves acoustic sound sources – instead of loudspeakers – and actuators used to simulate human actions (e.g. "strokes" and "scraping"). Similarly, the auditory processing stage is based on a dual classification task involving (human) concepts such as "instrument" and "playing technique".

The main objective of this compositional choice was to create *meaningful interactions* in which the software agent would be able to make decisions based on aesthetic criteria – instead of (pseudo-)random processes – and assume the same roles as the human (e.g. following and leading). Instead of randomly selecting an interaction mode, the robotic percussionist (aesthetically) assesses its current interaction with the musician and chooses to either "follow" him/her or "lead", by introducing musical changes.

At the same time, human decision-making is conditioned by a set of algorithmic instructions similar to those incorporated in the robotic percussionist's decision-making stage. For the human, following algorithmic instructions is the equivalent of aesthetically-driven decision-making for the computer: a task usually associated with machines, performed by a human. This trade of (anthropomorphic and mechanistic) attributes between the musician and the robotic percussionist aimed at exploring the intersection between human and computational decision-making.

As part of the compositional process, evaluation methods from human-computer improvisation were adapted into a framework for creative experimentation, fostering composer-performer collaboration. Particularly, the format of a *naïve rehearsal* [18] was used to explore unintended emergent behaviors in composed interaction scenarios. Data collected through observation and a semi-structured interview with the musician was used to inform the compositional and software development process, with the objective to balance compositional control with real-time interaction and decision-making.

The main limitation of this framework proved to be the concept of "naivety" itself: one cannot be "naïve" for too long and therefore a naïve rehearsal can only take place once. Further development of the framework could include conducting naïve rehearsals with more than one musicians and introducing additional data collection methods such as focus group discussions.

Acknowledgments

The author wishes to thank Marko Ciciliani and Gerhard Eckel for their critical readings of this manuscript. This research was funded by the Austrian Science Fund (FWF): AR 483-G24.

7. REFERENCES

[1] A. Kapur, "A History of Robotic Musical Instruments," *Proceedings of the 2005 International Computer Music Conference*, Barcelona, 2005, pp. 21–28.

[2] A. Kapur, M. Darling, D. Diakopoulos, J. W. Murphy, J. Hochenbaum, O. Vallis, and C. Bahn, "The Machine Orchestra: An Ensemble of Human Laptop Performers and Robotic Musical Instruments," *Computer Music Journal*, vol. 35, no. 4, pp. 49–63, 2011.

[3] G. Weinberg and S. Driscoll, "Toward Robotic Musicianship," *Computer Music Journal*, vol. 30, no. 4, pp. 28–45, 2006.

[4] G. Weinberg and S. Driscoll, "Robotic Musicianship – Musical Interactions Between Humans and Machines," *Human Robot Interaction*, vol. 22, 2007.

[5] G. Weinberg, S. Driscoll and M. Parry, "Musical Interactions with a Perceptual Robotic

Percussionist," *Proceedings of the 2005 IEEE International Workshop on Robot and Human Interactive Communication*, Nashville, 2005, pp. 456–461.

[6] G. Hoffman and G. Weinberg, "Interactive Improvisation with a Robotic Marimba Player," in *Musical Robots and Interactive Multimodal Systems*, J. Solis and K. Ng, Eds. Berlin, Heidelberg: Springer, 2011, pp. 233–251.

[7] A. R. Brown, T. Gifford and B. Voltz, "Stimulating Creative Partnerships in Human-Agent Musical Interaction," *Computers in Entertainment*, vol. 14, no. 2, pp. 1–17, 2017.

[8] G. Paine, "Interactivity, where to from here?," *Organised Sound*, vol. 7, no. 3, pp. 295–304, 2002.

[9] S. Emmerson, "Rebalancing the Discussion on Interactivity," *Proceedings of the 2013 Electroacoustic Music Studies Network Conference - Electroacoustic Music in the Context of Interactive Approaches and Networks*, Lisbon, 2013.

[10] G. E. Lewis, "Too Many Notes: Computers, Complexity and Culture in "Voyager"," *Leonardo Music Journal*, vol. 10, pp. 33–39, 2000.

[11] A. M. Gioti, "Neurons: An Interactive Composition Using a Neural Network for Recognition of Playing Techniques," *Proceedings of the 2018 Musical Metacreation Workshop*, Salamanca, 2018.

[12] N. Collins, "SCMIR: A SuperCollider Music Information Retrieval Library," *Proceedings of the 2011 International Computer Music Conference*, Huddersfield, 2011, pp. 499–502.

[13] D. Stowell, and M. Plumbley, "Adaptive Whitening for Improved Real-Time Audio Onset Detection," *Proceedings of the 2007 International Computer Music Conference*, Copenhagen, 2007.

[14] P. T. Ravikumar, K. Mcgee and L. Wyse, "Back to the Experiences: Empirically Grounding the Development of Musical Co-creative Partners in Co-experiences," *Proceedings of the 2018 Musical Metacreation Workshop*, Salamanca, 2018.

[15] M. Young, "NN Music: Improvising with a 'Living' Computer," *Proceedings of the 2007 International Computer Music Conference*, Copenhagen, 2007, pp. 508–511.

[16] M. Wooldridge and N. Jennings, "Agent Theories, Architectures, and Languages: a Survey", in *Intelligent Agents*, M. Wooldridge and N. Jennings, Eds. Berlin: Springer-Verlag, 1995, pp. 1–22.

[17] A. Linson, C. Dobbyn and R. Laney, "Critical Issues in Evaluating Freely Improvising Interactive Music Systems," *Proceedings of the 2016 International Conference on Computational Creativity*, Paris, 2016, pp. 145–149.

[18] W. Hsu and M. Sosnick, "Evaluating Interactive Music Systems: An HCI Approach," *Proceedings of the 2009 International Conference on New Interfaces for Musical Expression*, Pittsburgh, 2009, pp. 25–28.

[19] M. Applebaum, *Aphasia: for singer and tape*, 2014.