Factors affecting audience perceptions of agency in human computer musical partnerships

Andrew R. Brown, Toby Gifford and Bradley Voltz
andrew.r.brown@griffith.edu.au, t.gifford@griffith.edu.au, b.voltz@griffith.edu.au
Queensland Conservatorium, Griffith University
140 Grey St, Southbank QLD 4101 Australia

ABSTRACT
What design factors contribute to an illusion of agency in a computational system? Our previous research [1, 2] has investigated this question in the context of creative human-machine musical partnerships, where we identified musical behaviours implying machine agency from the perspective of a human performer. This paper investigates an audience perspective: what factors contribute to an impression of machine agency for in a musical performance? Audience feedback data was collected during a concert with four performances, each comprising a human musician interacting with a computer music system. Three performances utilized a computational agent, CIM [1], designed for this research. The fourth performance utilized an array of effect pedals designed in an extended instrument paradigm [3]. The audience feedback questionnaire queried whether a sense of machine agency was imparted, and to what degree visual, spatial, timbral and musical factors contributed to this impression. The results showed our CIM system succeeded in imparting a sense of agency, and that all four of the suggested factors contributed to that impression.

Author Keywords
Music, interaction, agency, performance, audience, partnership, generative

ACM Classification Keywords
H.5.5 Sound and Music Computing, I.2.11 Distributed Artificial Intelligence, H.5.3 Synchronous Interaction

General Terms
Human Factors; Design

INTRODUCTION
This paper reports on early investigations into two-part (duet) musical interactions, where one of the performing agents is a computer music system, and the other is a human player. We sought to determine whether or not an audience perceives that there are two performing agents, and to identify those aspects of live performance contributing to a sense of dual agency from an audience’s perspective. We consider machine agency in this context as being both generative and autonomous in ways that contribute to an ability to interact musically with a human partner.

Human-machine duets, where the machine is a music generator rather than simply an instrument, have a long history dating back to barrel organs and player pianos [4]. More recently, generative and responsive processes have been used to create digital instruments [3]. Our focus on duet interaction aims to produce a greater sense of computer autonomy than is typical in the ‘instrument design’ style of research. Our system, known as CIM, fits within the category of a synchronous duet system, where there is concurrent performance by both human and computer. In addition, we take into account more than the ‘musical’ output of the system and the ways in which the human interacts with the machine by considering performance production factors including visual stage presence, sound localization, and timbral differentiation.

The behaviour of the CIM system is based on a model of musical duet interactions we outline in another paper [5], and briefly outline here. This model emerged from analyses of notated and improvised duets, and identified various musical activities and cues for activities that characterize duet interaction. Six contextually generalizable activities were discovered:

- Initiate - introduce new musical material
- Loop - immediately repeat own material
- Echo - repeat partner within a short time window
- Restate - reuse own or partner's earlier material
- Shadow - play in unison or close parallel with partner
- Silence - don't play

The CIM system is capable of undertaking these musical activities. Changes of activity are triggered by cues from and analysis of performed input from a human player using a MIDI instrument. Currently, the system detects segmentation in the human performance and uses this, together with note information from the human performance, as a cue to engage in, or change activity. This
paper reports on using the system to present three contrasting duet improvisations in a concert performance, and audience experiences of this performance.

AUDIENCE RECEPTION STUDIES
A study by Stevens et al. examined audience responses to dance performances. They were interested in assessing the audience’s “cognitive, emotional and affective” [6] response. They developed a number of approaches including the Audience Response Tool (ART), which included a mixture of assessment devices for measuring both qualitative and quantitative responses. Our method resembles the ART approach in that we collected both qualitative and quantitative responses via an audience questionnaire for the purpose of understanding audience perception of machine agency in duet performances with a human performer.

Edmonds and Muller investigated audience engagement with interactive visual art works in an exhibition setting. This work is relevant because it also examines interaction in a creative context and focuses on works using generative computational processes. Their artistic research interest is similar to ours, in that the computer programs do not need to “stand as some kind of scientific theory of human thinking” but can be considered within a cybernetic frame “to model and implement animal-like behavior” [7]. For our study, we elected to focus on the audience recalling and unpacking their experiences, in the belief that this would provide sufficient data concerning the degree to which a sense of machine agency was imparted.

An approach to the evaluation of improvised music was outlined by Eisenberg and Thompson [8] in which they used the Consensual Assessment Technique (CAT) [9] to manage the complexity of aesthetic judgments and to connect features of improvisation to audience evaluations. We have employed the CAT technique to show the degree of consensus amongst our audience, applying this technique to evaluate the audience member’s perception of dual agency in the performance rather than to assess creativity or quality. This approach enables us to look for agreement amongst the audience without needing to understand completely the reasons for people arriving at these conclusions. We, like Eisenberg and Thompson, are also using the CAT analysis to help us clarify some of the characteristics—in our case the role of visual, spatial, timbral and musical indicators—that might lead to this judgment.

THE PERFORMANCE
The CIM system is visually presented to the audience in the form of a computer monitor which displays a series of colours, each corresponding to the musical activity that the system is currently engaged in. To further aid the audience’s perception of two performing agents, each was positioned apart on the stage and sound placements for each reinforced this physical separation. Contrasting synthetic sounds were used for CIM and the human performer to assist the audience in differentiating between parts.

We presented four performances, each comprising a human musician interacting with technology. The first three performances utilised the computational agent CIM. In the first and second performances, the human musicians played an electronic keyboard; the third performance was given using an electronic wind instrument. The fourth performance involved an electric guitarist utilizing an array of effects pedals. This guitar-with-effects performance was included to provide baseline data to gauge the degree to which the audience perceived there were two performing agents in the other performances with the CIM system. Performances were all improvised but varied stylistically, with specific sound sets used for each. Videos of these performances are available online.1

EVALUATION METHODS
Audience feedback on each of the four performances was gathered in the form of a survey, designed to provide a mix of quantitative and qualitative data concerning the degree to which they perceived two musical agents engaged in duet performance (two-part interaction). The following statements were provided, and members asked to respond using a five point Likert scale ranging from ‘strongly disagree’ to ‘strongly agree’.

1. There was a clear sense of two-part interaction.
2. Visual elements contributed to this impression.
3. Spatial audio elements contributed to this impression.
4. Timbral elements contributed to this impression.
5. The music played contributed to this impression.

In addition, short written responses were sought in respect of interactive aspects of the performances, and general comments concerning the creative elements. The sample audience included adults experienced with arts and technology but with varied degrees of musical background.

Reliability of the audience responses
To gauge the reliability of our audience evaluations - regarding the sense of agency exhibited by our computational system in this performance setting - we draw upon Amabile’s [9] Consensual Assessment Technique (CAT). This technique is designed for questionnaires that probe subjective experience, seeking consensus amongst subjects. Following Amabile we evaluated the “InterJudge Reliability” of each performance using Chronbach’s Alpha [10], yielding the following data:

<table>
<thead>
<tr>
<th>Performer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
</table>

1 http://www.youtube.com/user/CIMSResearch
A value of Chronbach’s alpha above 0.6 is generally taken to indicate ‘substantial agreement’, and above 0.8 ‘almost perfect agreement’ [11]. Thus the interjudge reliabilities for Performers 3 and 4 are in almost perfect agreement, whilst Performers 1 and 2 are in substantial agreement. Amabile recommends use of expert judges for CAT evaluations, and whilst our audience members were non-expert listeners, the interjudge reliability scores using non-experts provides us with confidence in the results.

RESULTS AND DISCUSSION

Of the fifty audience members who completed the survey, the majority indicated in their quantitative responses that they perceived a sense of two musical agents in the first three performances with the CIM system, albeit ranging in degree, while the sense of two musical agents in the fourth performance with the guitar and effects pedals was significantly lower.

The results of the quantitative part of the survey (the Likert scale questions) are shown in Table 2 below, along with a brief summary of the qualitative part (free text feedback). Performers 1 - 3 were interacting with the CIM system, whilst performer 4 (the ‘control’) was playing an extended traditional instrument.

<table>
<thead>
<tr>
<th>Reliability</th>
<th>0.63</th>
<th>0.68</th>
<th>0.82</th>
<th>0.85</th>
</tr>
</thead>
</table>

Table 1: InterJudge Reliability

For the remaining questions (Q2-5), the Kruskal-Wallis median difference test showed significant (p < 0.01 **) difference from neutral (“Undecided”) for each performer. Audience members were generally undecided about the contributions of visual elements, offering comments such as “the visuals need to be improved”, and that they were “not a rich representation”. We interpret this data not as a suggestion that visual elements are unimportant, but that they require further refinement.

There was agreement concerning spatial elements of the performances (Q3), and the location of sound sources in the performance was remarked upon, however, these responses were infrequent. This may indicate that the question was not well understood. For example, “[I am] not really sure what is meant by spatial elements, perhaps some notes were left speaker?” We suggest that, due to the acoustic properties of the performance space in combination with technical restrictions in the selection of amplification systems, the audience’s perception of a localized, independent sound for each musical agent may have been compromised.

Similarly, the bulk of respondents agreed that there was a substantial contribution by timbral elements (Q4). Written responses to this question were limited. Aside from a couple of comments concerning sound quality, such as “sound sample was poor, a little antiquated”, the feedback mainly related to the timbral differences between each part, with audience members noting that “tonal elements were distinct”.

Musical elements (Q5) were seen as important, and in particular pitch content, timing, style and relationship between the parts formed the basis for respondent’s comments to this question. Identifiable musical content (like when CIM repeats a music phrase by the performer) and differentiation of activity types between the human performer and the CIM system assisted (e.g., a combination of initiating and echoing) reportedly help the audience differentiate each part. They specifically commented favorably on “mirrored melody, [and] repeated notes” in the computer part and on the improvisational skills of human performers.

Comments on the Creative Partnership

Qualitative feedback was requested about the creative partnership. Eight audience members commented on the combined aesthetic effect of the performances, observing that the performances were interesting and that there was potential in pursuing the study. However, one respondent noted, less positively, that the feeling of a partnership “was largely contrived”, and that “vertical (harmonic) information did not carry through.”
It was clear from the responses to question 1, that in the performances with the CIM system the CIM’s contribution was recognized as a second part, while in performance 4 between human and various guitar effects pedals the pedals were not seen in this way. This reinforces an observation made by Rowe [13] that interactive music technologies fall on a continuum between behaving as instruments or players. In this case the effects pedals were perceived as instruments and the CIM system as a player.

Survey results also showed that each of the four factors focused on in this study—visual, spatial, timbral and musical—were seen as important in discerning two performing agents. These comments were heavily weighted towards the visual and musical aspects of the performances, with relatively few remarks provided in respect of spatial and timbral characteristics.

While an audience comprised mainly of musicians may have focused more on the timbral and spatial aspects, it is possible that these factors were conflated with comments about ‘musical’ aspects by non-expert musicians. More broadly however, it seems clear that what an audience ‘sees’ is an integral part of live performance, not only what they hear.

Our results are consistent with perceptual studies more broadly and Gestalt psychological theories that emphasize the tendency for people to look for connections between elements and try to make meaning of situations, even where it may not have been intended [14]. In this study audiences made many assumptions, especially about the role of unrelated equipment on stage. Many found it difficult to distinguish clearly between programmed and coincidental coordination between the human and computing system; both sonically and visually. For example, at times a technician was positioned near to or on the stage leading to comments like “[the] second person did not add to [the] performance.” We acknowledge that as system developers we may need to put less emphasis on complex musical processes inside the computing agent but pay more attention to clarifying the presentational aspects of the performance.

Despite a range of subjective evaluative comments from audience members, the CAT results indicate considerable agreement about the degree of duet interaction and a sense of agency from both the human and computer performers.

CONCLUSION

This paper has focused on audience reception of duet performances with an interactive music system. Audience survey data clearly reveals that members formed an impression of machine agency, and that each of the four factors focused on in this study—visual, spatial, timbral and musical—were seen as important in discerning two performing agents. While this has been a small study in a specific musical context, it points to implications and opportunities for further research on human interactions with generative computational systems more broadly, and underscores the diversity of factors that contribute to audience perception of agency.

REFERENCES