Towards Augmenting Communication in Human AI Music Improvisation, review

Henrik Frisk

# Review

## Is the subjenct matter and presentation appropriate for this special issue? Please explain if not.

It is stated in the introduction that the authors are interested in "understanding to what extend we can use AI to challenge and provoke human creativity, and to understand how we might build systems that can enable new kinds of creative practices between humans and AI systems collaborating in creative contexts". This aim should be appropriate for the special issue, but the design of the study is, it appears to me, rather geared towards understanding the musicians experience with performing with an AI. The actual system is neither described in enough detail, nor is it clear how it is used to challenge and provoke the participants.

## Does the manuscript meet its stated goals; are they meaningful for the material? Please explain if not.

The manuscript does meet its stated goals but extensive editing is needed. The paper discusses the ways in which humans and AI music systems can signal to each other but the focus is not on the AI, the ways it functions, or on the system itself. Instead the authors have performed a thematic analysis based on a group of improvisers and their experiences with playing with the system. The thematic analysis is well done and the method is, in general, used wisely. However, there are a few facts that makes it problematic:

1. First, although many of the references point to jazz improvisation, it is never stated what kind of musical practice the study is performed on. This reduces the usefulness of the results to a significant degree as the term 'improvisation', widely used throughout, is ambiguous at best. My guess is that the authors take for granted that the study deals with Western music, which is problematic by itself. The interactive requirements and expectations that a performer would have on an improvising AI varies to a great degree with what kind of improvisation the musician is comfortable with performing.
2. Similarly is it not clear whether or not the participants are all jazz musicians or what their musical expertise is, only that they are experienced musicians. If the reader can't tell if the participant is a fiddle player, a concert violinist or a free jazz improviser, many of the results are less useful.
3. The description of AI system, in essence a VMM, is solely focused on how the system chooses the next output based on the input but not in any way in relation to a musical logic. This is related to the first point above: not knowing what kind of musical system is being modelled, or why it has been designed in the first place (possibly only for the purposes of doing this study?), here makes it difficult to understand the analysis later in the paper. It is described to use MIDI only, and that the participants play a piano sample sound, but only a third of the participants are actually pinaists.

## Are the ideas presented clearly; is the presentation well organized? Please explain if not.

See above. There are a range of minor language errors and idiomatically incorrect phrasings. I believe, especially for this context, that the system should be better described. That it is using markov chains says very little about the way it works: does it use harmony? How is rhythm generated? Melody?

The structure of the paper is a bit unorthodox which makes this reader lose focus. The concept of silence, which appears very important to the authors (I am less convinced) is introduced after the description of the results of the study. It would be better to collect the theory in one place, before the presentation of the study, and state the suggested importance of silence already from the start.

## Have the authors ignored significant work or contradictory results? Please explain if not.

The results from the thematic analysis are well described but the way the authors connect them back to the aim is a bit vague. Since improvisation is mentioned the text would benefit from a more thorough overview of research on musical creativity in improvisation such as (Benson, Bruce Ellis, 2003, Borgo, David, 2005, Peters, Gary, 2009) to only mention a few key works. However, I am only guessing as to what the author's particular musical interest is.

## Is the manuscript significantly different from previously published work? Please explain if not.

I believe this manuscript is different from previously published work.

## Does the manuscript require any additional audio and/or video material in order to be evaluated by you or understood by readers? Please explain if so.

The thematic analysis would be greatly enhanced with music examples and/or notation that exemplifies what kind of musical phrase generated the code. Without this the paper, in my opinion, falls short.

## Recommendation: For this special issue of Computer Music Journal, the manuscript should be (select one):

X published after reviewing major revisions

## Please justify your recommendation, and provide general feedback about the manuscript.

Should the paper be complemented with the things I feel are missing, it could be published, but in its present form it should be rejected. What should probably be weighed in here is whether or not the strengths of this paper–the thematic analysis of musician's experiences with an automated system–is within the scope of the theme of this issue. Had the description of the system been better described this question would have less importance, but as it is, the paper has an emphasis towards musicology rather than computer music studies.

## if you have any confidential comments for the editors, please enter them here:

In summary, I am hesitant to whether this paper fits the theme, but this is something you obviously are better to judge.

## Overall comments

This paper discusses the ways in which humans and AI music systems can signal to each other

## This points to the complexity of improvisation.

The ability to improvise on stage with others is considered to be one of the great achievements or examples of human creativity. It requires empathy, quick-wittedness, an awareness of the genre, an awareness of the audience and each other, and a huge source of patterns that can be used appropriately given the correct contextual settings.

## This source is not necessarily relevant here, look at some of the references that discusses non-idiomatic improvisation.

``musical state and intention to others Seddon (2005).''

## It is not clear what 'this' refers to here: what musical experience?

``It is this notion of the musical experience''

## What previous work?

``from previous work by focusing on the exploration of''

## Reference?

This is a quite far reaching statement that is not really backed up. What is the basis for this conclusion?

, as improvising agents are implicitly (and sometimes explicitly) required to explain any decisions they are making,

## What were the participants experience with improvisation?

Participants: We conducted a qualitative study with nine participants (n=9) with a different range of music experience.

## How was the MIDI signal generated from voice and guitar?

While all participants could play at least 2 instruments, 3 considered that piano was the instrument they played best, 3 selected guitar and the other 3 were best with voice.

## Something wrong with the interpunctiation

``Table 1 shows the results of calculating the scores for the three selected factors using the GMSI Scoring App gms (b) and subsequently placing them within a percentile chart of the general population for which the index was generated gms (a); for instance, […]n''

## This section is not relevant here

Insights gained from understanding musicians’ physical mechanisms when improvising together provide a valuable resource to design communication mechanisms that facilitate improvisational interaction between human and AI musicians.

This section speculates on how "Insights gained from understanding musicians’ physical mechanisms when improvising together" can be useful and how the interface of an AI system may be improved, but this is not something that you have studied. I would suggest this is removed or moved to another section of the paper.

## On the thematic analysis

The theme Structured Improvisation explores the mechanisms participants identified that could provide more musical structure and help increase the quality of improvisational interactions with the system.

The usefulness of the thematic analysis, which is generally interesting and well performed would be furthered improved if the reader was allowed insight into what kind of musical gestures/interaction were coupled with the various tags. What is, for example, structured about this tag in a musical sense.

## Thematic analysis

The code system signalling to user explores the concept of bi-directional communication by giving the system autonomy to send signals to the human musician.

Since this is a thematic analysis, a code given cannot said to be given the system autonomy. the code, unless I've misunderstood something, signals that a certain aspect of autonomy in the system was observed by the improvisor.

## This statement could be interesting to unpack a bit more

, facilitating the discovery of the inner workings of the system should be a design consideration for improvisation music systems.

What follows here, the example of the Shimon marimba robot and the use of emoticons appear to rather be systems that improve multimodal communication in performance and perhaps not so much the inner workings of the system?

## Silence

``The role of silence in human-machine communication''

It is a bit odd that this section, in principle exploring new theory comes after the discussion. Now, the discussion should perhaps be renamed *Discussion of results* but even so, what is the reason for not placing this section within the rest of the theory?

## Is this a wish are a normative statement?

Ultimately, in any performative musical context there will be visual and (less frequently) vocal gestures which allow musicians understand the direction of the performance.

It is absolutely possible to perfrom improvisations in complete darkness or with eyes shot.

## What does 'here' refer to?

``f communication here is''

## Repetition

the Shimon marumba robot, which uses a human-like head that turns to show that it is ‘listening’ to the human performer.

## Does 'presence' provide information about the inner working sof the system?

Here we propose that all music improvisations systems should employ some type of ’presence’ that embodies the system in order to communicate aspects of the inner workings of the system, but also we propose that the interplay of moments of silence with some kind of embodiment, or for that matter, any kind of ‘presence’ that is implemented for an improvisational system (e.g. a digital visualisation), would play a significant role in conveying emotional meaning.

## It is not clear to me how this reference is relevant

Embodying silence in a music system can give it a more powerful effect in communication, invoking reactions from musicians and audiences alike in the performance.

## The evidence for this is weak in the paper

However, it has been extensively studied in the social sciences as a medium that carries meaningful and distinct information (depending on the context surrounding it).