Creative AI and Musicking Robots

Craig Vear [1][0000-0001-7879-7978]

[1] De Montfort University, Leicester, UK cvear@dmu.ac.uk

Abstract. This paper discusses the creative and technical approaches in a performative robot project called *Embodied Musicking Robots* (2018 – present). The core approach of this project is Human-centred AI (HC-AI) which focuses on the design, development and deployment of intelligent systems that co-operate with humans in real-time in a 'deep and meaningful way'. HC-AI is 'defined by two goals: (1) the AI system must continually improve by learning from humans while (2) creating an effective and fulfilling human-robot interaction experience' [1]. This project applies these core goals as a central philosophy from which the concepts of Creative AI and Experiential Learning in the context of performative robots are developed. At the centre of this discussion is the articulation of a shift in thinking of what constitutes Creative AI and new HC-AI forms of computational learning from inside the flow of the shared experience between robot and human. The central case study investigates the technical solutions and artistic potential of AI driven robots co-creating with an improvising human musician in real-time. This project is ongoing and is part of the Creative AI Research Group in the Institute of Creative Technologies at De Montfort University [2].

Keywords: Creative AI, Music robots, creativity, human-centred.

1 *Embodied Musicking Robots*

1.1 The goal

The aim of this practice-based research project is to investigate the technical solutions and artistic potential of AI driven robots co-creating with a human musician in realtime. This is an ongoing project that was initiated in 2014 at a workshop *Embodied Intelligence in Music* that investigated the meeting point of embodied cognition, artificial intelligence, music composition and performance, game and software philosophy, in order to develop the parameters of a new field of research.

The overarching research question with *Embodied Musicking Robots (EMR)* is:

If we want robots to join us inside the creative acts of music then how do we design and develop robot systems that prioritise the relationships that bind musicians inside the flow of music-making?

This question comes from the deep and meaningful experiences I had as a professional high-level musician for over 30 years and supports one of the core goals of HC-AI. Overall, I am curious about the technical innovation to make this happen; but more so by the new avenues of creativity that can open up to us humans. The overarching goal of this research is not to make machines more creative, but to make humans more

creative through deep and meaningful relationships with co-operating intelligent machines.

1.2 Definitions

Before I describe the solutions that I designed and deployed in the *EMR* project I need to simply define what I mean by the following terms within the context of this project. This defining process also helped consolidate the design and development of the Creative AI and the robotic systems with the goals of HC-AI.

Musicking – the creative acts of realtime music-making. Musicking is a term first created by Christopher Small to define a perspective that 'to music is to take part' [3]. Small wrote that taking part can happen 'in any capacity' [4] such as performing, composing and listening. Crucially meaning formation through musicking is formed in the relationships that are established within the realm of taking part with spaces, agents, sounds, and presences that are encountered here.

Flow – is the experience of musicking from inside the activity. Within the context of this project, the flow of musicking defines how 'musicians become absorbed in the music through a sense of incorporation within their environment (the soundworld), a shared effort (with the digital, virtual, AI and robotic agents) and a loss of awareness of their day-to-day wakefulness and bodily self-consciousness (embodiment with their instrument and into their music) [5].

Embodiment (in music) – the process in musicking of drawing the musician's sound into their bodily sense-of-being. This presumes that when musicians make music it is not a process of outputting sound into the world, but an embodied experience of *becoming* the sound they create in the flow of musicking. Equally, it describes the process of the musician reaching out from this sense of becoming and drawing in the sounds of others so they feel their *presence-as-sound*. This is a dance of sorts: to touch, to feel, to sense, to work with, to play with, to hide and seek and flirt and subvert, with others through the *flow*.

Creativity & Co-operation - when play turns into invention within the flow of musicking. However, creativity through play-involvement is not a constant, nor is it reliable or automatic; it needs nurturing with open, generous and cultivating energy. On the other hand, it can sustain bold and mischievous challenges or seemingly disruptive engagement designed to rail-road on-going trains-of-thought - so long as these are still giving in their nature. Within this project I define three sub-domains of creativity to highlight the human-robot relationships. These are based on my general experience as an improvising musician, and are used to identify the types of co-creativity within musicking from the human musician's perception (note – this project does not deal with notions of machine consciousness or perception):

Concurrent – a sense that both agents (human & robot) are playfully inventing in isolation but within the shared flow of musicking

Collaborative – a sense that both agents are contributing to a shared play idea, feeding a sense of collective invention through individual contribution and perspective

Co-creative – a sense that the robot and human agents are collectively inventing through a stimulated sense that each are in inside each other's head. By that I mean that

the relationship, as perceived by the human musician, is that the robot/AI is in-the-loop with the human and together they are inventing on a singular idea, feeding each other's play as if it were one train-of-thought.

Creative AI - includes practices that have AI embedded into the process of creation, but also encompasses novel AI approaches in the realisation and experience of such work.

With these definitions in mind the goal of the *EMR* project is to design and develop a Creative AI system that enhances human musician creativity by stimulating, inspiring, interacting and co-operating in the flow of embodied live improvised music-making. Therefore, to build a robot driven by a Creative AI system, it must (1) continually improve by learning from humans while (2) create an effective and fulfilling human-robot interaction experience [6].

2 The Project

The hypothesis to the research question posed above, involved the design, develop and deploy a robotic Creative AI that would have a presence within the co-creativity of the flow of musicking, and not be an AI-zombie. This approached reinforces the personal understanding that when a musician enters the world of musicking, the 'I' is coping in a very different world-of-concern than if they were walking down a street. In a sense 'I' becomes a different creature with a different set of priorities and concerns, outlooks and sensorial inputs than my normal, human wakefulness. The technical and artistic solution for *EMR* focused on a robot that was first and foremost a coping entity in this specific world-of-concern (the flow of musicking).

The solution was to develop a system based on these three principles: **Coping** – *EMR* needed to cope in real time within the realm of musicking, and be present as sound whose movements are embodied within such flow. This required a non-representational approach to how it related to the flow as the coping mechanisms needed to be open and dynamic enough to co-operate in any given musicking realm. Limiting the robot to a single representation of what musicking is, or might be, imposed onto the system by the human designer(s), would only work in a number of instances. **Creative AI dataset and Experiential Learning** – these concepts needed to be designed from within the realm of musicking, prioritising the phenomena of being inside this realm, and capturing an essence of what it means to be embodied within the flow. The concept of Experiential Learning was designed to support this(discussed below). **Belief** – the robot needs to believe in its view of the musicking world through limitations, embedded aesthetics and behavioural traits, even glitches and bugs in the system (not discussed in this paper).

From the human-centred artistic perspective *EMR* needed to address the following: the robot was not an extension of the musician; but should extend their creativity; the robot should not be an obedient dog or responsive insect jumping at my commands or impetus, but a playful other; it should not operate as a simulation of play, but as a stimulation of the human's creativity; it is not a tool to enhance the human's creativity, but

a being with presence in the world that they believe to be co-creating with them, and it should prioritise emergence, surprise, mischief, not expectation.

3 Technical Solution

The design of *EMR* was informed by two early papers by the robot innovator Rodney Brooks, specifically *Intelligence without Reason* [7] and *Intelligence without Representation* [8]. In these he lays out the foundation of his approach to designing and building robots that are first and foremost able to cope and therefore adapt to a dynamically changing environment within the parameters of specific and multiple goals. This research eventually led to his robots being used for space and sea exploration, military and medical application and the iRobot Roomba vacuum cleaner series. These Roombas are designed with *iAdapt* AI to be 'creatures' that cope in a specific world-of-concern; in real-time. They do not have a model of representation of their world (such as building a 3D model of the space through computer vision and object analysis), nor do they make one as it goes about its business, but uses goals and strategies to cope with whatever that world can throw at it (furniture, steps, chairs that get moved one day to the next).

Brooks' foundational theories, and observations of my own Roomba, guided the developed for my *EMR*, and generated this set of principles (adapted from Brooks [8]): *EMR* must cope in an appropriate musical manner, and in a timely fashion, with the dynamic shifts inside the musicking world; *EMR* should be robust to the dynamic environment of musicking, it should not fail to minor changes in the properties of the flow of musicking, and should behave appropriately to its ongoing perception of the flow; *EMR* should maintain multiple goals, changing as required and adapting to its world by capitalising on creative opportunity; *EMR* should do something in the world of musicking, 'it should have some purpose in being' [9].

The multiple goals were (in order of priority):

- *Self-preservation* – the robot must avoid obstacles, not crash into the other musician or fall off the stage.
- *Instinctual behaviour* – if left alone the robot would make music. This was driven by the Creative AI dataset (discussed below), which operated as its DNA of musicking creativity.
- *Dynamic Interaction* – the robot can, in certain conditions, be affected by the sound of the live musician. Using a process of simulated affect linking [10] the Creative AI could leap between related, abstracted or unexpected datasets. Metaphorically, the robot's internal trains-of-thought would be triggered by phrasing (short-term temporal limits) and the dynamic impetus of the human.
- *Short-term memory* – the robot would have a memory buffer of the live music, which would influence its responses and operate as a thematic repository to be called upon as an option.

A critical feature of the design of *EMR* is that each of these goals directly move the wheels. It was essential that each goal is not part of an elaborate, logically-flowing,

representation of a thought process, mimicking some-kind-of-mind. As such, the overall design of robotic system was modular, with each system directly accessing the wheels when in operation.

3.1 Creative AI dataset and Experiential Learning

This project innovated a different approach to computational learning that involved a human in-the-loop and an in-the-groove approach. This Experiential Learning (EL) approach trained the AI on-the-job and crucially inside the flow of embodied musicking. Furthermore, that the EL process collaborated with a human musician who was equally learning about this new musicking system. This approach supported both the human and the AI to automatically learn and improve from experience.

The EL process focussed on capturing the physical phenomena of an improvising human musician in the flow of creative musicking (see appendix 1). The sensing mechanism used 3D depth tracking of the human musician's body using a Kinect sensor (simply x, y, z movement of both hands, body centre and head), and Fast Fourier Transfer (FFT) analysis of the live sound (fundamental frequency and amplitude). The resulting dataset reflected the position and rotation of an embodied musicking body in motion together with the amplitude and frequency analysis of the actual sound made by such movement, without preserving the performer's mass, musculature, melodic shape or music. Thus, embodied musicking movement is extracted from the performer's body while they are making music; in a poetic sense the dataset contains the meta-level DNA of musicking without the specifics, or a representation of the music. Recorded audio-video capture of the music performance would always anchor the dataset to a specific person and point in time, whereas the meta-level data could become the building blocks for the virtual composition. As data phrases can be edited, treated and re-purposed by the robot's AI again and again without the risk of repetition.

The initial process involved seeding the Creative AI dataset by capturing the live performance of an improvising musician. Once a small set had been generated, this musician then worked with the robot through a series of training sessions, with this new live data being added to the Creative AI dataset. The more they worked together the more meta-level DNA of their shared creativity would be put back into the dataset system, thereby improving the AI's knowledge base of experience of embodied musicking.

The EL approach learns through embodied interaction inside the flow of musicking. It prioritises the capture of the meta-level DNA of its improvising partner – the human musician – and extracted elements from the dataset (any data randomly chosen by the system as it is all endowed with meta-level creativity) into its AI processing, and then outputs the resultant sound as music. This EL process enhances the dataset through experience by its embodied coping inside the flow of musicking. The human musician perceives meaning in the robot's musicking who in turn co-operates in the making of music (generally perceiving the relationship through one of the perspectives of creative co-operation discussed above) and responds with a creative solution through music. This is then captured using the sensing mechanisms and stored back into the Creative AI dataset. Thus, the cycle of EL continues to enhance and improve the dataset and

enlarging the Creative AI memory bank of deep and meaningful interactions between human and robot which in turn forms the basis for future interactions.

4 Discussion

Using the principles outlined above the *EMR* project has created a co-creative system that responds to the interaction with a human musician through a cyclical relational process. It is important to note that the interaction with the musician begets movement as its primary goal for musicking, and that this movement is embedded with essence of embodied musicking because of the Experiential Learning process. Following this, the movement begets sound, which begets music such that all relationships between human and AI are informed by phenomenon data captured within the embodied flow of music-making: either from the Creative AI dataset or through live interaction.

The ultimate goal of this research is *not* to find solutions to replace human creativity, but to enhance it and move it forward into new discoveries. In short, this research is seeking to find experiences like those emergent through DeepMind and Alpha Go's interaction with the professional Go players. In the 2019 film [10] several of these professionals reflected that when they played with AlphaGo they 'see the world different […] Maybe it's beautiful', and 'like move 37, something beautiful occurred there"; "in a broad sense move 37 begat move 78 begat a new attitude, a new way of seeing the game ... he improved through this machine, his humanness was expanded after playing this inanimate creation' [12].

4.1 Further challenges

Current developments in *EMR* concentrate on five branches of investigation. The first is an *EMR* performing with an *EMR* in a series of "boxing matches" that resulted in recorded albums [13 & 14]. The second is translating the symbolic AI script from Max/MSP to Python. This opens out possibilities in the third branch which involves training Recurrent Neural Networks (RNN's) with the raw Creative AI dataset and for these to control the movement of the robot. The current design of this AI system is for the live audio signal to still influence *Instinctual behaviour* and *Dynamic Interaction* (discussed above) but for the data to be generated through prediction in the RNN and for this to move each wheel, which in turns begets sound, which begets music, as described above. The fourth branch is an investigation into the nature of this dataset. A research project funded by a Human-Data Interaction Network grant [15], is designing and building a more sophisticated dataset aiming to capturing a truer picture of embodied musicking. This Creative AI dataset research is ongoing and due to finish in 2021. And finally, the fifth is bringing together a large ensemble of recycled robots into an improvising orchestra, due to premiere at the Art-AI festival in 2021, with a possible follow-on performance at the *Dubai Expo2020* in 2021.

References

1. MIT (2019) Website. https://hcai.mit.edu last accessed 2020/10/23
2. Creative AI Research Group Homepage https://www.ioct.co.uk/creative-ai, last accessed 2020/10/23
3. Small, C. (1998) *Musicking*. Middletown: Wesleyan Press.
4. Ibid.
5. Vear, C. (2019) *The Digital Score*. New York: Routledge. pp 192
6. MIT (2019) Website. https://hcai.mit.edu last accessed 2020/10/23
7. Brooks, R. (1991) *Intelligence without Reason* https://people.csail.mit.edu/brooks/papers/AIM-1293.pdf last accessed 2020/10/23
8. Brooks, R. (1987) Intelligence without representation. *Artificial intelligence*, *47*(1), 139-159.
9. Ibid.
10. Gelertner, D. (1994) *The Muse in the Machine*. New York: Free Press.
11. AlphaGo (2017), Film. Director Greg Kohs.
12. Ibid.
13. 'Patabots (2019) *The Seven Pleasures of Pris*. Music album. https://patabots.bandcamp.com/album/seven-pleasures-of-pris last accessed 2020/10/23
14. 'Patabots (2019) *The Voight-Kamff Test*. Music Album. https://patabots.bandcamp.com/album/the-voight-kampff-test last accessed 2020/10/23
15. HDI Network homepage. https://hdi-network.org/ last accessed 2020/10/23

Appendix 1

The EL process focussed on capturing the physical phenomena of an improvising human musician in the flow of creative musicking