Audio-Matic: Exploring Artificial Intelligence Through Generative Music Composition

Elgin-Skye McLaren
Simon Fraser University
Surrey, BC
emclaren@sfu.ca

Abstract
Artificial intelligence (AI) is increasingly ubiquitous. Yet the public remains apprehensive of AI systems. Ethicists recommend users interact with, and reflect on, complex AI applications. With this in mind, I propose an interactive generative music system, *Audio-Matic,* as a novel means of encouraging playful, accessible experiences with AI technology. Users interact with *Audio-Matic* through a piano keyboard, and the system responds by creating new melodies based on user input. These melodies are generated using Google's deep learning tool Magenta. The generated response is played back for users through a series of servo motors mounted on a stringed instrument. In the following paper I provide a theoretical framework, propose design recommendations for this type of generative music system, and describe how these recommendations were applied in the development of the *Audio-Matic* prototype. I conclude with a discussion of possible future work.

Author Keywords
Generative audio; Artificial Intelligence; Deep Learning; Interactive Systems; Tangibles; Music;

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User Interfaces –*Evaluation/methodology; interaction styles.*

Introduction

Artificial intelligence (AI) is increasingly a part of our everyday lives. Many of us interact with AI systems on a daily basis without even realizing it. Meanwhile, much of the public is apprehensive about AI, influenced largely by pessimistic and alarmist media [19]. As AI grows more ubiquitous, ethicists propose that measures be taken to ensure that the public is aware of how the technology works [12,18]. This means ensuring that people, especially those from non-technical backgrounds, have an opportunity to interact with and reflect on complex AI systems. Herein lies the challenge for researchers. AI systems are, by definition, technical. If understanding how these systems work is important, how do we create entry points which are accessible to all?

I propose an interactive generative music system, *Audio-Matic,* as a novel means of encouraging playful engagement with a complex AI tool. Users interact with *Audio-Matic* through a piano keyboard. The system responds by creating new melodies based on user input. These melodies are generated using a type of artificial intelligence called deep learning. The generated response is played back for users through a series of servo motors mounted on a stringed instrument. Other generative music systems have proven effective in appealing to users and audiences [2,11,16]; however, to my knowledge no other systems have featured call and response interaction using a deep learning tool and a tangible stringed instrument. In the following paper I provide a theoretical framework, propose design recommendations for this type of generative music system, and describe how I applied these recommendations when developing the *Audio-Matic* prototype. I conclude with a discussion of proposed future work.

Theoretical Framework

From a theoretical perspective, a tangible generative musical system could reduce a user's apprehension by offering a playful and accessible point of entry to complex AI. This might be particularly appropriate for people from non-technical and non-musical backgrounds. In this section I briefly review relevant theory and define key terms.

Interactive Tangible Installations

Tangible interaction facilitates learning through hands-on experimentation. It also has unique advantages when compared to screen-based interaction. According to Shaer, "tangible interfaces have an instant appeal to a broad range of users. They draw upon the human urge to be active and creative with one's hands and can provide a means to interact with computational applications in ways that leverage users' knowledge and skills of interaction with the everyday, non-digital, world[15:4]." Tangible interactions respond to a user's natural inclination to touch and be tactile.

Museums and art galleries use tangible interactive installations to engage visitors. This is increasingly popular for musical installations. Cappelen et al. suggest that interactive installations can transform an instrument from the tool of professional musicians to a means of *musicking*. This refers to "a meaning making activity that includes everyday listening, dancing, creating and performing music [5:511]." The use of

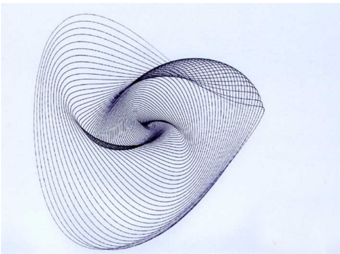

Figure 1: Drawing generated using a Harmonograph [7]

Figure 2: Example of the cut-up technique

digital instruments in public settings enables individuals to participate in the creation of music, regardless of their musical capacity [5].

Artificial Intelligence
There is considerable debate in the literature about the definition of AI [14]. For the purposes of this paper, I use the definition proposed by the AI100 Standing Committee and Study Panel: "a science and a set of computational technologies that are inspired by—but typically operate quite differently from—the ways people use their nervous systems and bodies to sense, learn, reason, and take action [19:4]." The development and application of deep neural networks is one of the most significant recent advances in AI technologies [10]. Deep neural networks are a form of machine learning modelled after the neural pathways in the human brain. They are unique because they allow for multiple levels of abstraction [10]. In other words, a deep neural network analyzes data in layers, looks for particular features in each layer, and builds on its findings in subsequent layers. For example, when analyzing an image, the first layer of abstraction may recognize a group of pixels going from dark to light, the next layer may recognize that the group of pixels is part of a line, and a future layer may recognize the line as the side of a house. The deep neural network can only recognize these patterns if it has been trained on similar data (ie: other images of houses). Significant advances in photo image recognition and speech recognition, among others, are credited for this cutting-edge technique.

Generative Music
Artists and researchers have long exploited novel technologies for creating artistic works [4]. Generative processes in the practice of art pre-date the advent of digital computers – for example, the use of harmonographs for drawing [7] and the "cut-up" technique for writing poetry [6] (figures 1&2). The development of computer generated music dates to the mid-1950's, when Brooks et al. used statistical analysis to create scores that mimicked an existing corpus [9]. The term *Generative Music* was popularized by Brian Eno with the release of a floppy-disk based generative album in 1996 [11]. Technological advancements, particularly in AI, have led to rapid growth in the practice of generative music and art [7]. Today's tools are more powerful and easier to use than ever before [1,9]. *Interactive* generative music is an increasingly important area of study for many researchers. Biles, for example, created an evolutionary computational model to jam with live musicians [2]. Nuanáin developed a tabletop system that allowed algorithmic composition using tangible blocks[11]. And Whalley used human gestures as input to a genetic algorithm mapped to sound [16].

Design Considerations

I propose a new system that lets individuals make music with an AI deep learning system. Based on a review of the literature, I develop three design guidelines on which to base a prototype:

1. *The system must be intuitive and accessible to people of all skill levels*. All users should be comfortable interacting with the system [5]. Blain et al.'s work on collaborative musical interfaces suggests that call and response-style interaction is the most effective for engaging both novice and professional players in an interactive music installation [3].

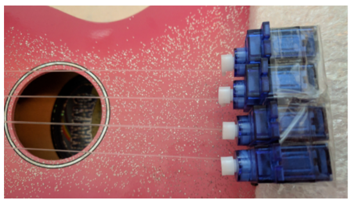

Figure 3: *Audio-Matic* prototype

2. *The user must be able to differentiate the sounds he/she plays from those generated by the system*. Previous work on interactive music installations suggests that users may have a difficult time recognizing their individual contributions to the overall sound [3]. This can be confusing or discouraging.

3. *The intended interaction must be playful.* To encourage interaction, this type of tangible system should be fun, engaging and pique curiosity [15].

Audio-Matic Prototype

Based on these design recommendations, I created the *Audio-Matic* prototype (figure 3). It enables a call and response interaction between users and Google's Magenta, a deep learning tool designed for music [1]. The system has three parts (figure 4): 1. a USB MIDI keyboard; 2. A computer running audio processing software (Max/MSP 6) and the deep learning tool (Magenta); and 3. A series of Arduino-powered servo motors mounted on a ukulele or other stringed instrument. The prototype is designed to be a proof-of-concept system of a larger work which could be placed in a public space, such as a gallery or museum.

To initiate interaction with the system, users play notes on the keyboard. Simultaneously, the audio processing software outputs the corresponding sounds from the computer speakers to emulate a real piano. The user's MIDI input is processed by Magenta, which analyzes the melody and compares it to a model that has been trained on thousands of songs in a MIDI database. Based on this analysis, Magenta generates a new melody that fits within the model. The computer's generated MIDI response triggers the servo motors, which play the corresponding notes on the stringed instrument. The use of an external stringed instrument helps the user differentiate his/her sounds from the generative component. Magenta's generative compositions can yield mixed results (musically speaking) but the system often produces interesting interpretations based on user's input. This unpredictability encourages curiosity and continued interaction with the system.

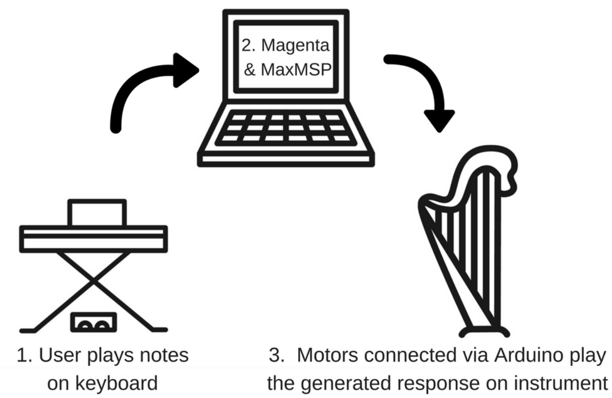

Figure 4. Audio-Matic interaction

Discussion and future work

AI researchers face a puzzle: How to create accessible entry points for systems that are, by definition, technical? The contributions of this research are twofold: 1. I propose draft design guidelines for an AI-based, tangible generative musical system; and, 2. I construct a working proof-of-concept prototype called

Audio-Matic. The purpose of this system is to provide such an entry point for users to interact with, and reflect on, complex AI systems like deep neural networks.

In the future, I will expand the project's scope in several ways. First, I will build a larger prototype that has at least a full octave range. This will offer a more accurate mapping of generated MIDI to instrument strings. Second, I will run a small mixed-methods study to evaluate the system from an interaction perspective. Using a "research through design" approach [17], I will investigate the relationship between user experience and the interactive generative music interface. I will then refine and improve the existing system. Lastly, in accordance with Hsu's recommendations for incorporating HCI principles into evaluating interactive music systems [8], future research may evaluate the generated compositions from subjective (ie: musical and artistic) viewpoints.

References

1. Martín Abadi, Ashish Agarwal, Paul Barham, et al. 2016. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv:1603.04467 [cs]*.

2. John A. Biles. 2007. Improvizing with Genetic Algorithms: GenJam. 137–169.

3. Tina Blaine and Tim Perkis. 2000. The Jam-O-Drum Interactive Music System: A Study in Interaction Design. *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, ACM, 165–173.

4. Margaret A. Boden and Ernest A. Edmonds. 2009. What is generative art? *Digital Creativity* 20, 1–2: 21–46.

5. Birgitta Cappelen and Anders-Petter Andersson. 2011. Expanding the Role of the Instrument. *Proceedings of the International Conference on New Interfaces for Musical Expression*.

6. Alan Dorin. 2013. Chance and Complexity: Stochastic and Generative Processes in Art and Creativity. *Proceedings of the Virtual Reality International Conference: Laval Virtual*, ACM, 19:1–19:8.

7. Alan Dorin, Jonathan McCabe, Jon McCormack, Gordon Monro, and Mitchell Whitelaw. 2012. A framework for understanding generative art. *Digital Creativity* 23, 3–4: 239–259.

8. Willliam Hsu and Marc Sosnick. 2009. Evaluating Interactive Music Systems: An HCI Approach. Hsu, William T., and Marc H. Sosnick. "Evaluating Interactive Music Systems: An HCI Approach." NIME. 2009.

9. Allen Huang and Raymond Wu. 2016. Deep Learning for Music. *arXiv:1606.04930 [cs]*.

10. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature* 521, 7553: 436–444.

11. Cárthach Ó Nuanáin and Liam O' Sullivan. 2014. Real-time Algorithmic Composition with a Tabletop Musical Interface: A First Prototype and Performance. *Proceedings of the 9th Audio Mostly: A*

Conference on Interaction With Sound, ACM, 9:1–9:7.

12. Cathy O'Neil. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype.

13. Christiane Paul. 2016. *A Companion to Digital Art*. John Wiley & Sons.

14. Stuart J. (Stuart Jonathan) Russell and Peter Norvig. 2010. *Artificial intelligence: a modern approach*. Prentice Hall, Upper Saddle River, N.J.

15. Orit Shaer and Eva Hornecker. 2010. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends® in Human–Computer Interaction* 3, 1–2: 4–137.

16. I Whalley. 2010. Generative improv . & interactive music project (giimp),. *Proceedings of the international conference on new interfaces for musical expression*, 255–258.

17. John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research Through Design As a Method for Interaction Design Research in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 493–502.

18. 2016. *Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems*. The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems.

19. 2016 Report | One Hundred Year Study on Artificial Intelligence (AI100). Retrieved March 25, 2017 from https://ai100.stanford.edu/2016-report.