Summary of the 2023 Senior Seminar

Personalized Music Generation with NeuroEvolution

Noriko Otani Laboratory
2072073 So Hirawata

1. Introduction

The rise of automation has been transformative in various domains. In particular, advances in automatic composition technology are providing composers with new insights and sources of inspiration. Otani assumed a scene where musicians create a song using an automatic composition system, and proposed a method for generating a melody based on an evolutionary computation algorithm for the system [1]. Its effectiveness has been validated by professional uses and subjective evaluations. However, it faces the challenge of reflecting the dynamic sensibilities of composers. Ando's Interactive Composition Aid System, CACIE, uses interactive evolutionary computation and tree topology for music composition [2]. This system facilitates the creation of complex or atonal pieces of music but the representation of large pieces in a tree model can be overly complex, potentially hindering user's understanding and manual intervention in the composition process.

This study aims to provide inspiration in the composition process and to enable the intuitive generation of long musical pieces that reflect the composer's dynamic sensibility. An interactive composition method using NeuroEvolution is proposed.

2. Music Generation Method

In the proposed method, a melody generation model, defined as a neural network, generates a melody following to the opening melody consisting of various notes and rests. A total of M such melody generation models are employed, and each model generates N unique melodies. The number of bars and beats per minute (BPM) of the melody are specified by human. The generated melodies are then rated on an 11-point scale, from 0 to 10, based on their musical quality by human.

Based on ratings, the parameters of the melody generation models are updated using Particle Swarm Optimization (PSO). To address the slow exploration of PSO, the position update formula is applied R times in a single updating cycle. This strategy is aimed at accelerating the adaptation of the models to the characteristics of highly rated melodies, thereby reducing the rating burden on human. Furthermore, when determining for the global and personal bests in PSO, the most recent model is selected from those with the best rate. This ensures that the melody generation models are continually refined to adapt to the dynamic sensibility and assessment criteria of human.

3. Experiment

An experiment was conducted using a web application that implemented the proposed method shown in Figure 1. RNN was adopted as the melody genertion model, and three models were prepared (M=3). The first one was trained on a diverse dataset comprising 45,129 pieces from the Lakh MIDI Dataset, covering various genres. The other two models were finetuned to specialize in jazz and classical music, respectively. Each model generated two pieces (N=2), and was updated by appling the position update formula 50 times (R=50).

Two experts participated to the experiment. The

participant A, with expertise in algorithmic composition, a background in classical and jazz music, and skills in choir, guitar, and saxophone. The participant B, a professional in pop and rock music, skilled in keyboard and drum performance. They were instructed to compose freely using the application and were asked for their experiences and insights through semi-structured interviews. Their words were analyzed using SCAT (Steps for Coding and Theorization).

Participant A found that the application often generated pieces that exceeded expectations, introducing an element of positive unpredictability. Participant B found the melody generation to be a departure from his usual creative process, highlighting the application's potential to inspire new directions in compositional activity.

The analysis of the number of notes and rests in the pieces generated in each cycle revealed a significant aspect of the melody generation model. The model not only adapted the composition of notes based on the highest-rated pieces but also showcased a sensitivity to the dynamic sensibilities of the participants. For instance, pieces with many sixteenth notes and rests were generated for the participant A, who assumed to play an instrument. On the other hand, pieces with mostly quarter and eighth notes were generated for participant B, who assumed to sing. This distinction underscores the model's ability to capture and reflect the dynamic properties of human sensibility.

4. Conclusion

The core of the proposed method in this study is a melody generation model based on neural networks, which is updated through human evaluations based on Particle Swarm Optimization (PSO). The experimental results show that the model's potential to generate melodies that resonate with the sensibility, stimulating unexpected creativity. Notably, the melody generation model showed an evolutionary response tailored to specific human characteristics.

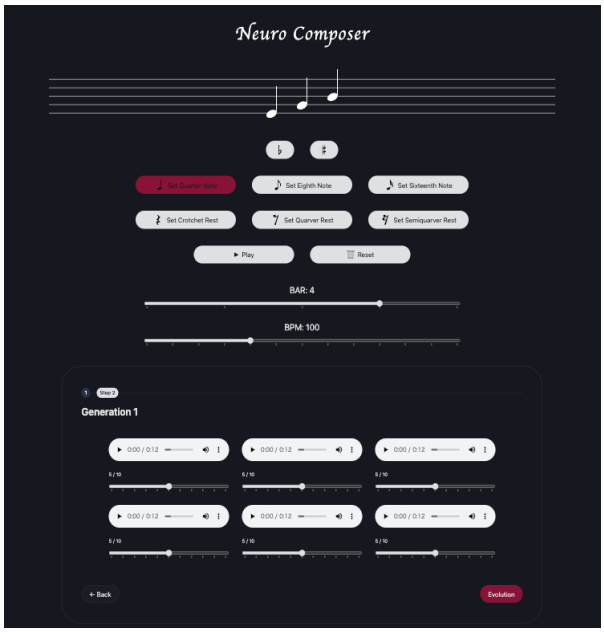

Figure 1. Application Image

In the experiment, participants pointed out the need of playing click sounds, generating chord progressions, and capturing the key to the music more effectively. Future work will focus on addressing these feedback in order to enrich the composition process, making it more expressive and user-friendly.

References

[1] N. Otani, D. Okabe, M. Numao, "Generating a Melody Based on Symbiotic Evolution for Musicians' Creative Activities," Proceedings of the Genetic and Evolutionary Computation Conference, 2018, pp.197–204

[2] D. Ando, H. Iba, "Interactive Composition Aid System by Means of Tree Representation of Musical Phrase," Proceedings of 2007 IEEE Congress on Evolutionary Computation, 2007, pp.4258-4265