JAMMING WITH YATING: INTERACTIVE DEMONSTRATION OF A MUSIC COMPOSITION AI

Wen-Yi Hsiao, Yin-Cheng Yeh, Yu-Siang Huang, Chung-Yang Wang, Jen-Yu Liu, Tsu-Kuang Hsieh, Hsiao-Tzu Hung, Jun-Yuan Wang, and Yi-Hsuan Yang
Taiwan AI Labs, Yating Team (yhyang@ailabs.tw)

EXTENDED ABSTRACT

Recent years have witnessed great progress in using deep learning models for automatic music generation. While many researchers choose to demonstrate their work by providing audio samples of the generated music only,[1] some researchers take the extra step to build demonstrations that are aimed to be interactive, accessible, and easy to use [1, 4, 5]. Along with the believe that a major goal of such AI models is to play music in real time with human musicians [2], we also expect to see in the coming years more interactive demons of music AI that include human to the loop of music creation, such as the one presented by AI Duet [3].

With this paper we like to present such a demo, as part of our endeavor in creating interactive music listening and creation experiences at the Taiwan AI Labs. The demo is made up with two AI functions: *jamming* and *auto-accompaniment*. For jamming, a human player plays a short phrase of music (via for example singing or piano) and the AI will response to it with an automatically generated music phrase. For auto-accompaniment, the AI generates accompany tracks such as bass and drums for the leading one provided by human.

We call our AI "Yating" (雅婷), naming after a common female name in Taiwan.

To showcase the two functions at once, we design a music creation process that is composed of two rounds. The process is illustrated in Figure 1. The first round is a jam session, a dialogue in piano between human and AI. Every time a human performer plays a 4-bar piano phrase,[2] Yating will continue and response to it with another 4-bar piano phrase. Meanwhile, Yating will generate the accompanying bass and drums tracks for this 8-bar phrase (i.e., 4+4) in the background, which are to be played in the second round. This interaction comes to an end after a few turns. In the second round, we mix the generated bass, drum and piano tracks together and loop them, so as to create a backing loop. The human performer can then start to record new tracks over the loop, for example via singing or playing the guitar, to create a multi-track live music.

Please visit `https://youtu.be/9ZIJrr6lmHg/` for a demo video. A snapshot of our musician jamming with Yating can also be found in Figure 2. As Yating is currently learning to compose Jazz music, the generated loop is a Jazz piano trio. We refer readers to [6] for some technical details.

We implement our demo based on a popular digital audio workstation (DAW) called REAPER,[3] which offers a set of APIs to create extended functionalities. With the APIs, we can write Python scripts to seamlessly connect REAPER with external MIDI keyboards or microphones, and the AI music composition models we built. Using DAWs as the interface allows musicians or programmers to customize many things related to music production in a professional way, such as the choice of virtual instruments and mixing parameters. This is not available if we use only common web applications to build the demo.

For better user experience, we implement two threads for parallel programming, one for the piano response part and the other for generating the bass and drums. While performing, the two threads also need to record some desired input (i.e., indicated by the black segments in Figure 1) and generate the results (i.e., the yellow

[1] Please visit `https://github.com/affige/genmusic_demo_list/` for a list of demo webpages we compiled.

[2] This is done via a MIDI keyboard in our current implementation, but can be replaced by using a microphone to record the sounds from a piano and then using a piano transcription model [6] to get the transcripts.

[3] `https://www.reaper.fm/` (accessed September 4, 2019)

segments). I/O latency (i.e., the red arrow) has also been accounted for. The demo is now ready to be played with in real time. In the future, we may further enrich this AI-human cooperative music creation process by adding more threads or other AI functions.

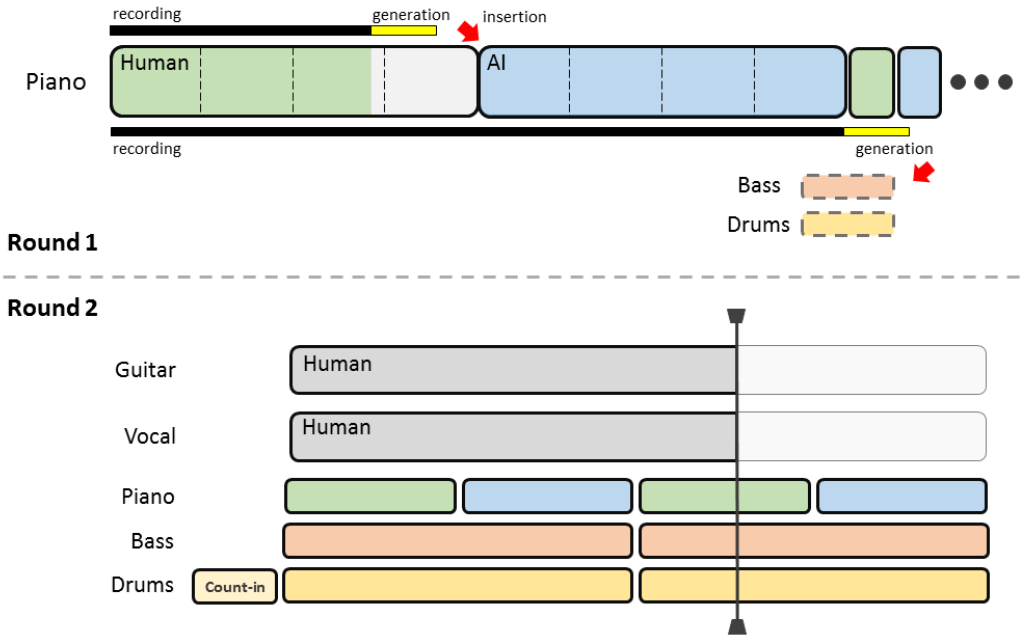

Figure 1. The flow diagram of the "Jamming with Yating" demo.

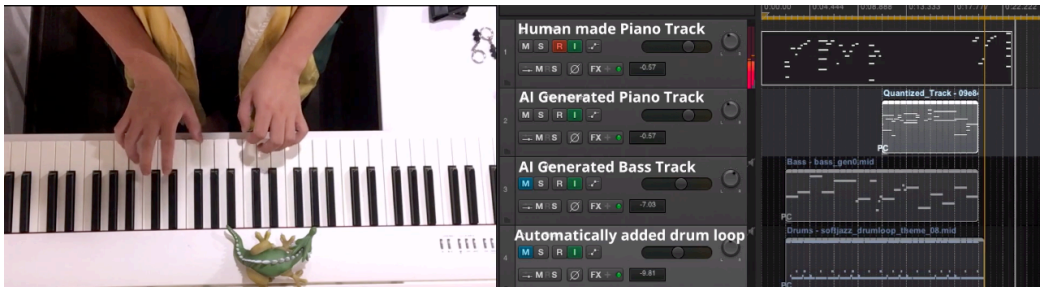

Figure 2. A musician jamming with Yating: (left) bird's eye view, (right) screenshot of REAPER.

REFERENCES

[1] Hamid Eghbal-zadeh et al. A GAN based drum pattern generation UI prototype. In *Proc. ISMIR, Demo paper*, 2018.

[2] Charles P. Martin et al. Deep predictive models in interactive music. *arXiv:1801.10492*.

[3] Adam Roberts et al. Interactive musical improvisation with Magenta. In *Proc. NIPS, Demo paper*, 2016. [Online] `https://experiments.withgoogle.com/ai-duet/`.

[4] Adam Roberts et al. A hierarchical latent vector model for learning long-term structure in music. In *Proc. ICML*, 2018. [Online] `https://magenta.tensorflow.org/music-vae/`.

[5] Vibert Thio et al. A minimal template for interactive web-based demonstrations of musical machine learning. In *Proc. Workshop on Intelligent Music Interfaces for Listening and Creation*, 2019.

[6] Yin-Cheng Yeh et al. Learning to generate Jazz and Pop piano music from audio via MIR techniques. In *Proc. ISMIR, Demo paper*, 2019. [Online] `https://soundcloud.com/yating_ai/sets/ismir-2019-submission/`.