Tutorial 2

Designing Generative Models for Interactive Co-creation
Anna Huang, Jon Gillick and Chris Donahue

Abstract

Recent advances in generative modeling have enabled AI systems to create realistic musical outputs, and additionally offer exciting potential to assist a broad range of music creators. However, these models have limitations that impose a unique set of challenges when embedded in interactive environments.

How do we design generative models and interactions that enable new creative (or co-creative) possibilities, while at the same time addressing real musical needs and user goals? This tutorial covers a range of considerations that come into play when using AI as a design material for creative musical interactions: identifying user needs and interaction opportunities, translating our high-level interactive objectives into actionable ML problems, (crudely) codifying desired behavior or aesthetics into quantitative metrics that can be hill-climbed during model development, and identifying cadences for evaluating progress with users in controlled experiments and in the wild. This process might span multiple projects or papers, with each diving deeper on different aspects of the process. We will draw from our own experiences and projects, highlighting choices that we made and reflecting on what we might do differently next time as we trace the lifecycle of these projects from research to the real world and back.

This tutorial will be geared towards anyone with some experience in MIR who is not already working at the intersection of music generative modeling and human interaction but may be interested in learning more. The primary purpose of this tutorial is to demystify this daunting process: we will offer guidelines and point out pitfalls, keeping in mind that there is no one-size-fits-all protocol. We hope that attendees will leave the tutorial with a clearer understanding of the challenges associated with designing, building, and evaluating interactive music AI experiences, and strategies which may help them overcome these obstacles.

Biographies of Presenters

Anna Huang is a Research Scientist at Google Brain, working on the Magenta project. Her research focuses on designing generative models and tools to make music more interactive and approachable. She is the creator of Music Transformer, and the ML model Coconet that powered Google's first AI Doodle, Bach Doodle. She holds a PhD from Harvard University, masters from the MIT Media Lab, and a dual bachelor's degree in computer science and music composition from University of Southern California. She is currently co-advising students at Mila, the Quebec AI Institute. She is also a guest editor for TISMIR's Special Issue on AI and Musical Creativity, and a judge and organizer for the international AI Song Contest.

Jon Gillick is a PhD Candidate at the School of Information at UC Berkeley, where he is also affiliated with the Center for New Music and Audio Technologies (CNMAT). His research centers around exploring new ways of creating and interacting with music and sound using machine learning. Before coming to Berkeley, he studied Computer Science at Wesleyan University in Connecticut and Music Composition and Production at the Pyramind Music Production Institute in San Francisco, and he spent time working in the Bay Area as both a freelance composer/audio engineer and as a software developer.

Chris Donahue is a postdoc at Stanford University in the computer science department. The primary goal of his research is to build AI systems which help humans be more creative. In practice, this often involves both improving generative models and designing new interactive environments which make these models more useful to humans. In a music context, he is particularly interested in how generative models may allow non-musicians to unlock their dormant musical creativity. Before Stanford, Chris completed his Ph.D. at UC San Diego, where he was co-advised by Miller Puckette (music) and Julian McAuley (CS).