Aaron Rourk
Computational Musicology
Fall 2018
Annotated Bibliography

Project Summary

This project seeks to contribute to the continued development of projects like Robert M. Keller's "Impro-Visor", the Experiments in Musical Intelligence of David Cope and Belinda Thom's "BoB". The goals of these efforts are to, through the analysis of pre-transcribed music, create a stylistic model of a particular artist, capable of predicting, and replicating, music in that style. I am interested in developing the capabilities of such systems for real-time, improvised scenarios. Obviously this involves some key differences in approach, but much of the underlying goal remains the same. Ideally, the system I envision would be able to transcribe monophonic, melodic improvisation and generate automatic, imitative improvisation in the style of the input. Rather than taking MIDI information as its data source, this system would analyze real-time audio input. Additionally, in a non-realtime system, multiple improvisations could be recorded, analyzed and categorized in advance of the generation of new, stylistically relevant music. In the situation I propose, there would be no time for such a labor-intensive process. Instead, the computer would have to assemble its own experience with every new iteration of the program. Some might protest that the computer already lacks memory, experience and culture, how could we deny it these things by design? But modern computer far surpasses human beings in terms of speed of analysis, sorting and execution, so on the whole it seems like a fair trade.

What I have described is, of course, a huge endeavor, and I will begin will just a small part. Rather than attempt to tackle the whole problem, I will focus my efforts on phrase-detection, specifically detection of the *ends* of phrases. For, when a phrase is over, the next sound will commence the next phrase. As such, there is no need for a system that can recognize a phrase *beginning*.

But when is a phrase completed? At first, it would seem that looking for periods of time beneath a certain amplitude threshold would guide one sufficiently. But what about when phrases are composed of several smaller melodic chunks that begin, then end, only to begin again? In fact, where are the lines between a "fragment," a "phrase," a "section" and an entire "piece." Are any of these actually clearly defined, or are they contextually dependent? Is a phrase a different sort of entity in contrasting musical styles? These are complex questions, but questions that I will attempt to address in the course of my research for this project.

Annotated Bibiography

Cope, David. "Recombinant Music: Using the Computer to Explore Musical Style." University of California, Santa Cruz, 1991.

In his research project of the title Experiments in Musical Intelligence, David Cope developed "an expert system that employs pattern recognition processes to create recombinant music—music written in the styles of various composers by means of a contextual recombination of elements in the music of those composers." Cope's idea of

using “disassembled, reorganized, and reassembled” portions of existing music serves as a crucial example for the project I have at hand.

Dubnov, Shlomo, Gerard Assayag, Olivier Lartillot, Gill Bejerano. “Using Machine-Learning Methods for Musical Style Modeling.” *Computer Magazine*, 2003, pp. 73-80.

In this article, the researchers describe their efforts in applying machine-learning methods for the use of modeling musical styles, work that has already had ramifications for the generation of music for computer-aided composition. Using machine learning, the researchers uncovered mathematical models from sets of musical examples, capturing the regularity apparent within. The work has implications for prediction and generation of works in imitation of existing musical styles.

Gillick, Jon, Kevin Tang and Robert M. Keller. “Machine Learning of Jazz Grammars.” *Computer Music Journal*, vol. 34, no. 3, 2010, pp. 56-66.

This article describes the automated learning of the probabilistic grammars that form the core of the automatic improvisation generator from Keller’s “Band-in-a-Box” software, extending it with “a technique for representing melodic *contour*.” The system the researchers have developed accepts a corpus of transcriptions and attempts to improvise solos in the style of that transcribed material by extracting representative contours and sequencing them using Markov chains. Using the basic building blocks of jazz improvisations, which the researchers refer to as “slopes,” the system combines these slopes into a grammar, which then follows the relevant tonal requirements of the jazz style.

Keller, Robert M. and David R. Morrison. “A Grammatical Approach to Automatic Improvisation.” *Proceedings SMC’07, 4th Sound and Music Computing Conference*, 2007.

In this article, the authors describe their approach for automatically generating melodies in the jazz style, in real time, through the use of probabilistic grammars, incorporated in a free software tool called Impro-Visor (short for “Improvisational Advisor”). The authors intend the tool to be used as an assistant for musicians learning to improvise in the jazz style. At first, the researchers created a database of melodic fragments which could be combined together to create melodic lines. This proved to be a labor-intensive approach, given the proliferation of keys and chord progressions present in jazz music. Instead, the researchers developed the approach described above.

Thom, Belinda. “Learning Models for Interactive Melodic Improvisation.” Carnegie Mellon University.

Building on the ideas of the Band-in-a-Box software (which utilizes Keller’s “Impro-Visor”), Belinda Thom “addresses the problem of the computer interacting with a live, improvising musician,” and introduces Bob, “a model of improvisation that enables the computer to trade solos with a musician in an adaptive, user-specific manner.” Most notably, the researchers have developed “an architecture that naturally enables

unsupervised learned knowledge, perception and generation to be tightly coupled.”