CS280r Final Project Report
Αρμονία (Harmonia): A System for Collaborative Music Composition

Mark Goldstein, David Wihl

{*markgoldstein,davidwihl*}*@g.harvard.edu*

Abstract

Increasing productivity of music composition has many positive benefits. Listeners would appreciate individually tailored music to their emotional needs and context. Composers would be facilitated by greater and more diverse cooperation yielding more innovative music. Composition agents could assist in the generation of repetitive or experimental musical forms. Therapists can use music as part of a treatment plan for autism and many other disorders. The system we propose attempts to address these myriad needs by offering two key innovations: a SharedPlan with collaborative versioning to mediate the workflow of a composition, an algorithmic evaluation of a composition against the intention of the SharedPlan to provide guidance to both human and agent composers.

1. Introduction

TODO Should contain an overview of the problem to be addressed, the approach taken to address that problem, and the results of that approach. Should provide the reader with a road map for how your argument will be developed in the other sections of the paper.

RESEARCH TODO

- Mark: talk to composers and incorporate their feedback and UI suggestions
- David: speak to David Greenberg to incorporate feedback and UI

2. Related Work

Our work builds on several areas of research related to music, computer science, and creative cognition. First we discuss related work in collaborative ideation, both in general and specifically in music. Next we discuss intelligent music systems that facilitate human composition and improvisation, and related work in Music Information Retrieval (MIR). Finally we describe previous work that applies information theory to the analysis of musical structure.

2.1. Collaborative Ideation

Collaborative Ideation (CI) seeks to improve the productivity of individuals and groups in generating ideas through collaboration. The people involved are interested in creating related objects (we all want to write birthday cards, come up with solutions to social problems, create music) and seek either feedback or examples of others' work to enhance their individual process. Collaboration is usually centered around a shared workspace that allows for communication and sharing of ideas. The space may be physical or virtual. The produced ideas may be for individual use, or there can be a shared artifact such as an essay or a piece of art attributed to several ideators. The dynamics of sharing ideas through the shared place may be real-time or not, though increasingly, today's settings are real-time and virtual.

The simplest forms of CI help each participant increase productivity in brainstorming solutions about a specified topic by showing all participants each other's ideas. While this approach has naturally been explored in human culture for millennia, the design of intelligent computer systems today aims to facilitate these activities to allow for increased creativity and productivity. Considering a setting where it is overwhelming for each ideator to view all participants' ideas, a computer system may prune and sort the idea space to show each participant only the most relevant and inspiring ideas. IdeaHound [Siangliulie 2016]. addresses at-scale collaboration in this setting. The system creates a semantic map of all ideators' generated ideas that allows each to easily view their work in the context of the entire solution space. The map is automatically generated by prompting users to interact with a personal "whiteboard" where they can cluster their ideas and separate them by semantic distance, and then by computing a global map from the collection of whiteboards. This approach bypasses the need for external workers to power semantic analysis of ideas. Using this map, IdeaHound can recommend diverse suggestions that allow each ideator to span a wider part of the solution space, without needing to devote cognitive effort to search large lists of ideas. Our system takes a lot of influence from this archetypal example. However, some questions that arise when moving to collaborative music composition specifically are: What happens when the generated objects are structured rather than unordered collections? What happens when ideators need to explicitly build on each other's ideas rather than only seek inspiration from them?

CI has surfaced in the space of online blogs and services designed to share visual art and music. Ideas range from small, unfinished efforts seeking directions to finished pieces seeking critique. Artists improve upon their ideas using the large-scale feedback. SoundCloud is an example of a hybrid music streaming service and CI platform. Though much of the hosted music is presented in finished form, people also post incomplete projects. Artists sometimes share "stems" to their music, which are individual sound files that feature isolated instrumental tracks, with the intention that others seeking inspiration remix their pieces into new ones. A newer platform called Blend makes the sharing of source files explicit. By default, artists share their works in progress in the format of music production software source files, which allows others to quickly pick up on the artist's work and take it in new directions. This setting is closer to our area of application and supports building on one another's ideas. However, what changes when several ideators intend to create a shared piece? With SoundCloud and Blend, one may take another's piece in a totally different direction. In our work, a collaborative composition has a goal associated with it through the duration of its existence. It is up to the composers and the system to stay close to this goal.

2.2. *Computer Facilitated Composition and Improvisation*

Computer agents with the ability to facilitate and take part in music composition and improvisation are on the spectrum of a wide number of intelligent musical systems under question in the field of Music Information Retrieval. Many of these approaches have in common a requirement for the system to "understand" music at multiple levels, including low-level acoustic signal, mid-level theoretical constructs such as harmony and rhythm, and high-level level aspects such as mood, genre, and style. Music recommendation system such as Spotify seek to analyze music and extract a measure of relevance for a function such as "study music". In systems that actually create music, people are interested in not replacing human composers, but assisting them. Perhaps the composer has good "seeds" ideas, but the system may recommend variations of ideas, or re-orderings of ideas, to make them more conveying. This system knowledge often comes from large-scale corpus analysis that mines for patterns and idioms common to the corpus.

ChordRipple [Huang 2016] is a recent system that takes as input a progression of musical chords from a composer, and suggests substitutions of intermediate chords that preserve the original semantics of the input while serving to replace conventional choices with more interesting ones. If the composer agrees to make one of the recommended changes, the system then assists the composer in interpolating between original and substituted material before and after the original change, resulting in further mixing of human and system generated music.

While our current work seeks primarily to assist teams of human composers to enrich and organize their work, we intend to design the system such that intelligent computer agent composers may be further in the loop. The Google Brian team has recently launched the Magenta Project for exploring machine intelligence in music. Magenta is an integrated environment of software tools and music-related datasets. Recently, Magenta released AI Duet, a computer system that reacts to human improvised gestures. Improvisation is an important part of composition. Even in steps where a human is composing, it may be beneficial to have an agent for the human to go back-and-forth on ideas with, much like two friends would iteratively vary and refine their ideas. In settings where a piece is defined by a specific enough set of guidelines, such as in a therapy use case where a listener may need music at a certain tempo and with a simple beat [See section INSERT SECTION] , powerful information retrieval systems make effective machine composition agents possible. Human composers may be placed at later steps of collaboration to ensure that the piece meets requirements in a humanly perceptible way.

2.3. *Information Theory and Music Analysis*

Our systems relies on the ability to model musical structure in a way that supports automated feedback for collaborating composers, where feedback is in the form of suggested rearrangements of musical ideas that help a composition reach a mutually specified structural goal. In this direction, there has been a rich body of work in automated analysis of musical structure from the 1950's to present. A prominent direction is to model musical form by way of listener perception and the expected dynamics of their attention and surprise. Understanding musical structure and its impact on listeners is a cogent goal to music theorists, cognitive scientists, and machine learning researchers. Many of these approaches have drawn on probability theory and information theory. A survey of approaches historical to contemporary can be found in the Con Espressione Manifesto [Widmer 2016], which is a strong position paper on the coming devade of research directions for music information reitraval

Our work builds on the the Information Dynamics Approach [Abdallah et al 2012]. Abdallah uses predictive information rate, a entropy/divergence based metric that measures how a listener's mid-piece distribution over future musical events is continually revised as new information is presented, to compute a curve that summarizes aspects of surprise in redundancy in a piece of music. Our work assumes that musical structure can be effectively summarized by this criteria. We assume that pieces from a particular genre/mood are defined by characteristic balances of surprise and redundancy over time, with peaks of information content (communicated by the composer) in genre-specific locations. We leverage this metric as the foundation of our automatic analysis system, which compares a collaborative work-so-far against the characteristic curves for the genre and mood specified by the mutual plan for the piece, and suggests edits to the composers to keep them in line with their goal.

3. System Design

[perhaps include a workflow image and a MIDI image side-by-side here]

3.1. *Workflow Overview*

someone creates a shared plan (individual or composers)

information retrieval system gets characteristic metrics using as much of the sharedplan info as it can

composers iteratively work on the piece

Within a particular composer/system interaction, automated analysis is run on the piece as it is. suggests some actions to the composer, particularly suggesting to switch two parts of a piece. Composer can follow suggestions or do their own modifications, or do nothing and finish.

3.2. GIT

git + intention, attributing credit, version control not available in existing midi systems

3.3. Genre + Mood

existing information retrieval systems

3.4. MIDI and actions to be taken on piece of music

3.5. Shared Plan + Metadata, inter-composer communication

git + intention + algorithmic eval
How do composers or users express intentions?
What sharing of information occurs between users and composers? Decoupling (be explicit)
failure mode: how to avoid revision wars

4. User Interface

TODO

5. Automated Analysis of Musical Structure

5.1. Brief information theory background

Entropy
KL Divergence
Mutual Information

5.2. Current Design

6. Use Cases

6.1. Individual User, Individual Composer

Our first use case considers the following scenario: a listener who may be a non-musician would like a new piece of music, perhaps to use for a function such as study music. We consider the case that the listener specifies a new project defined by a mood and genre. At this point, multiple composers could collaborate on the music specification, but we first consider the case that a single composer iterates over the piece with assistance from our system until the requester is happy.

6.2. Multiple Composers

Our second use case considers the case where multiple composers create a music specification together, and then collaboratively compose music that stays on track with the original specification.

TODO: include failure modes

6.3. *Therapist with Agent - Human Composition Team*

Our third case considers the situation where a music therapist would like music to use with their patients. These pieces may have a more highly-refined specification than music for casual listening. The specification may follow a treatment plan and may require a specific tempo or special therapeutic timbres (sound qualities).

high volume necessity

given the detailed specification, an artificially intelligent agent may do a large amount of work, which is then checked by a human composer

7. Discussion

7.1. *Enhancing or Stifling Creativity*

Notes: evaluation is optional. Can be ignored by committer.

7.2. *Limitations*

Collaboration is offline, not real-time

Current music representation is discrete MIDI, not audio. Limits for vocals, ocean sounds

Presume that reliable corpus-based genre and mood classification solutions exist, particularly information retrieval procedures

8. Conclusion

Two Novel Contributions:

- Collaborative music composition system Intentionality, SharedPlan and Agents
- Algorithmic evaluation of composition against intention

9. Future work

- Improved agent composition
- Intelligent ad hoc composition
- Facilitator of scalable music composition
- improved evaluator, possibly RNN based

10. References

Meyer, L.B., 1956. Emotion and Meaning in Music. Chicago University Press, Chicago, IL.

Narmour, E.. 1992. The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model.

D. Huron. 2006. Sweet Anticipation: Music and the Psychology of Expectation. MIT Press, Cambridge, MA.

Abdallah, Cognitive Music Modelling: An Information Dynamics Approach. 2012 3rd International Workshop on Cognitive Information Processing (CIP)

Widmer, Gerhard. 2016. Getting closer to the essence of music: The Con Espressione manifesto. ACM Transactions on Intelligent Systems and Technology

Engel et al., Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders, 2017

Wiggins, Auditory Expectation: The Information Dynamics of Music Perception and Cognition. 2012 Topics in Cognitive Science

Moles, A.. 1966. Information Theory and Aesthetic Perception. University of Illinois Press, Urbana, IL

Two Multivariate generalizations of Pointwise Mutual Information Tim Van de Cruys, Association for Computational Linguistics 2011

Cohen, Joel E., Information Theory and Music , Behavioral Science, 7:2 (1962:Apr.) p.137

Schillinger, Joseph The mathematical basis of the arts 1948

Pierce, Electronics, waves, and messages. 1956

Pierce, Letter Scientific American 1956

Youngblood, Style as information 1958 Journal of Music Theory

The mathematical theory of communication. Shannon, Claude Elwood 1948. Bell Tel Labs Monograph