UNIVERSITY OF CALIFORNIA
SANTA CRUZ

METHOD FOR SIMULATING CREATIVITY TO GENERATE SOUND COLLAGES FROM DOCUMENTS ON THE WEB

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF MUSICAL ARTS

in

MUSIC COMPOSITION

by

Evan X. Merz

December 2013

The Dissertation of Evan X. Merz is approved:

Professor Larry Polansky, chair

Professor Benjamin Carson

Peter Elsea

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Copyright © by

Evan X. Merz

2013

Table of Contents

List of Figures

Table of Illustrations

Abstract

Method for Simulating Creativity to Generate Sound Collages
from Documents on the Web

Evan X. Merz

To create algorithmic art with documents available on the internet, artists must discover strategies for organizing those documents. In this project I used a graph structure based on Melissa Schilling's model of cognitive insight to reorganize sounds on the web using aural and lexical relationships. I was then able to generate music with these graphs using several different activation strategies.

In section one I introduce my goals for this project. In section two I review other approaches to this problem and art that has influenced my approach. In section three I demonstrate techniques for organizing and collaging sounds from freesound.org. Sounds can be organized in a graph structure by exploiting aural similarity relationships provided by freesound.org, and lexical relationships provided by wordnik.com. Music can then be generated from these graphs in a variety of ways. In section four I show how my software was inspired by theories of creativity. Specifically I show how my software is an illustration of Melissa Schilling's graph model of cognitive insight. In section five, I elaborate on the pieces I've generated for this dissertation using this software and several other novel sound generating programs.

Acknowledgements

Thanks to all the people who helped me in graduate school, and in completion of my dissertation. Thanks to the many teachers who advised me throughout this project, including Ben Carson, David Cope, Peter Elsea, and Larry Polansky. Thanks also to the people who guided me throughout my graduate career, including James Phelps, David Evan Jones, Paul Nauert, and Amy Beal. Also thanks to Laura McShane for keeping me on track.

Thanks to the people at freesound.org and wordnik.com with whom I am unacquainted. Without the people at those sites, this work would not have been possible.

Thanks to my parents Don and Mary Ann Merz for their encouragement and support throughout a musical childhood. Thanks to my brother Isaac for working out many musical ideas with me. Thanks to my younger brothers Vincent, Angel, Mishe, and Julio for keeping me young. Thanks to my sisters Emily and Abbigail for putting up with me.

Finally, thanks to my wife Erin, for more things than I could possibly list here, and thanks to my daughter, Lucy, for making me happy.

1. Introduction

In this dissertation, I explore my compositional and political ideals in several pieces of software, as well as a six algorithmic, fixed-media compositions. The collection is titled *Disconnected* and each piece deals with a different aspect of my political beliefs. The group as a whole works toward two of my lifelong goals as a composer.

1. Write music using all sound.
2. Write new music.

In the next few chapters I will define these goals, along with my political beliefs, and show how they come together into this collection of music.

Section two is titled *Context*. It deals with the historical precedents of the music presented here. It starts by discussing the goal of composing with all sound, and the other musicians who have worked toward this goal. Then it discusses politically charged music, with a special emphasis on industrial music, which shares a handful of characteristics with the music in this collection.

Section three is called *Composing with All Sound*. That section is a step-by-step walkthrough of the algorithms at work in my software. It shows how graph structures are created from the relationships between sounds on the web. Then it demonstrates several techniques I've developed for activating those graphs.

In section four, *Connections with Creativity*, I describe the theories that inspired my software. I talk about Margaret Boden's theory of creativity, and how Melissa Schilling molded it into a graph model that is a practical tool for algorithmic artists. I also propose a data structure that can store web documents in a way that is related to the way ideas might be stored in our minds.

Section five, *Disconnected*, deals with the music itself. It reviews the compositional process for each piece, then shows how the resulting piece can be interpreted as relating to my compositional goals.

Sections six, seven, eight, and nine are the conclusion, appendices, supplemental files, and bibliography.

2. Context

The ideas incorporated into this project have their origins in early 20th-century art movements such as collage, Merz, Dada, cut-up, remix, mash-up, DJ set, and more. My doctoral work fits into several new categories of art and music that reinterpret Marcel Duchamp's readymades by using 21st-century technology. These categories include net art and database art, which represent new ways to recycle cultural artifacts.

In this project, the material being recycled is thematically inspired by the politics of the world in my adult life. This includes recordings of political speech, sounds related to politically-controversial subjects, and sounds drawn from copyrighted works. Although this work bears little sonic similarity to what is normally called protest music, I share compositional intent with some of the protest music of the past.

2.1. Collage, Found Sound, Sampling, Database Art

One of my compositional goals is to write music using all sound. In the software created for this project, I translate that goal into two guidelines.

1. No types of sounds are forbidden.

2. All types of sounds are possible.

These might at first look like only a single rule, but in practice these rules are different. The first rule relates to compositional intent. Often, a composer will narrow the focus of a composition as a first step in the process. Sometimes this means bracketing out certain types of sounds. Perhaps the composer might choose an instrumental ensemble, or a synthesizer for a piece, and in doing so eliminate other sources of sound. Or a composer of electroacoustic music might decide not to include speech sounds in a piece. All of these are ways of precluding sounds from occurring in a piece.

The second guideline relates to the practical limitations on the execution of a piece. If a composer only has a piano to work with, then it's impossible for him to incorporate the sound of a jet engine into a piece even if he wants to compose with all sound. Similarly, if a composer doesn't own a specific instrument, and only has access to a non-networked computer, then incorporating samples of a particular instrument is impossible.

In summary, the goal of composing with all sound translates into two ideas. First, no sounds should be eliminated in the pre-compositional phase. Second, it must be theoretically possibly for any type of sound to occur in a piece.

These were aesthetic goals for the algorithmic portion of this dissertation. The final musical compositions incorporate varying degrees of algorithmic composition and other compositional techniques. The software that executes these ideas is the main topic of this document.

2.2. Sampling in the Internet Age

In the 21st-century, composers have expanded the number of samples in their music by utilizing the vast amount of sound available on the world wide web. For some composers, using many samples is both a musical and a politial imperative. In 2008, Johannes Kreidler created *Product Placements*, a piece containing 70,200 copyrighted samples. His goal was to criticize the outdated structure of copyright law in the internet age.

> "With this performance, I've 'programmed' conflict by presenting GEMA with an enormous and completely absurd amount of administrative work in an effort to comply with their arcane registration requirements. Online registration is not possible for this composition, because the outdated copyright law forbids a composer from registering works that contain quotations online." (Kreidler 2009)

Lauren Redhead argues that Kreidler's approach also challenges the concept of creativity in general. By incorporating thousands of unidentifiable snippets of recorded music *Product Placements* "becomes about the nature of creativity, and does not accept the myth that the musical work and the composer are or should be inherently creative" (Redhead 2011, 5).

Other artists have created live web streams that constantly combine and remix the media posted to the web. Ben Baker-Smith's *Infinite Glitch* is a piece of software that combines videos from video sharing sites on the fly. It exists as a permanent web installation at infiniteglitch.com.

> "Every day an incomprehensible number of new digital media files are uploaded to hosting sites across the internet. Far too many for

> any one person to consume. Infinite Glitch is a stream-of-conciousness representation of this overwhelming flood of media, its fractured and degraded sounds and images reflecting how little we as an audience are able to retain from this daily barrage.
>
> Infinite Glitch is an automated system that generates an ever-changing audio/video stream from the constantly increasing mass of media files freely available on the web. Source audio and video files are ripped from a variety of popular media hosting sites, torn apart, and recombined using collage and glitch techniques to create an organic, chaotic flood of sensory input." (Baker-Smith 2012)

The works by Baker-Smith and Kreidler are extreme examples of the type of composition that is possible with access to the media on the internet. Still, in both cases, the samples are mangled to the point of unrecognizability. The samples in *Product Placements* are so brief that they cannot be identified. The videos used in *Infinite Glitch* are so heavily processed that they are difficult to comprehend. Kreidler is making a political statement, and Baker-Smith is commenting on the glut of media available to everyone. Neither artist is trying to combine the disparate media in a way that considers the content of the media files, and how it relates to its new context.

Since the late 1990s, mashup artists have been combining popular songs to form new compositions. These artists are relevant to this project because they recombine samples created by other people, based on the judgment that those samples work well together, which is precisely the process that is being automated by my software. The usual formula for a mashup is to combine the vocals from one recording with the instruments from another. These type of mashups are called A+B mashups, but mashups can get much more complex. One of the most famous A+B mashups is DJ Danger Mouse's 2004 album titled *The Grey Album*. *The Grey Album* is a notable achievement in part because it draws on only one album as source for the vocals, and one album as source for the instruments.

> "Danger Mouse, a DJ and producer, has married a cappella tracks from rapper Jay-Z's 2003 CD, 'The Black Album,' to musical beats and phrases found on the Beatles' 1968 offering, commonly known as 'The White Album' -- hence the title, 'The Grey Album.'"
> (Graham 2004)

Mashups are also significant because they transform listenership into a participatory experience. This is enabled by various message boards and websites such as gybo5.com (Get Your Bootleg On). Mashups are an extension of the growing participatory culture movement that has blossomed in the internet age.

The story of participatory culture often starts with the so-called *trekkies*. Initially scorned by the media and the general population as trogloditic outcasts (Jenkins 2006), the movement has grown uncontrollably since the internet has allowed distant fans to connect. Today, dressing up as your favorite fictional character, or writing fan-fiction based on characters created by another artist, is practically mainstream. Authors such as Pamela Aidan have built literary careers by writing new novels about Jane Austen's characters, and cosplay is one of the main draws to conventions such as the San Diego Comic Con.

Illustration 1: A fan's version of Marty McFly interacts with the Muppets at SDCC 2013

Jenkins expands on how fans interact with their favorite creations.

"fans are central to how culture operates. The concept of the active

> audience, so controversial two decades ago, is now taken for granted by everyone involved in and around the media industry. New technologies are enabling average consumers to archive, annotate, appropriate, and recirculate media content." (Jenkins 2006)

Websites such as freesound.org explicitly enable participatory culture. FreeSound allows users to upload, annotate and share audio files for music creation. Creating art with databases that are accessible or modifiable by users has become a small movement on the internet called *database art*. Although the label *database art* is primarily associated with the visual arts, my work shares one distinct characteristic with that movement. The software described here essentially acts as an agent for the artist, recontextualizing data from a database based on a process created by the artist.

> "The data readymade in database art has two important characteristics: a resistance to artist-made content and the transformation of the conventional role of the artist. Although the artist of this genre must contribute the core concept to the artwork, he or she creates a process more than a final product. It is this process that causes a flow of meaning between the old and the new contexts." (Hsu 2013, 81)

It's also important to note that other composers have written software that generates music using online audio files. Indeed, that was the initial goal of the creators of the FreeSound website. Before FreeSound, one of the creators worked on a project called *Public Sound Objects*, which provided a shared musical space for real-time collaboration.

> "The Public Sound Objects (PSOs) project consists of the development of a networked musical system, which is an experimental framework to implement and test new concepts for online music communication. The PSOs project approaches the idea of collaborative musical performances over the Internet by aiming to go beyond the concept of using computer networks as a channel to connect performing spaces. This is achieved by exploring the internet's shared nature in order to provide a public musical space where anonymous users can meet and be found performing in collective sonic art pieces." (Barbosa 2005, 233)

This project was expanded after FreeSound was established in a project called *FreeSound Radio*.

> "*FreeSound Radio* [is] an experimental environment that allows users to collectively explore the content in Freesound.org by listening to combinations of sounds represented using a graph data

> structure. Users can create new combinations from scratch or from existing ones. A continuous supply of potential combinations is provided by a genetic algorithm for the radio to play." (Romo et. al. 2009)

Recently, composer Miles Thorogood build the *Audio Metaphor* software to generate soundscapes using FreeSound and social media. Thorogood's program finds sounds on FreeSound using words in natural language queries. Then it tries to find additional queries using social media.

> "The words in the natural language query are processed to generate audio-file recommendations; however, these recommendations may be narrow in scope or number. To broaden the scope and number of results, Audio Metaphor generates a second set of database queries by searching Twitter for Tweets using the words in the natural language query." (Thorogood et al., 2012)

These earlier projects differ from this project in several ways. Like Thorogood's software, my software uses lexical relationships to discover sounds in the FreeSound database. Thorogood uses social media to find related sounds, whereas I draw on the word relationships provided by Wordnik. The largest distinguishing aspect of my software, however, is that it explicitly attempts to mimic the creative process as modeled by Melissa Schilling. So the discovered sounds are stored in a data structure that is based on Schilling's graph model of creativity. This will be explained in the fourth section of this document.

3. Composing with All Sound

3.1. Introduction

In this section, I present software that organizes sounds from the internet into a graph structure, and then activates that graph to generate electroacoustic music. This software incorporates both aural and lexical relationships to find related sounds. By drawing on two types of sound relationships, I connect this software with computational theories of creativity. This connection will be explored in the following section, while this section will explain the specific algorithm used by the software to generate music.

3.2. Constructing Sound Graphs

Search engines such as Google and Yahoo aren't very useful for creating music. While both Google and Yahoo allow users to search for sound files, neither search engine is optimized to search for audio. FreeSound.org provides a searchable database of user-contributed sounds that can function as a sound search engine. The FreeSound database contains over 66,000 sounds, as well as audio analysis information and user-submitted tags.

FreeSound.org provides access to their user-generated database of sounds through an application programming interface (API). My program makes requests to the FreeSound website, which returns the requested data. The FreeSound API provides multiple search criteria for retrieving sounds, including by text, content, and sound similarity. I use text search to find sounds relating to the initial search term as well as sounds with related tags, while the similarity search is used to build networks around those sounds. I don't use content search, although it may be explored in the future.

The FreeSound text search checks nearly all the text and numeric fields associated with a sound, including file name, file id, description and tags. According to the FreeSound API documentation, "searching for '1234' will find you files with id 1234, files that have 1234 in the description etc" (FreeSound API Documentation, accessed April 18, 2013). This leads to results where all interpretations of the search terms are included. For instance, when I searched for the term metal, the first two pages of results contained sounds made by hitting metal objects, while the third page contained a sample of a heavy-metal band playing a riff.

The FreeSound similarity search uses a distance metric that is calculated by an algorithm that combines features from several other analyses. The algorithm calculates a Manhattan distance, which differs from a standard straight-line Euclidean difference. The Manhattan distance is the sum of the absolute differences of the coordinates of two points. In high dimensional spaces, the Manhattan distance is slightly easier to calculate than the Euclidian distance.

> "the similarity measure used is a normalized Manhattan distance of audio features belonging to three different groups: a first group gathering spectral and temporal descriptors included in the MPEG-7 standard; a second one built on Bark Bands perceptual division of the acoustic spectrum, using the mean and variance of relative

> energies for each band; and, finally a third one, composed of Mel-Frequency Cepstral Coefficients and their corresponding variances." (Martinez et. al., 2009)

There are many alternatives to FreeSound's search algorithms. I could have used a database engine like Tom Stoll's CorpusDB, which is optimized for working with large databases of sounds (Stoll 2013). Or I could have created my own database of sounds, my own search functions, and implemented my own audio similarity algorithms. I chose to use FreeSound's search mechanisms primarily because they are convenient, but a second benefit is that by searching the entire FreeSound database, I don't have to limit my search space to a smaller database of local sounds.

These search-based mechanisms are used to build graphs of related sounds. A graph is a mathematical abstraction consisting of vertices and edges. The vertices are symbols, while the edges represent connections between those symbols. Graphs have been used to represent social networks, the spread of infection, and the connections between computers on the internet. The following diagram shows a graph with four vertices and three edges.

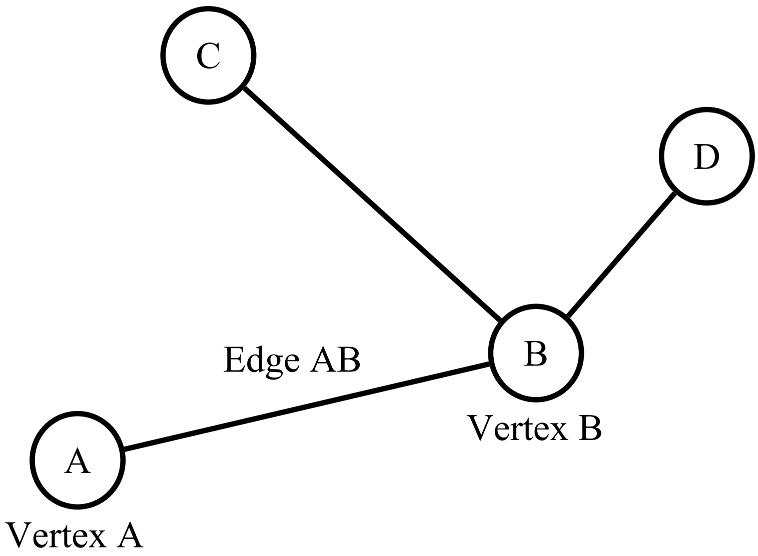


Illustration 2: A graph with four vertices and three edges

In this project, a new graph is created for each composition. The vertices in these graphs are sounds, and the edges represent an aural or lexical relationship between two sounds. Each graph contains the materials for a piece, and the connections that might be explored in building a piece. The graph is not a composition, however, it is merely a static data structure. The sounds in the graph must be activated in order to generate sound.

In each graph, the initial sound is obtained by using a text search on search terms provided by the composer. The first sound returned by that search is used as the first vertex in the resulting graph.

Next a basic similarity search is used to attach sounds with a similarity distance less than a pre-specified threshold. The similarity search provided by the FreeSound API never returns fewer than fifteen sounds, but it may be the case that the sounds returned aren't very similar to the original sound. The composer can specify a distance threshold that tells the program to ignore sounds that are not very similar to the sound under consideration and a maximum number of sounds to connect in this way. Some or all of the sounds returned in a similarity search may have a similarity distance of greater than the specified threshold. The FreeSound database contains over 66,000 sound files, so there are usually several sounds that are audibly similar to the starting sound. The program adds these sounds to the graph, and adds weighted edges connecting the new sounds to the target sound. Aural edges are weighted by their similarity rating. Sounds with a smaller distance rating are connected by edges with a larger weight.

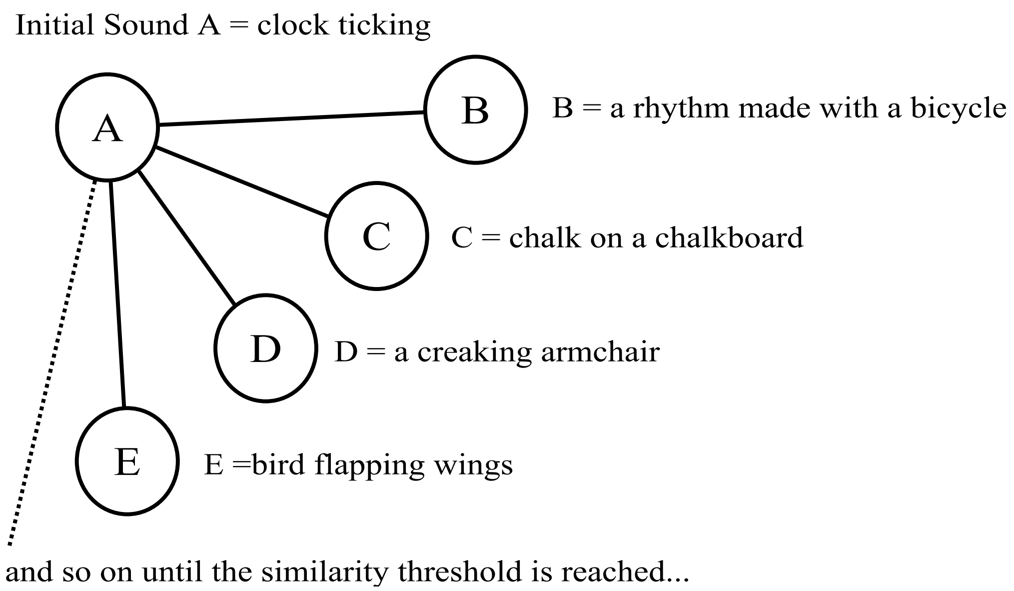


Illustration 3: Basic Similarity Search

After the basic similarity search, a recursive similarity search is employed to get a list of sounds, each being the most similar sound to the previous sound. In other words, a similarity search is executed on the most similar sound to the original sound. The resulting most similar sound is connected to that sound. Then this process is repeated up to a specified depth (n), creating a chain of sounds of length n. This allows a composition to move gradually from a specified sound to sounds that are less and less related to the original at each step.

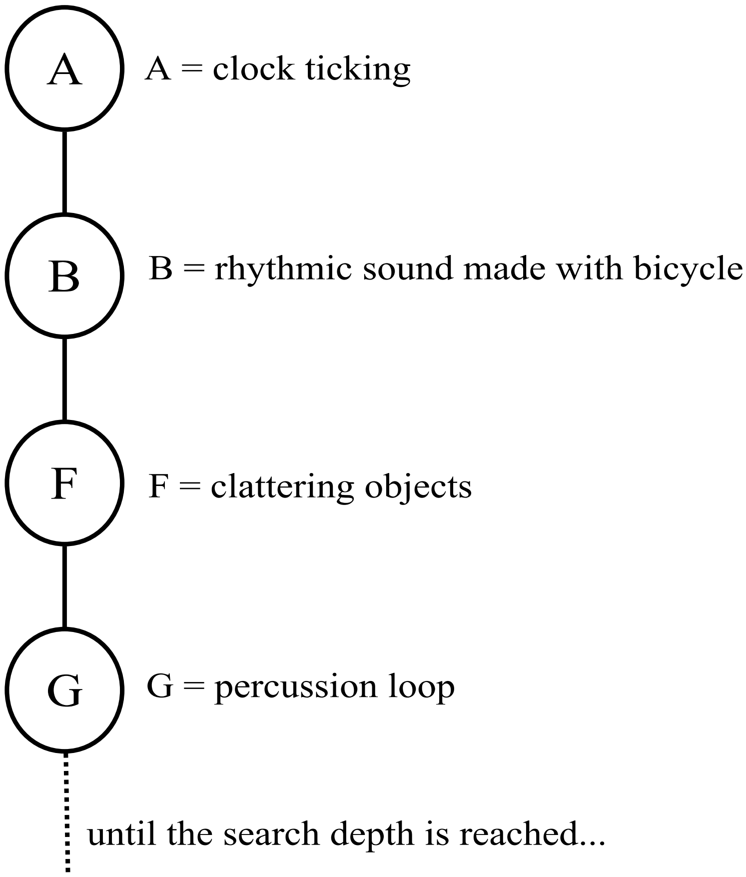


Illustration 4: Recursive Similarity Search

So far all of the sounds in the graph other than the initial vertex are discovered via FreeSound's similarity search. There is no way to connect one group of sounds to sounds that may be related but aren't audibly similar.

Wordnik.com provides an API that allows a programmer to find words that are related to a target word. The related words fall into categories that include synonyms, hypernyms, hyponyms, same-context, reverse dictionary, tags, and more. Definitions of these relationships and all possible relationships can be found at the *About Wordnik* page at wordnik.com/about. The following chart shows the first word returned under each category of words related to the search term *clock* on November 3, 2013.

Relationship	**First Result from Wordnik**
synonym	timepiece
same context	lamp
rhymes	Bach
tags	monosyllable
tagging	horology
reverse dictionary	regulator

Figure 1: Words related to clock

The tags on the original sound on freesound.org are used as search terms to queries of the Wordnik API. The related words returned by Wordnik are then used as new search terms in FreeSound queries. This provides a mechanism for the software to link aurally disparate groups of sounds.

Once these three searches are completed on the initial vertex, they can be repeated on any other vertex in the graph. This facilitates building arbitrarily large sound graphs that contain as many sounds or sound-areas as desired.

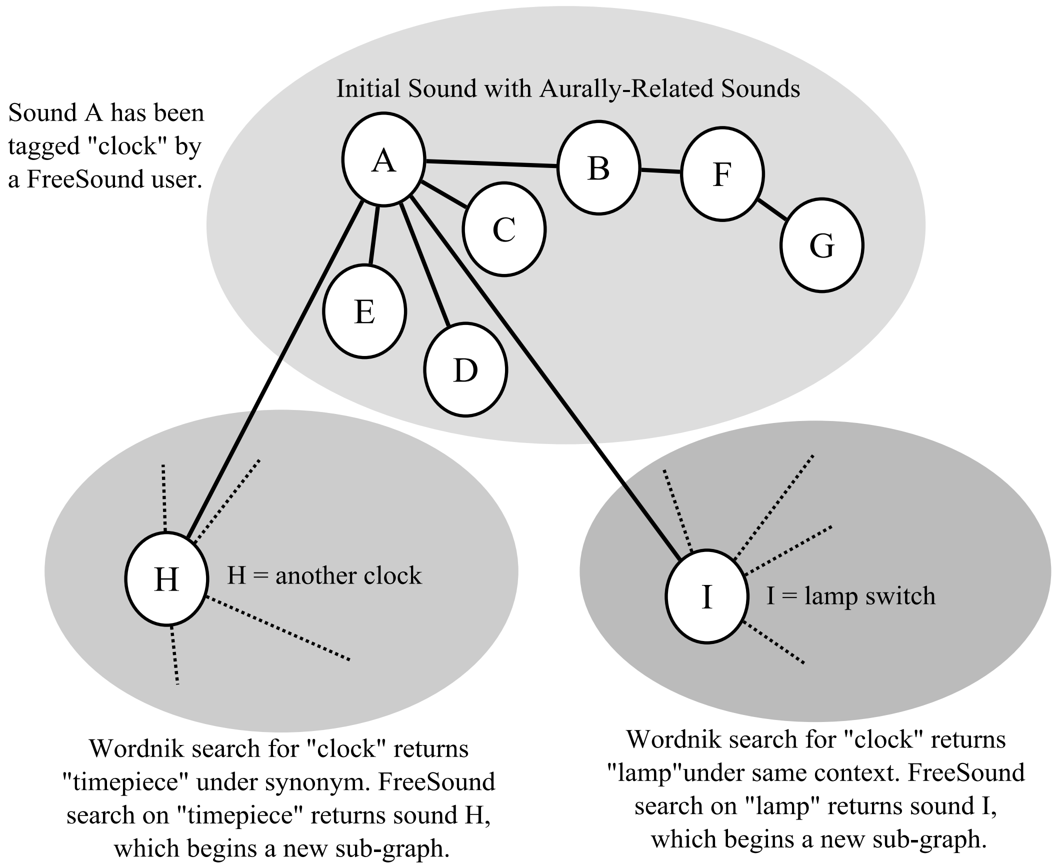


Illustration 5: Lexically-related sub-graphs

One caveat to this approach is that it assumes that the relationships generated by FreeSound and Wordnik are salient to a listener. The music generated in this project sonifies the underlying relationships in those databases. The resulting music will only be intelligible to listeners who can comprehend those relationships. These two data sources could be replaced by other sites, such as ccmixter or dictionary.com, or any databases that are judged to work better. An expansion of this project might compare audio similarity algorithms, and evaluate the quality of word relationships returned by Wordnik.

3.3. Creating Music by Activating Sound Graphs

After a group of sounds has been organized in a graph structure, those sounds must be accessed and triggered in some way. Activating these sounds is the second part of music creation using the system described here. The sounds can be activated in a variety of ways. A program might randomly wander from vertex to vertex, triggering sounds as it moves. Or a search strategy might be used to wander through the graph in a directed way. The goal of the activation step is to create a collage of sounds. In early versions, the software wrote collages directly to audio files. The current version writes to a format that can be subsequently manipulated in a digital audio workstation (DAW).

The activation, or collage-generation step writes to a Reaper Project file. The Reaper file format was created by Cockos for use in the Reaper DAW. Reaper files can contain all the track, media, midi and processing information that might be created in a music production project. In this case, the software creates many tracks in the output file, and organizes the audio collage onto those tracks. So the output file contains an editable Reaper project which gives me the opportunity to clean up or modify the algorithmic composition. Typically, several files in the collage will be unreadable, and several will have audio errors. I usually remove these files from the project and add compression on the entire mix before rendering it. In a few pieces where I wanted to explore a different type of collaboration between myself and the software, I edited the output more heavily.

I explored the collage-creation step with several different graph-activation algorithms. First I used an algorithm similar to a depth-first search which seeks to the end of the longest path in the graph. Specifically, I created a graph consisting of one long chain of sounds by using recursive similarity search to a depth of thirty. Sounds are then activated starting from the root vertex, the starting sound, and seeking to the leaf vertex, the final sound in the chain. The sounds along the path are activated in turn when 66% of the current sound has been played. This graph creation and activation strategy sonifies the similarity algorithm employed by freesound.org. In pieces generated this way, a listener can hear how the similarity algorithm recasts his own understanding of aural similarity.

My next attempt at sound-graph activation employed cellular automata. Cellular automata are deterministic programs where many sub-programs, or cells, act simultaneously. Every cell has a state and a rule set. Each cell may have a different state, but all cells have the same rule set. The rules determine the state of a cell at any time t, based on the state of that cell's neighbors at time t-1. It's difficult to code a standard cellular automaton that works on a sound graph because in the standard model of cellular automata, all cells are connected in the same way. In other words, they all have the same number of neighbors. In the sound graphs created by my software, this isn't the case. So creating a rule set that leads to the emergence of non-cyclical patterns is very difficult. In most of the rule sets I tried, the automaton either quickly moved to a steady state where all cells were off, or it quickly settled into an oscillating pattern where the same cells were activated repeatedly. Neither of these outcomes are desirable in music composition where variation over time is intended. As a result of these early experiments with cellular automata, I turned to a variation on cellular automata suggested by Peter

Elsea.

In this variation on cellular automata, a king vertex is selected that will act as the beginning of the activations. Activations radiate outward from the king. When 66% of the king sound has been heard, all vertices that are neighbors to the king are activated. When 66% of the longest sound in that group has been heard, all vertices that are neighbors to those sounds and haven't yet been activated, are activated. This continues until all sounds in the graph occur in the collage.

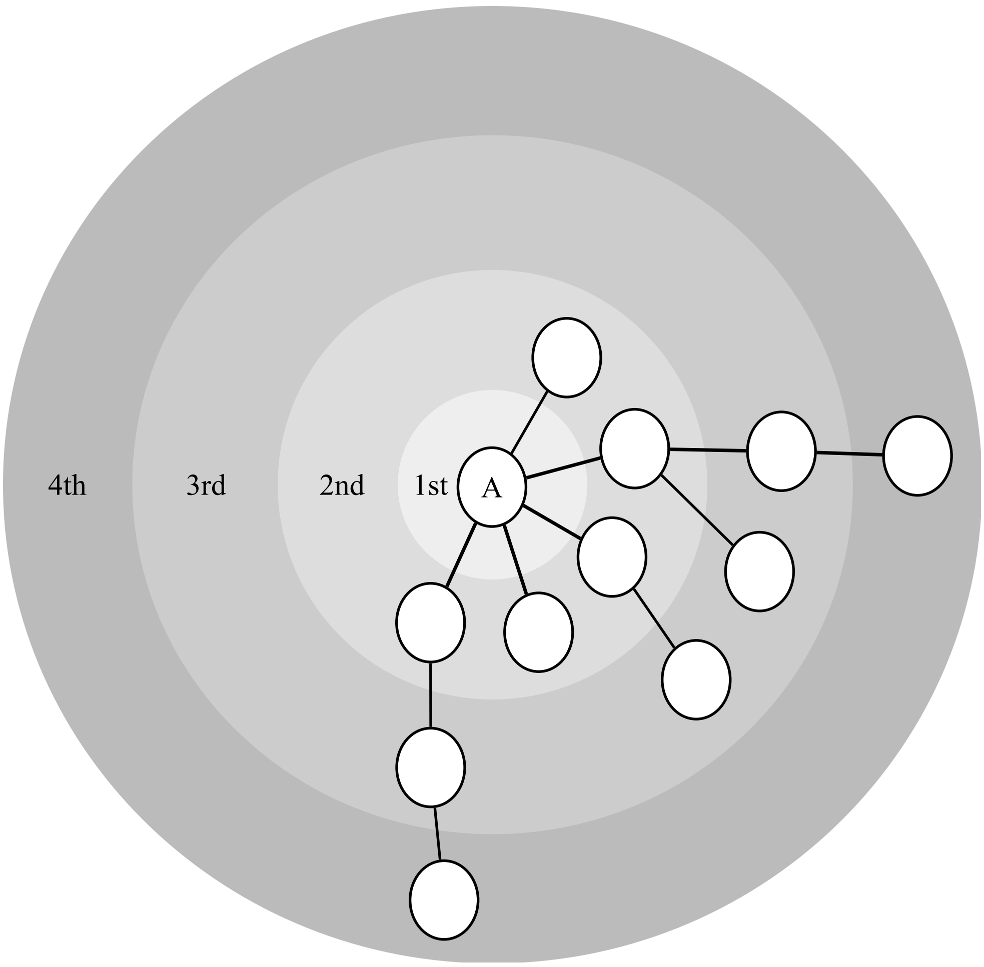


Illustration 6: King automaton activation strategy

The third graph activation strategy employs swarm intelligence. The program uses a variation on Craig Reynolds' *boids algorithm*. The boids algorithm has three rules: separation, alignment and cohesion. The separation rule says that boids shouldn't run into one another. The alignment rule says that a boid should move in roughly the same direction as the boids around it. The cohesion rule says that the boid should stay near the boids it can see. These rules must be slightly modified to work within the discrete space of a graph, as opposed to the continuous cartesian space of a typical boids simulation. The separation rule becomes an over-crowding rule. If too many boids are on a vertex, then that vertex becomes undesirable. The alignment rule is eliminated because in a graph of the type used by my software, direction doesn't exist. The cohesion rule stays the same. Boids try to move with the swarm as long as the space isn't too crowded. In short, the boids move as a loose swarm around the sound graph. They move around in roughly the same area of the graph without all being on exactly the same vertex. They move around the graph exploring the distinct neighborhoods that were created in the graph creation step.

3.4. Conclusion and Future Work

In this section, I described the organization of online sounds using aural similarity measures provided by freesound.org, and lexical measures provided by wordnik.com. After organizing the sounds in graphs based on these relationships, I showed three ways the sounds in the graphs can be activated to create music. These activation strategies are only three out of many ways that these sound graphs might be used to generate music. I am still working on ways to use this algorithm to generate vernacular style electronic music with a tonal center and a steady pulse. I am also working on ways to use the concepts here to generate visual art.

4. Connections with Creativity

4.1. Introduction

In this section I elaborate on the theories that motivated the graph structure described in the previous section. The structure of the graphs, and the specific steps taken to create them, is based on a graph model of creativity by Melissa Schilling. I will show how this model of creativity is implemented in my software.

4.2. Boden's Computational Model of Creativity

In her 1990 book *The Creative Mind: Myths and Mechanisms*, Margaret Boden suggests three mechanisms for creativity. Her model has been an effective starting point for artists, programmers and mathematicians (Cope 2005).

In Boden's model, a conceptual space defines the types of ideas that can be thought. Conceptual spaces typically arise from a given culture or education, as opposed to arising uniquely in each person. A conceptual space is defined by how a problem is parameterized, and how a particular parameterization may create or abolish conceptual barriers.

The first mechanism is called combinatorial creativity. It involves making unfamiliar connections between familiar, closely related ideas. An example is poetic imagery: creating metaphors and employing fanciful description. When Romeo asks if he should compare Juliet to a summer's day, he's pondering combinatorial creativity. Other examples might include creating a collage, or analogies. This type of creativity does not alter the mode of thinking or expand the boundaries of the conceptual space. It simply creates a new thought by combining thoughts that already exist.

The second mechanism proposed by Boden is called exploratory creativity. Exploratory creativity involves exploring the less well-defined regions of a conceptual space, expanding the boundaries of a conceptual space without fundamentally altering it. Boden argues that if the conceptual space is rich enough, then simply exploring it may lead to novel ideas. In the English language, exploratory creativity might occur when two words are combined to form a new one. This is occurring rapidly. For example, exploratory creativity occurred when someone first spoke of a motorcar, a handheld, or a chatroom.

Boden's third creative mechanism is called transformational creativity. Transformational creativity involves connecting two formerly unconnected conceptual spaces.

> "The deepest cases of creativity involve someone's thinking something which, with respect to the conceptual spaces in their minds, they couldn't have thought before. The supposedly impossible idea can come about only if the creator changes the pre-existing style in some way. It must be tweaked, or even radically transformed, so that thoughts are now possible which previously (within the untransformed space) were literally inconceivable."
> (Boden 2004)

Examples of transformational creativity in language are more difficult to conjure up. One example might be when words from other languages bring foreign concepts to english speakers. For instance, there isn't a native english equivalent for the concept of *ennui*. The incorporation of that word into the english language transforms the conceptual space, enabling new thoughts, new poetry, and new exploration.

4.3. Cope's Emily Howell

In the late 1990s and early 2000s, David Cope built software that applies Boden's model of creativity. The final product is a program that he calls Emily Howell, and the process of creating it is described in *Computer Models of Musical Creativity* (CMMC).

It's not possible to thoroughly explain Emily Howell's algorithm in this limited space. At its core, Emily Howell learns a complete, weighted graph of musical data from a database of musical examples. The vertices in the graph are measures of notated music, although it was programmed to allow data flexibility. Emily Howell's graphs are complete. In other words, all symbols are connected to all other symbols. The weights of these connections are set by the proximity between occurrences of the symbols in the learning database.

Music is generated from these graphs by employing a traversal strategy that mimics Boden's three mechanisms for creativity. When choosing the next symbol, Emily Howell considers both the edges from the current vertex, as well as the weights of the edges that connect the neighborhood of current vertex with the neighborhood of the vertex under consideration.

Cope defines three types of induction that correlate to the three tiers of creativity proposed by Boden. He calls them local, regional, and global creativity respectively. In local creativity, the next output symbol is chosen because it is strongly connected to the current symbol, and their neighborhoods are also strongly connected. In regional creativity, the same strong relationship between two symbols enables the program to select a series of symbols from the corresponding neighborhood. Finally, in global creativity, a symbol is chosen based on weak connections. In direct analogy to transformational creativity, a symbol is chosen that was hitherto only very weakly connected to the current symbol. This is a very brief overview that leaves out many important and interesting aspects of the Emily Howell program.

4.4. Schilling's Model of Cognitive Insight

Although Cope's model is capable of generating very compelling music, Cope refrains from connecting his program with a generalizable mathematical model. Other researchers have explored ways that creativity might be modeled using concepts from the field of graph theory. Melissa Schilling synthesized various approaches to creativity in the psychology literature and proposed a graph-theory model of cognitive insight that is similar to the algorithm represented in Cope's Emily Howell.

Schilling examined five mechanisms for cognitive insight proposed by various psychologists: completing a schema, reorganizing visual information, overcoming a mental block, finding a problem analog, and random recombination.

She drew graphs representing how these strategies work on specific problems, then showed that all of them are highly congruent (Schilling 2005). She showed that they can all be captured by the same underlying graph.

> "If association networks are sparse, and search is constrained or costly ... these results suggest that random or atypical connections can have a disproportionate payoff ... Applying small-world network principles to the realm of cognitive networks thus provides a compelling explanation for insight. The quantum leap of understanding that occurs during insight, and the affective response it produces, may be due to the formation of a new small-world network of interconected representations: One unlikely combination between two seemingly distant knowledge clusters suddenly results in a much shorter path length between a large web of connected representations." (Schilling 2005)

Ultimately, Schilling points out that all of the proposed creative mechanisms can be captured with the simple idea of the average path length decreasing in the graph under observation. Schilling argues that creativity occurs when a bridge is added to a conceptual network, thereby decreasing the average path length between two neighborhoods in the network.

Schilling does not speculate about the properties of an actual conceptual network in a real mind. She does not suggest specific values for the average path length or degree distribution of the conceptual network in the mind. Schilling's key claim is that conceptual networks in the brain are small-world networks enabled by bridges which are created in moments of cognitive insight.

It's important to note that Schilling's model deals with the idea of cognitive insight, which isn't necessarily the same thing as creativity in general. Still, scholars in both areas are dealing with the way that ideas are structured in the brain. The transformational creativity mechanism proposed by Boden and demonstrated by Cope seems to be functionally identical to Schilling's model of cognitive insight.

This model clearly presents many opportunities for future research in psychology and other fields. As an artist and a programmer, however, my goal in this project was to see if I could format data on the internet in a way that is analagous to the way Schilling suggests ideas are stored in the brain. If that is possible, then it opens the door to simulating creativity using data from the internet.

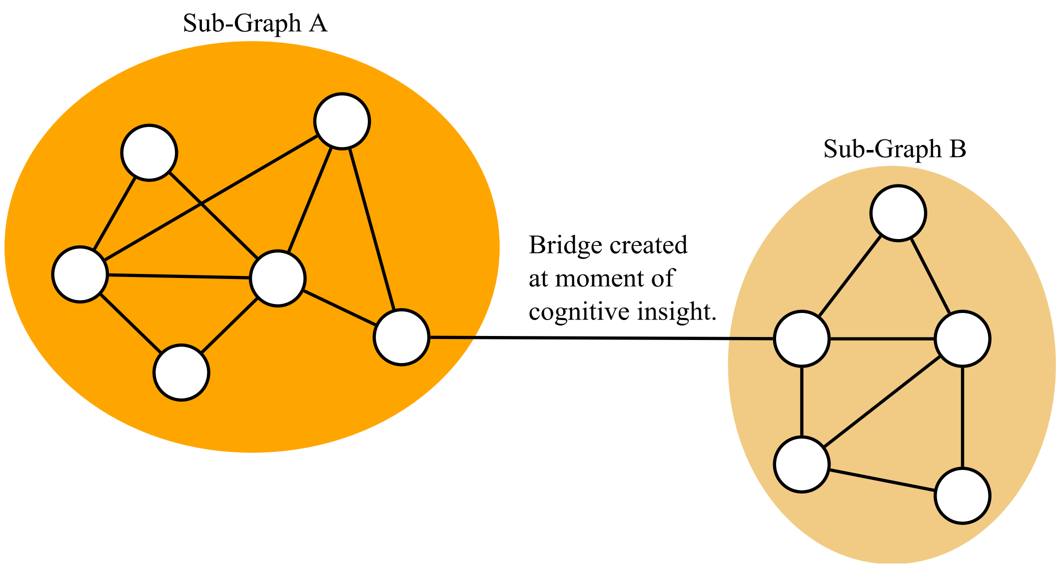


Illustration 7: Schilling model of cognitive insight

4.5. Barabasi's Study of the Internet

Due to the bridges in Schilling's model of the mind, ideas in the mind are modeled as small-world graphs. Small world graphs are graphs where almost every vertex is reachable by almost every other vertex in a very small number of steps. Specifically, the average path length of a small-world graph is correlated with the logarithm of the number of vertices in the graph. The small-world property of the mind is also found in the graph of documents on the web. Researchers have consistently reported a logarithmic relationship between the number of nodes and the average path length.

> "Despite the large number of nodes, the WWW displays the small world property. This was first reported by Albert, Jeong and Barabasi (1999), who found that the average path length for a

> sample of 325,729 nodes was 11.2 and predicted, using finite size scaling, that for the full WWW of 800 million nodes that would be around 19. Subsequent measurements of Broder et al. (2000) found that the average path length between nodes in a 200 million sample of the WWW is 16, in agreement with the finite size prediction for a sample of this size. Finally, the domain level network displays an average path length of 3.1 (Adamic 1999)." (Albert et. al. 2002)

Based on an understanding of other graphs, one might assume that the web relies on bridges to achieve this lower path length, but subsequent research revealed that this isn't the case (Barabasi 2002). In fact, the web relies on hub websites in order to achieve shorter path lengths. Hubs are super-websites that form the backbone of path-finding through the web. Hub sites are linked to by far more sites than the average website. These massively linked hubs provide short routes from one document to another.

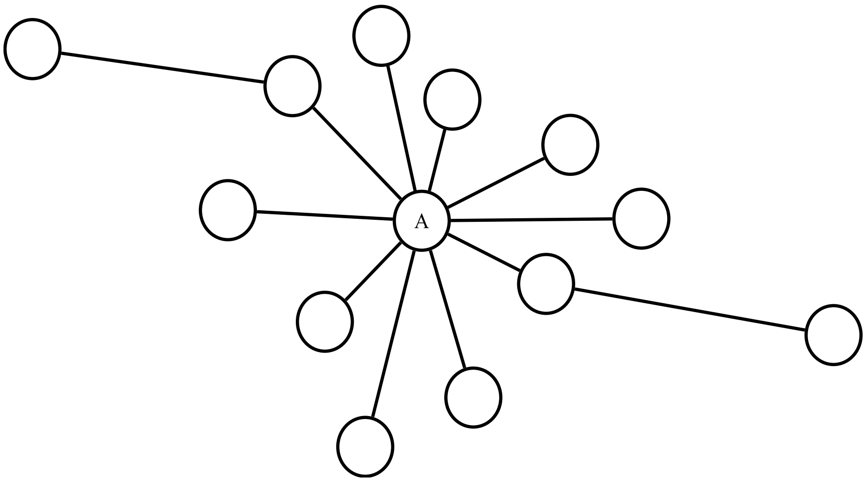


Illustration 8: A hub node

That both the internet and the mind seem to share the small-world property is not very revealing. Not only do they seem to achieve this through different mechanisms, but the small-world property is misleading, especially when trying to find something in a small-world graph.

> "The six/nineteen degrees phrase is deeply misleading because it suggests that things are easy to find in a small world. This could not be further from the truth! Not only is the desired person or document six/nineteen links away, but so are all people or documents. In other words, six —– or ten or nineteen —– can either be a very small number or a very large one, depending on what you're trying to do." (Barabasi 2002)

Although the graph of documents on the internet shares some mathematical properties with the graph theory model of creativity proposed by Schilling, the two are fundamentally different.

4.6. The Properties of Creativity Networks

In this section and the next, I propose a data structure that can work as a theoretical starting point for dynamic-content algorithmic art. This data structure is a broad class of structures which I call creativity networks based on the previous work by Boden, Cope, and Schilling.

Any graph that can be used to simulate creativity must have:

1. at least one vertex with degree greater than one (combinatorial creativity)
2. at least one path with length greater than one (exploratory creativity)
3. at least one bridge (transformational creativity)

These requirements are deliberately weak. I am not suggesting a data structure that can capture the creative potential of the mind, rather I am suggesting three characteristics that we must look for in any algorithm that attempts to mimic creativity.

Cope's graphs in Emily Howell are complete, every vertex is connected to every other vertex, hence they cannot fulfill the conditions above. There can be no path greater than length one in a complete graph, and there cannot be a bridge in a complete network. However, if we consider a graph containing only the vertices which are actually used in a composition by Emily Howell, then Cope's graphs do conform to the above three rules. Cope's traversal strategy, which mimics Boden's creative mechanisms, has the effect of transforming the graphs into creativity networks. In this sense we can see Cope's Emily Howell as one way of employing a creativity network.

Creativity networks were designed to capture Boden's creative mechanisms, but it's important to note that they also capture the spirit of other models of creativity. Consider Edward De Bono's distinction between vertical logic and lateral thinking (De Bono 1970). In vertical logic, a solution to a problem is found using a step-by-step procedure, while in lateral thinking a shorter solution is found by using an alternate approach to the problem. DeBono's modes of thought relate to exploratory and transformational creativity respectively, and they are represented by path length and bridges in graph theory.

4.7. Meaning from Web Search

There exists one final obstacle in constructing a theoretically-grounded program that simulates human-like creativity according to Schilling's model. When concepts are linked in the mind, they are not linked haphazardly or randomly (Schilling 2005). Rather, links are created through life experience, and only maintained if they prove to be useful.

Links in web pages can be placed anywhere, but web search provides a different way to connect two documents on the web. When a user searches for something in a search engine, they create a small network that connects their search terms to the returned documents. So, when I search for "Downton" using

Google, I connect the search term to the Downton Abbey wikipedia page, the imdb page, the PBS page and so on.

The context of the search engine gives objective meaning to the resulting graph. For example, if you searched for "Downton" before the show was created, then Google might respond with "Are you sure you didn't mean Downtown?" and display results relating to your local downtown area. The two different states of Google's graph of the web led to two different meanings for the word Downton.

By using web search rather than html links, we alter the graph significantly. That's part of the goal of my software —– to alter the graph of the web in order to represent web documents in a way that is related to how ideas are organized in the mind.

In this project, I am assuming that this definition of meaning for a document on the web also applies to sounds, both within a graph structure, and within a piece of music. In both cases, the relationships between the sounds gives meaning to those sounds

This is a key assumption of electroacoustic music, and much of the discourse surrounding the field. In *Music and Discourse*, Nattiez points to the context-dependent nature of *objet sonore* as a problem with using them to understand a piece of music. He argues that "perception of a sound-object within a work... never operates independently of what precedes and follows that element" (Nattiez 1990, 94). He finally points to Steve Reich's minimalist works as a concrete example of how the context of a sound changes the interpretation of that sound.

> "In Reich's minimalist works, the same pattern is heard to the point of saturation, then a tiny variation is introduced. We need a little time before realizing that we are in a new sequence, because the pregnant quality of the previous sequence acts upon our perception of the following one. In other words, we need simultaneously to grasp the characteristics of a single object, and to experience the link established in the *succession* and *enchaining* of object-sequences." (Nattiez 1990, 101)

Joanna Demers argues that recontextualizing sounds is one of the principle motivations for artists who create sample-based music. She writes that all electronic music is "a meditation on the act of listening to sounds both old and new, therefore a meditation on the cognitive processes that accompany listening" (Demers 2010). Demers argues that some composers use samples to incorporate meanings from a sample's previous context into its new context. She writes that "sampled materials are 'read' as texts whose original meanings and associations contribute to the meaning of the new composite work" (Demers 2010).

In *Capturing Sound*, Mark Katz shows how this recontextualizing of samples to change their meaning is used by hip hop artists. He argues that by sampling their heroes, Public Enemy puts their work into context, and by recontextualizing the words of civil rights leaders, they criticize other pop artists such as Bobby McFerrin (Katz 2010, 164).

One of the great powers granted by a social media-sharing site like FreeSound is the ability to quickly yield a new perspective on a sound by placing it in a new context. This power has already been recognized by the founders of the site.

> "One important aspect of creative use of sounds is the possibility to cross established boundaries and use recordings of dierent types in unintended ways. Thus, in current popular music it is an established practice to use recordings from very diverse domains: voice recordings, environmental sounds or fragments of existing music are commonly reinterpreted by putting them in new auditory contexts. The Freesound.org database is actually a good example of this tendency towards eclectic use of sounds. The site contains as of january 2009 more than 60000 sounds of the most diverse nature, with the only common denominator that they are expected to be reused in some new context." (Roma et. al. 2009)

4.8. Constructing Creativity Networks Using FreeSound

As discussed in the previous section, my software employs three search mechanisms that simulate the three types of creativity. Basic similarity search relates to Boden's combinatorial creativity. Recursive similarity search relates to Boden's exploratory creativity. Lexical search using the Wordnik API relates to transformational creativity.

The final search mechanism is the key. The underlying graph of aurally-related sounds already exists within the FreeSound database. This can be compared to a person's brain before a moment of cognitive insight. By using lexical search to connect disparate parts of that network, we can simulate Schilling's model of creativity. In this type of simulation, it's necessary to relate vertices in multiple ways.

4.9. Relationship to Emily Howell

My software owes part of its methodology to Cope's Emily Howell, however, my software takes a different approach to core concepts shared by both programs.

Both my software and Emily Howell create music by recombining material from a database. Cope's program learns a graph by examining musical fragments within a musical context. In my software there is no learning step, nor previous musical context. The sounds from FreeSound are aurally related in the FreeSound database, but there is no musical context from which my program can learn musical relationships. My methodology assumes that a listener can hear the aural and lexical relationships extracted from FreeSound and Wordnik. The choice of data sources, however, is central to the artistic success of the final product, and other data sources may exist which may be more fruitful.

Another similarity between both programs is that they both create graphs. The character of the resuling graphs, however, is distinctly different. Emily Howell creates complete, directed, weighted graphs, whereas my software creates a sparse, undirected, unweighted graph.

One distinct difference between the two programs is that Emily Howell employs only one graph traversal strategy. Cope's traversal strategy summons Boden's creativity mechanisms. In my software, the creativity mechanisms are embedded in the graph. The graph has structures that allow an arbitrary activation algorithm to take advantage of some or all of Boden's creativity mechanisms. Cope's program is tied to one traversal strategy, while my software can employ a number of possible activation strategies.

Cope's Emily Howell can be seen as one way of using creativity networks to create art. In this section of my dissertation, I've tried to show that there is a more general underlying algorithm that can be used to simulate creativity. My program is an instantiation of this algorithm, but by choosing different data sources or activation strategies, more variations on this methodology could be created.

4.10. Conclusion

In this section I explained the theoretical basis for the graph structure used in my software. I showed how, by storing meaningful relationships in the edges in a graph, any activation algorithm can be seen as a way of allowing meaning to emerge from that graph. I showed how my software is related to Emily Howell by David Cope, and how both programs are examples of a broad class of creativity simulations.

5. Disconnected

Disconnected is a collection of pieces composed using three categories of techniques.

1. Purely algorithmic using FreeSound composition software
2. Interactive algorithmic using a custom synthesis interface
3. A combination of the above two techniques

5.1. Mindphaser Edward Snowden

5.1.1. Composition

Edward Snowden entered public consciousness in the spring of 2013 when he revealed the extent of the NSA's national surveillance program. As an exile, Snowden is a victim whose crime seems to be telling the truth.

Compositionally, this piece combines and extends most of the techniques in the rest of this collection. It is a collaboration between myself and the software, and incorporates a custom synthesis interface, resynthesized sounds, and samples from commercial recordings.

The piece began by creating a sound graph using the search term *spy*. The resulting graph was difficult to predict. I predicted sounds of spy cameras, or spy gadgets, but the resulting graph is diverse and difficult to characterize. It contains several loud alarm and bell sounds which figure prominently in the output. However, the first sound returned from the search is titled *scaryscape_spyone*. This sound is 43 seconds of a high, piercing synth sound that alternates with sharp, distorted synth attacks. Although it is an interesting sound, the first tag on that sound is *funk*. This led the lexical search to look for sounds relating to jazz. Although the resulting graph doesn't contain jazz riffs, it does contain a lot of instrumental sounds at similar pitches. This includes sitars, and western orchestras, as well as unrecognizable synthesized drones. These sounds may be only distantly related to the initial search term *spy*, but in the resulting output, the internal inter-sound relationships that emerge are novel and compelling.

Graph Details for *Mindphaser Edward Snowden*	
Node Count	120
Edge Count	390
Degree Distribution	
Degree	**Occurrences**
1	78
2	11
3	13
4	2
5	1
9	1
12	2
13	1
14	2
15	1
16	2
17	2
18	1
19	1
23	2

Figure 2: Graph details for Mindphaser Edward Snowden

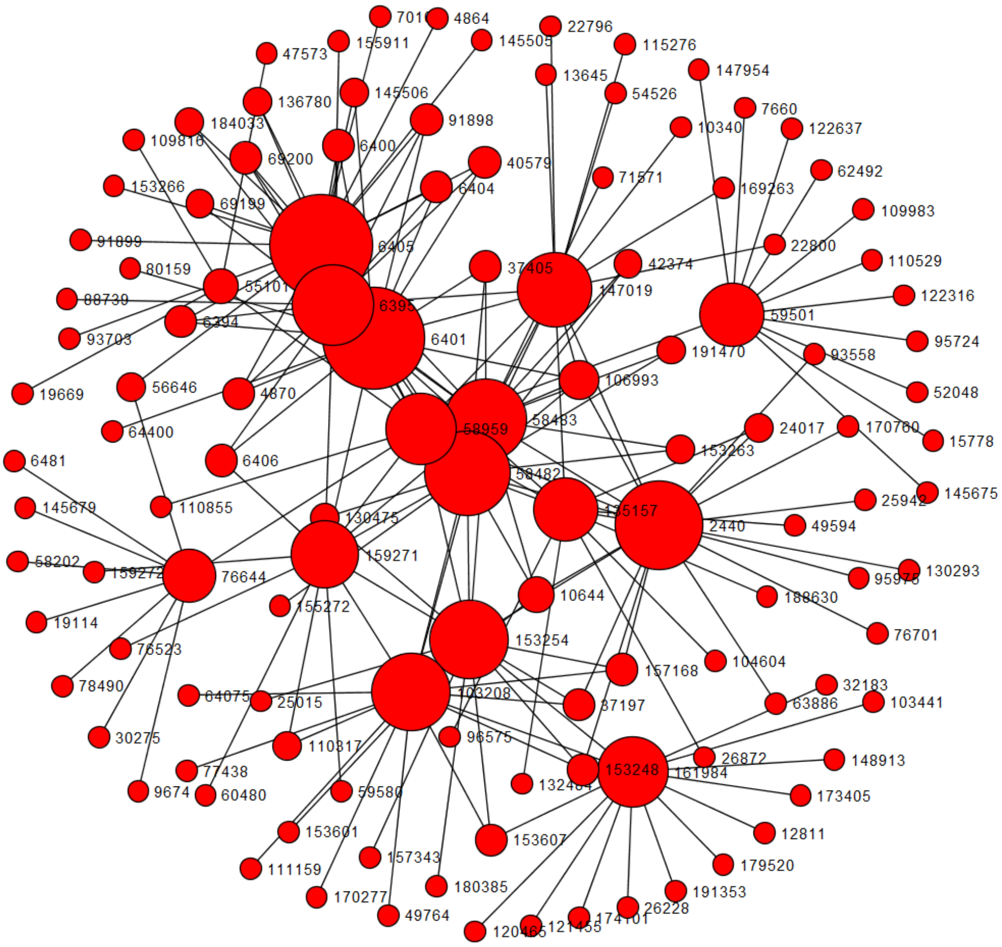


Illustration 9: Sound graph for Mindphaser Edward Snowden

I activated this graph by using the swarm intelligence traversal method to create a five-minute sound collage. Rather than using the complete raw sound collage as a piece itself, I listened to the piece to discover where the swarm created natural divisions in the soundscape. I located short silences, and places where the stream of sounds changes character. These were used to extract five different sections of material. I rendered both a reversed and a regular version of each section for flexibility when incorporating the music into a larger composition.

I wanted to incorporate vocal sounds and instrumental sounds with the edited sound collage. I had intended to use speech by politicians regarding the NSA spying program, but I thought that selected quotes from Edward Snowden were more to the point. I focused on two Snowden quotes which ultimately became the only spoken material in the piece. In one quote Snowden says "any analyst at any time can target anyone." In the other quote he says "the NSA specifically targets the communications of everyone. It ingests them by default. It collects them in its system and it filters them and it analyzes them and it measures them and it stores them."

These two quotes were modified in several ways. I stripped out almost all of the high frequency content above 1000Hz, then I time-stretched each sound significantly. These transformations resulted in quiet, smoothly changing sine tones that I used as a harmonic ground for parts of the piece.

In addition to extracting the harmonic character of his voice, I also wanted to extract the noisy character of his voice. To focus in on this aspect of the sound, I applied a custom, interactive sound granulator of my own design. Like many of my interfaces, this granulator uses the mouse as the primary input device.

Granular synthesis is a process whereby many tiny grains of sound are combined to form larger sound phenomena. It is a variation on additive synthesis, where the sounds being combined are very short in duration. In this interface, the left mouse button makes the grains shorter, while the right mouse button makes them longer. The position of the mouse along the y-axis is used to control the time between grains, on an exponential scale. The position of the mouse along the x-axis is used to control the position in the sound file from which grains are drawn. Finally, the loudness of the grains is determined by the speed of the mouse cursor. I created a granulated version of each quote from Snowden.

In this piece I wanted to manually use aural and lexical relationships to bring new sounds into the piece. To discover aurally related material, I turned to John Cage. His work with gamuts, and his open scores, often result in a chaotic mix of many different types of sounds that is similar to the collages generated by my software. I found a nice performance of *Variations III*, and cut out a twenty-five second sample.

Next I needed a sound that is lexically related to the underlying concept. I looked for music about spies or the cold war. I cut a snippet from a song called *Spies* by Coldplay. The introduction to *Spies* contains a long synthesized tone with a rising inflection at the end. I sliced out this nine-second sample to work it into the new composition.

I arranged and modified the materials according to my intuition. I kept the sections of algorithmic sound collage in order, although they appear in modified form. I interspersed these sections with quotes from Cage, Coldplay and Edward Snowden.

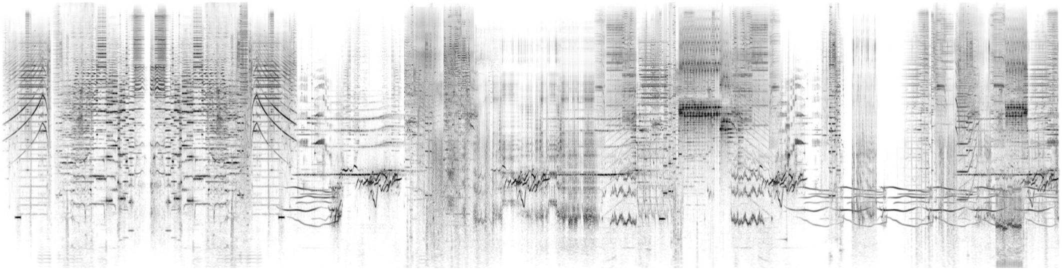


Illustration 10: Sonogram for Mindphaser Edward Snowden

5.1.2. Interpretation

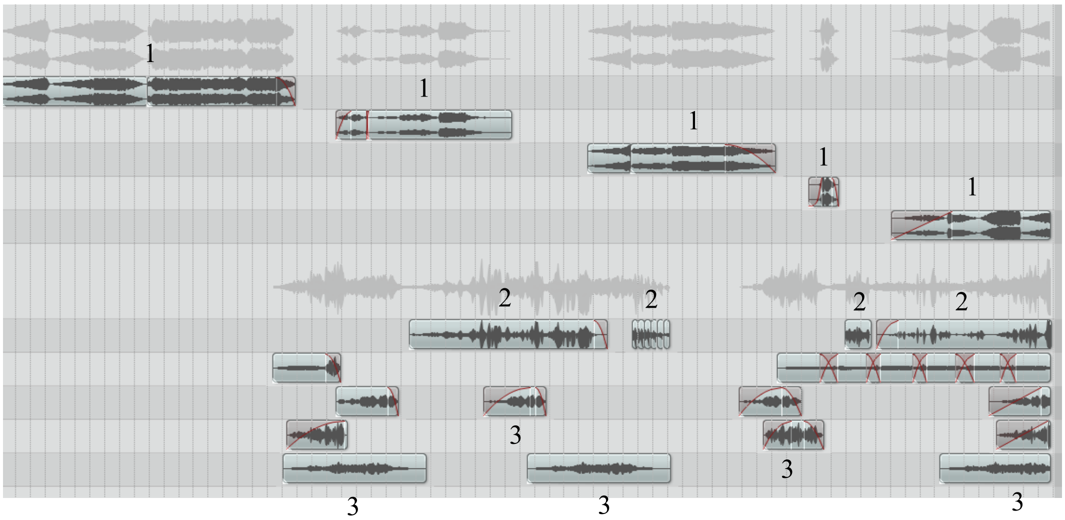
By choosing the title *Mindphaser Edward Snowden*, I put this piece into context both musically and politically. The title *Mindphaser* is a title with a pedigree that has been used by several artists in the past, most notable in a song by the group Front Line Assembly. They chose the title based on its previous use by Klaus Schulze, and an industrial group called Whitehouse. By incorporating samples from various media and naming their piece after other music they appreciated, Front Line Assembly "summarizes and remixes its own historical context" (Reed 2013). By once again reappropriating the *Mindphaser* title, and sampling from Cage, Coldplay, and Snowden, I similarly remix the historical context of this piece.

The sound collage aesthetic is apt in a piece that deals with large-scale surveillance. Such surveillance requires storing large amounts of data, most of which is irrelevant to national security. Similarly, my program pulls together mounds of user-contributed data, and creates sound collages which might be perplexing. These chaotic juxtapositions of many different types of aural material capture the messy melange of multimedia data that must be stored and sorted by the spies at the NSA. It is fitting that the spy program that forces itself onto its own citizens is represented by a database of sounds that were voluntarily contributed by some of the same people.

This technique can be tied to several characteristics of industrial music, including detournement, the cut up, and the formal-aesthetic principles of the industrial movement. First, the technique used here isn't literally detournement; I'm not taking someone else's words and turning those words against them. Rather, I use a data-collection method similar to that of the NSA as commentary on the NSA. I gather sounds from the people to comment on a group that is recording sounds of the people. Secondly, the cut-up technique, with a name derived from William Burroughs' text collage technique, has a clear relationship to the software created here. The sound collages are more-or-less algorithmic cut-ups. Each sound is taken out of the context in which its creator intended, placed into a new context with sounds that may at times seem almost unrelated, and given a political twist. Finally, the piece borrows the aesthetic ideals of Industrial Music in its formal contour. Reed argues that listening to Industrial Music often "feels like a plateau with no peak; it resists traditional differentiations of structure" (Reed 2013). *Mindphaser Edward Snowden* does have an essentially traditional structure, but on the surface, the sound collage elements have a similar feeling, which is revisited to various degrees in every piece in this collection. The sound collage elements feel like constantly flowing chaos with no ebb.

There are three primary elements that guide the form of *Mindphaser Edward Snowden*, algorithmic sound collage, manually edited samples, and performed granular synthesis. The following illustration shows how these elements interact in what I call intersecting form.

Intersecting Form



1 = Algorithmic Sound Collage

2 = Granulated or Unprocessed Snowden

3 = Manually Modified Samples (Cage, Coldplay, Snowden)

Illustration 11: Form of Mindphaser Edward Snowden

I call this form intersecting form because the elements intersect more and more as the piece progresses. A section of sound collage is presented first, followed by an interlude of the manually modified samples. Then a new section of sound collage is heard, followed close on by the granulated performance of sampled Snowden speech. All of these elements overlap. They continue to overlap more and more until the final section where all three elements combine to close out the piece.

5.2. Fundamental Disconnect

Fundamental Disconnect differs from many of the pieces in this collection in that it does not employ the software that is the primary subject of this dissertation. *Fundamental Disconnect* is an example of a piece that employs a novel algorithm and interface created for a single piece.

5.2.1. Composition

On June 26, 2006, the late Senator Ted Stevens described the internet as "a series of tubes." The octogenarian's words instantly became a meme employed by net neutrality advocates in support of their cause. For me, the quote was symbolic of the fundamental disconnect between people living in the 20th-century and people living in the 21st-century. The internet isn't a series of tubes. It isn't even a truck. In fact, it is unique and distinct from previous inventions. There is no apt analogy to earlier technology that can capture what the internet is today. So Stevens' attempt to understand and legislate the internet by comparing it to other technologies, was both unfortunate, and detrimental to the discussion of net neutrality.

Many of my compositions begin with an algorithm. I write code to experiment with different ways of modifying or interacting with sound. In early 2013, I was experiementing with applying frequency modulation to audio files. In typical frequency modulation synthesis, one sine wave is used to control the frequency of another sine wave. In my interface, a sine wave is used to control the playback rate of an audio sample. The position of the mouse cursor along the x-axis is then used to control the frequency of that sine wave. The position of the mouse cursor along the y-axis controls the frequency of another sine wave which is used to apply ring modulation to the frequency modulated audio file. In other words, the mouse position along the x-axis controls frequency modulation while the mouse position along the y-axis controls ring modulation.

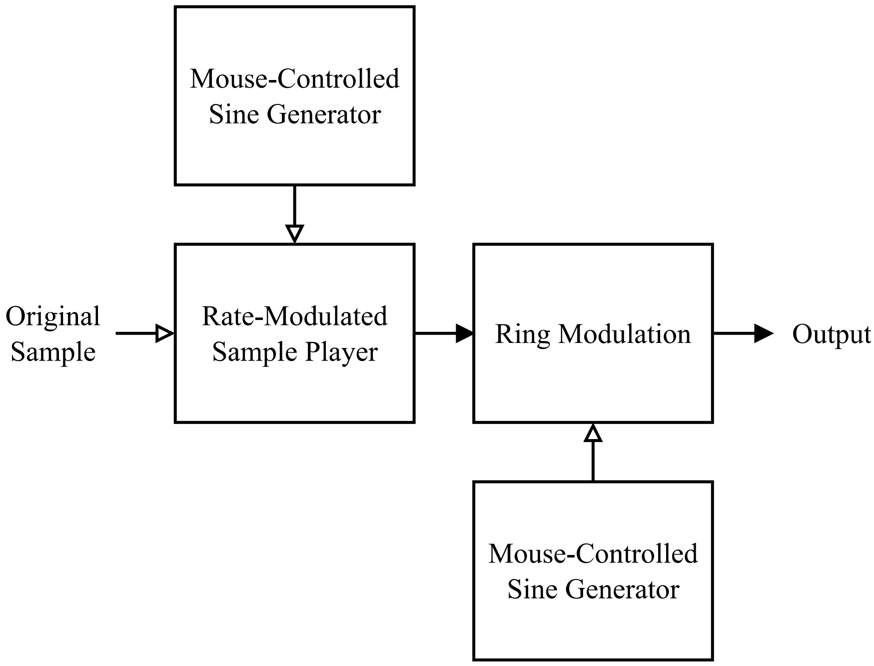


Illustration 12: Program flow in Fundamental Disconnect

The mouse position along the x-axis is also used to control how much frequency modulation is applied to the sound, allowing me to let some of the original sound shine through at key moments.

The final composition was created by loading the sample of Stevens' words into this interface, then improvising with it until I found a range of interesting sounds.

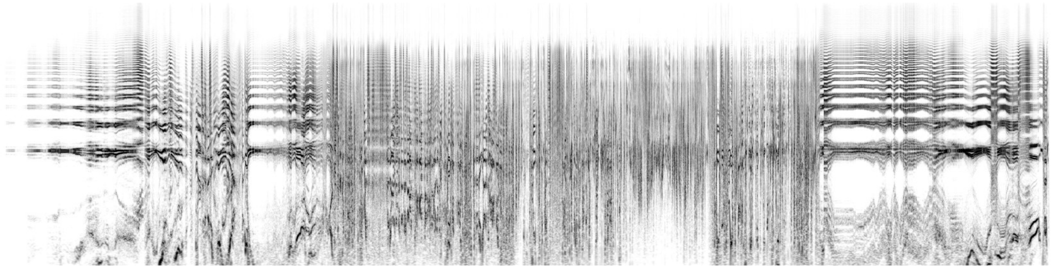


Illustration 13: Sonogram for Fundamental Disconnect

5.2.2. Interpretation

Fundamental Disconnect employs the technique of détournement. By modifying and recontextualizing Ted Stevens' words, *Fundamental Disconnect* reveals the mistake of comparing the internet to a series of tubes. Where his analogy was intended to clarify complex technology for older senators, it only muddied the debate, and made it more difficult to understand whether net neutrality was good or bad. *Fundamental Disconnect* explores the effect of his words by making them even less intelligible. Several minutes of noise are presented before it is hinted that the noise is being generated from a short twenty-five-second sample of Stevens' voice.

Although the noisy surface of the piece might sound like a neo-industrial soundscape, the form of the piece owes more to traditional western music. It is a traditional arch form, with an ending that is a varied repetition of the beginning and a climax that occurs around two-thirds of the way through the piece.

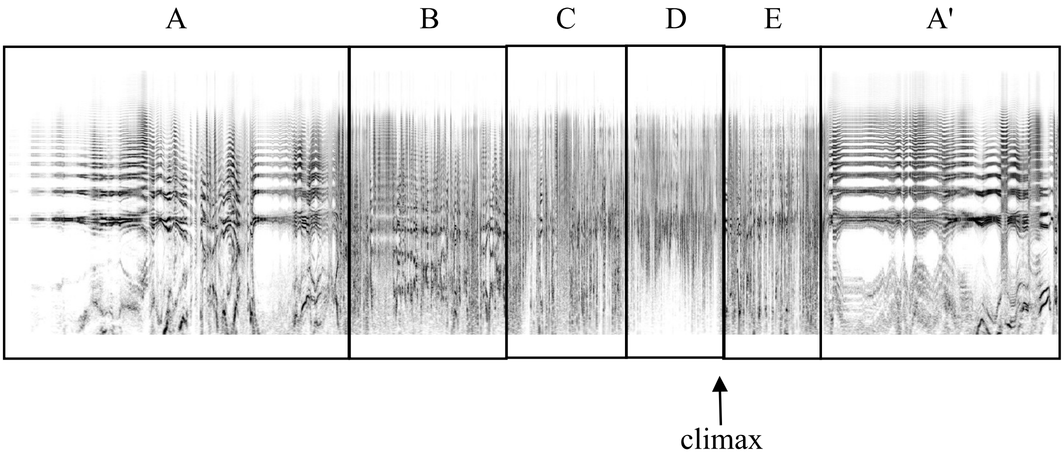


Illustration 14: The form of Fundamental Disconnect

5.3. For the Sake of Transparency

For the Sake of Transparency is one of two collaborative pieces in this collection. Much of the source material was generated by my software, then I edited the results.

5.3.1. Composition

For the Sake of Transparency was inspired by our copyright laws. Currently, copyright laws benefit middlemen, while hindering creative artists and consumers. The title of the piece is a play on the time and money that an artist must invest into authorizing samples that they want to use in their work. The task is trivial for corporations with legal departments, but nearly insurmountable to mashup artists like Girl Talk or CCC.

I began the piece by creating a sound graph. Wanting to find sounds relating to the concept of transparency, I used the search term *window*. The first sound returned was a foley-type sound effect of a window being opened created by user RutgerMuller and titled *Window Sound Effects 2*. Through basic similary search, this sound was connected to sounds of walking on gravel, a book dropping repeatedly, and someone playing with knick-knacks on their desk. All of these sounds involve several actions, or one action repeated several times. So in many of the sounds in the resulting graph, there is more than just a single attack.

Graph Details for *For the Sake of Transparency*	
Node Count	120
Edge Count	380
Degree Distribution	
Degree	**Occurrences**
1	82
2	12
3	3
4	2
5	1
6	3
8	1
10	1
12	2
13	2
14	4
15	5
16	1
19	1

Figure 3: Graph details for For the Sake of Transparency

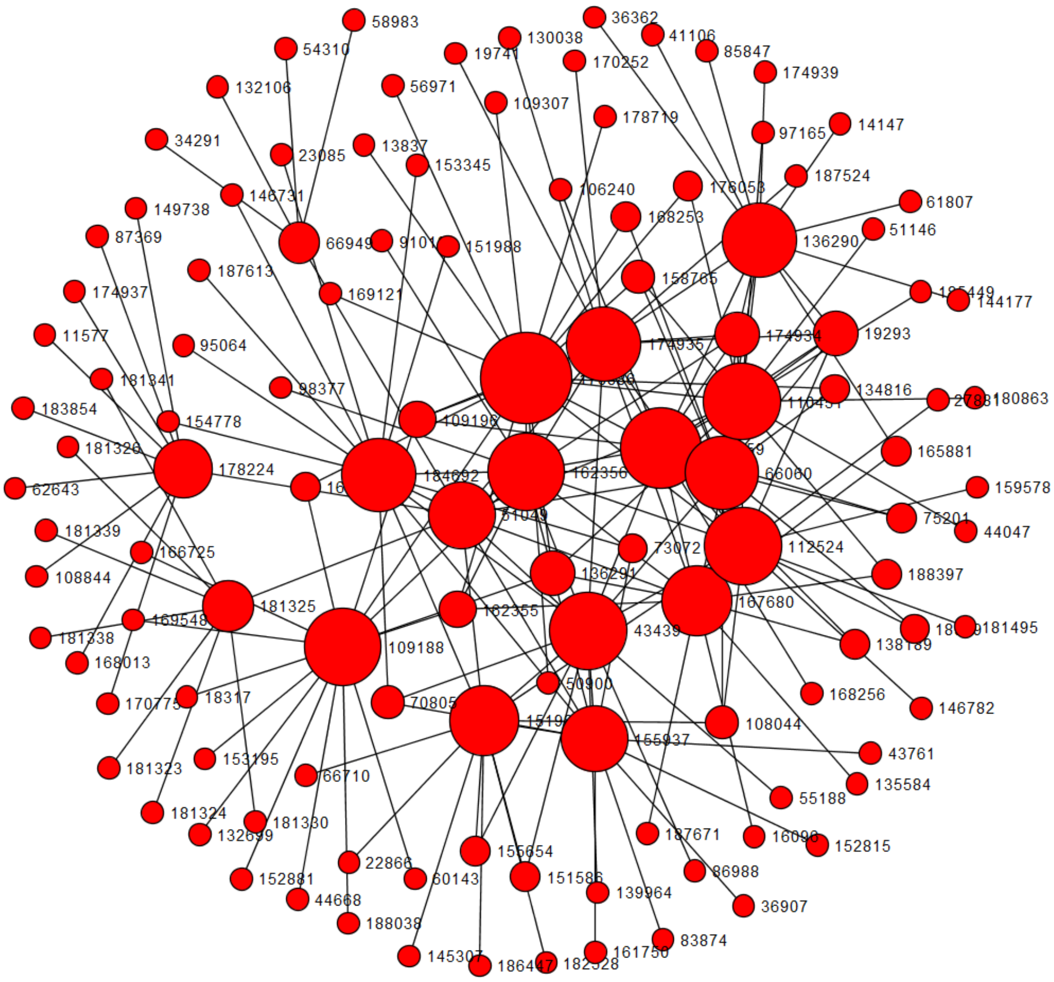


Illustration 15: Sound graph for For the Sake of Transparency

In this piece I wanted to explore the contrast between activating a graph using swarm intelligence and activating it using the king automaton. I wanted to contrast the progression of sounds generated by the swarm with the cacaphony generated when the king automaton triggered many sounds at once. I activated the graph using both strategies and stored the results.

After generating the two sound collages using those activation strategies, I split each one up into what seemed like natural subdivisions. Then I interspersed them, first inserting a section generated by the swarm, then a section generated by the automaton. This led to an ebb-and-flow of sound.

Because the piece is about copyright law, I wanted to incorporate a sample from a hit song. I used a Smashing Pumpkins album from the mid 1990s called *Melon Collie and the Infinite Sadness*, and I sampled the piano chords from the opening track. The selection of this piece was partly arbitrary, and partly guided by my compositional goals. I wanted a tonal piece that would work well when reorganized. After slicing out many different chords, I trimmed out most of the higher harmonics in these chords, time stretched them significantly, then resynthesized them to new audio files. The resulting pad-like chords are a shadow of the original, even though they do not retain the surface-level character of the original.

After resynthesizing these chords, I intuitively matched them with the mix of algorithmic output. I finished the piece by manually diffusing it to two channels, and applying compressing and equalization.

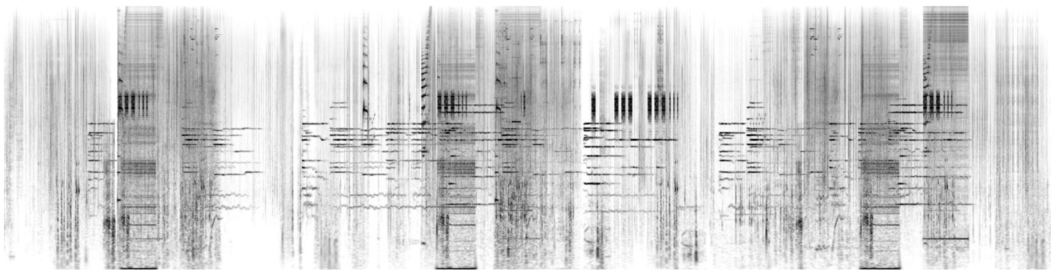


Illustration 16: Sonogram of For the Sake of Transparency

5.3.2. Interpretation

For the Sake of Transparency shows that sound graphs can be used to generate a number of different textures. Although it only uses one sound graph, it shows how a graph can lead to both smoothly-evolving chains of sounds as well as cacaphonic justapositions of sounds.

For the Sake of Transparency does not directly confront copyright laws the way that mashup artists like Girl Talk, or web artists like Johannes Kreidler have done. The lack of instantly-recognizable copyrighted material means that the piece can be heard without hearing the underlying politics. However, since all samples were downloaded from FreeSound, or derived from commercial music, it shares a compositional approach with those artists.

The form of *For the Sake of Transparency* is partly traditional, and partly non-traditional. The listener may grab onto the harmonic flow of the dreamlike chords that undercut much of the piece, but these are not structural elements. The harmonies flow along in a more or less random rushuffling, similar to how harmonies might unfold in a minimalist or ambient piece of music. The harmonies may distract the listener from the rather traditional form of the recombined collages.

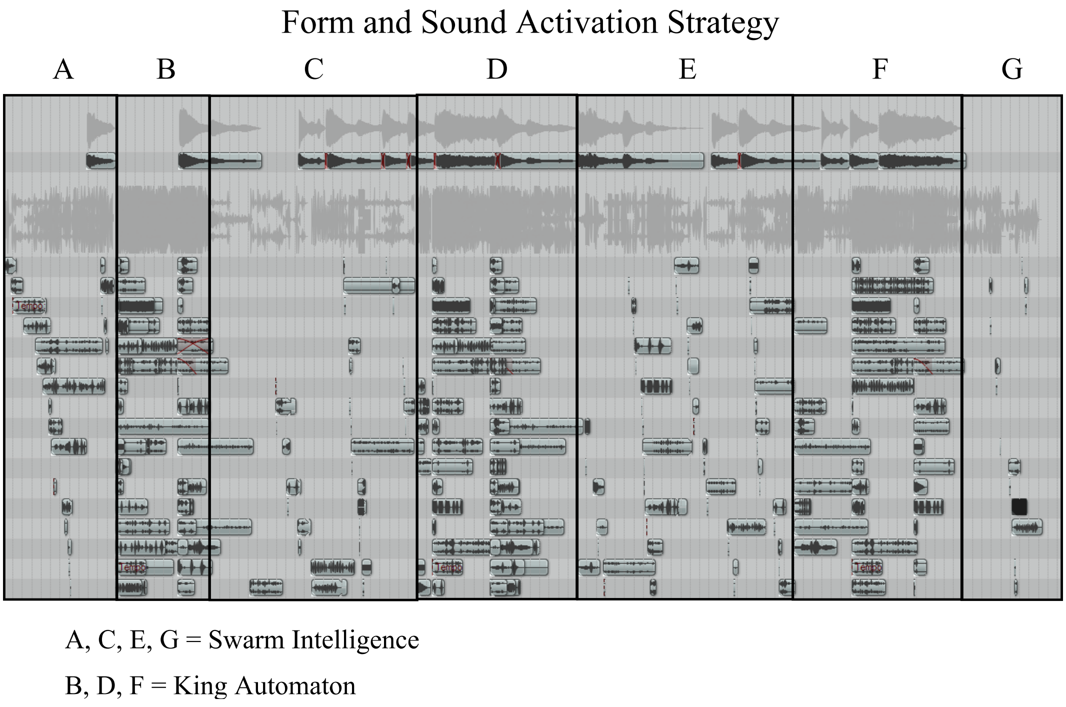


Illustration 17: Form of For the Sake of Transparency

The form is a varied binary form. Although there are no explicit repeats, the activation strategies alternate in each section. There are three sections filled with content from the king automaton activation strategy, and five intersecting sections filled with content from the swarm intelligence activation strategy. There is a small amount of overlap between sections, and occasionally there is a period where only the chordal synthesizer can be heard.

5.4. The End of My Career

5.4.1. Composition

The End of My Career is one of two purely algorithmic pieces in this collection. The sound graph began with the search term *orchestra*. This search returned the sound with id 145679. Uploaded by user LukeSharples, this sound is called *Space Orchestra A3*, and it sounds like a pipe organ mixed with an orchestra playing a sustained chord. Through basic similarity search, this sound was connected with several other *Space Orchestra* sounds uploaded by the same user, then to sound 56646, a thick synthesized sound with a linearly rising pitch. This thick spectrum with linearly rising or falling frequencies characterizes many of the sounds that were eventually included in the graph. This includes sounds of people shouting, synthesized rocket-like sounds, the doppler-effect sound of race cars passing, and synthesized tones.

The graph contains three main sound areas. One characterized by sounds with linearly rising or falling spectra, another characterized by instrumental sounds and riffs, and a third characterized by vocal sounds. Any music created with this graph will be shaped by the interplay of those groups of sounds.

Graph Details for *The End of My Career*	
Node Count	100
Edge Count	327
Degree Distribution	
Degree	**Occurrences**
1	68
2	12
6	1
7	1
9	2
11	3
12	6
13	5
14	1
20	1

Figure 4: Graph details for The End of My Career

I used the swarm intelligence-based activation strategy to generate the sound collage for *The End of My Career*. By moving between the areas the swarm generates an interplay between similar and contrasting sounds.

In the resulting collage I removed sounds that were unreadable by my computer. Then I manually diffused it into a two channel mix, and applied compression.

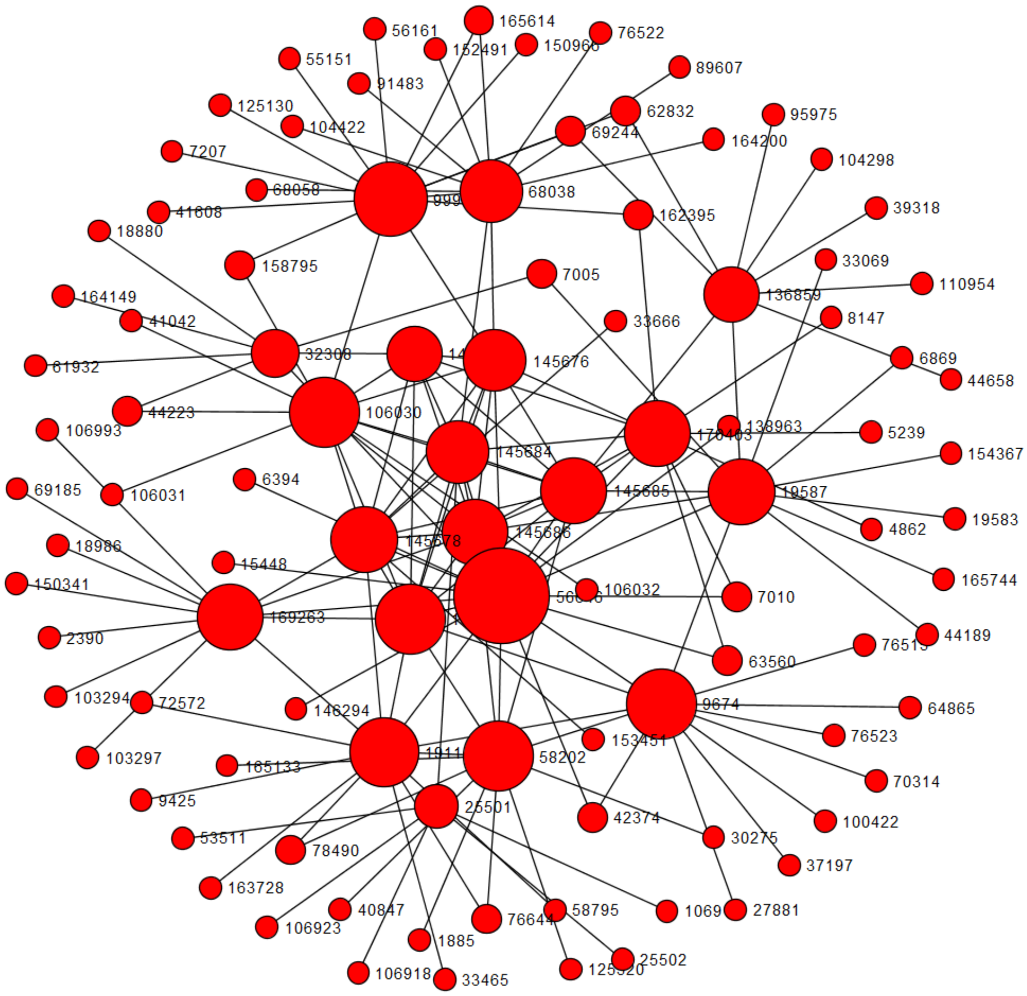


Illustration 18: The sound graph for The End of My Career

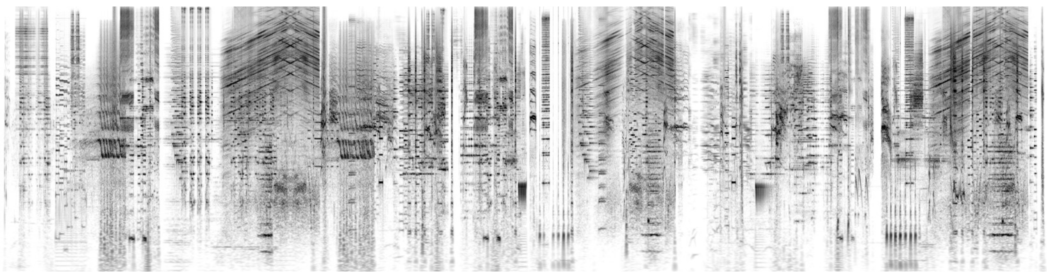


Illustration 19: Sonogram of The End of My Career

5.4.2. Interpretation

In *The End of My Career*, form emerged from the swarm's motion despite the fact that it wasn't explicitly programmed to make musical forms. In essence, the form is a side-effect of the swarm algorithm. This approach to the problem of form is a way of creating structure in music without explicitly applying patterns developed by previous musicians.

In earlier swarm intelligence work I developed a way for swarms to store a stigmergic record of their movements. I added simulated grass to the model, converting the swarm into simulated sheep. The sheep roam around the field looking for food. They eat the food where they stop, and they are constantly moving toward areas with more food in areas that haven't been recently visited. In this way, a swarm can generate long-term structure even though it doesn't have a plan.

In this piece, the swarm generates form by weaving together the three sound areas in the graph, as shown in the following three diagrams. The first diagram shows the recurrences of sounds with sharply linearly ascending or descending spectral components. The jet engine sound seems to mark the beginning, middle and end of the piece, even though the end of the piece was determined by a five minute limit.

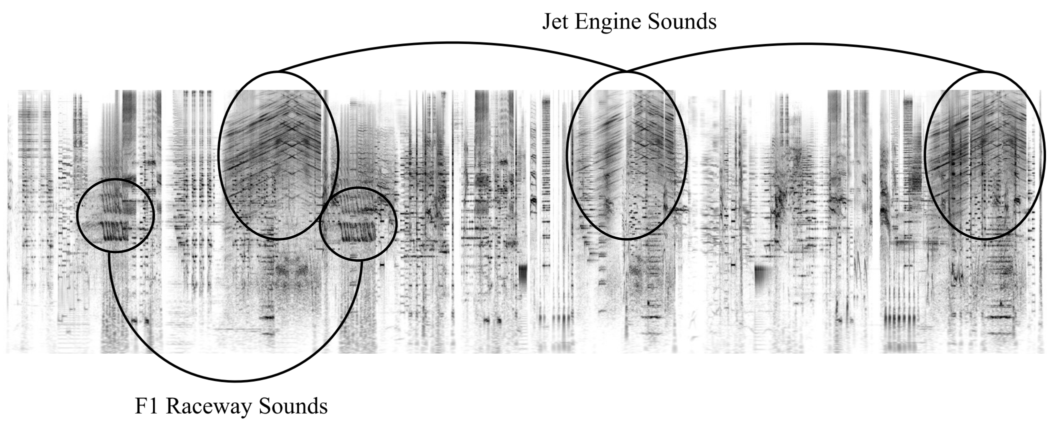


Illustration 20: Recurrences of sounds with sharp linear spectra in The End of My Career

In the second form diagram for *The End of My Career*, you can see how the instrumental sounds are reused. In the latter two appearances of the jet engine sound, it is combined with organ-like sustained chords. The nearness of these areas in the graph, along with the behavior of the swarm, leads to repetition with variation.

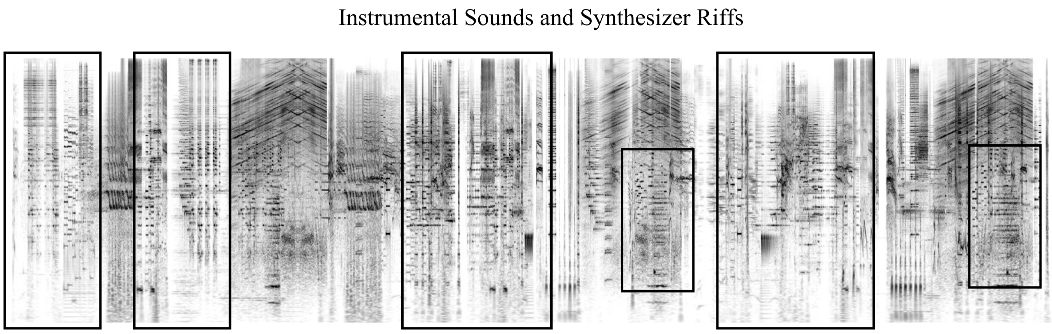


Illustration 21: Recurrences of instrument and synthesizer riffs

The final diagram shows the recurrences of shouting and screaming sounds. These sounds mark transitions between the two sound areas in the previous graphs, although they are difficult to distinguish in the sonogram. The way these shouting and screaming sounds blend with their surroundings is one of the most striking surface-level features of this piece. This is an example of how the swarm can take advantage of the aural similarity relationships encoded in a sound graph.

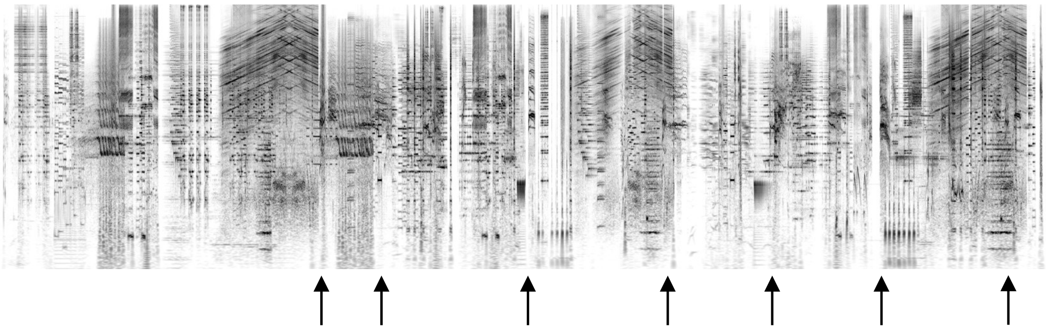


Illustration 22: Recurrences of screaming and shouting sounds

Part of the novelty of using swarm intelligence in music composition is that it's impossible to fully explain the structure. However, it's clear that distinct recurring temporal structures are created by the actions of the swarm on the graph. The ability of the swarm to generate temporal patterns, and varied repetition, without resorting to a planning process, is one of the unique demonstrations in this work.

5.5. Wordless

Wordless employs a novel algorithm and interface created for a single piece.

5.5.1. Composition

In *Wordless* I address the way that the voice of the people has been silenced. I began by searching for sounds related to coughing, wheezing and asthma. This led to sounds of asthma attacks by users shopp and powerswitch, and a recording of a person breathing through a gas mask by user lysander-darkstar. I followed the links to find aurally similar sounds such as ambient recordings of warehouses, trains, vacuum cleaners and some airy synthesized sounds.

The composition of this piece began with a new interface for mangling sounds. The interface allows the player to select samples, then play them back with effects that are modified by the position of the mouse cursor. The mouse position along the x-axis controls the amount of delay applied to the sounds, the frequency at which they are filtered, and the duration of the attack envelope. The position of the mouse cursor along the y-axis is mapped to the frequency of the modulating sine wave that is used in ring modulation, and the duration of the decay envelope.

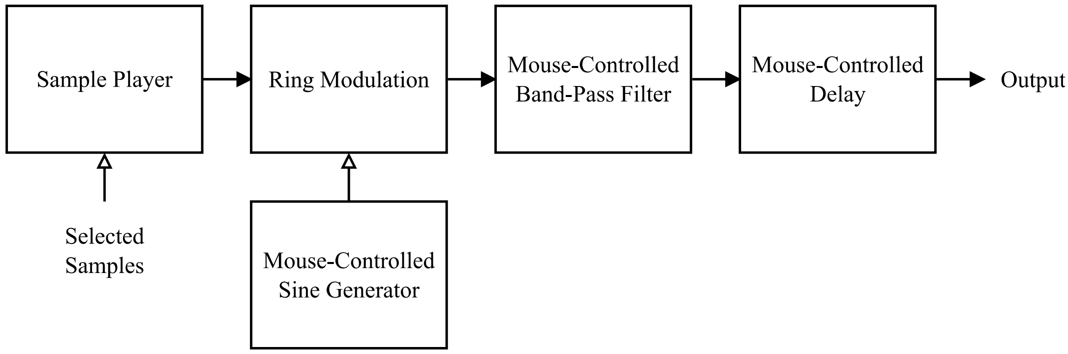


Illustration 23: Program flow for the software used in Wordless

The program allows the player to modulate several samples simultaneously. The samples are listed in toggle buttons along the bottom of the interface. The selected samples are triggered by left-clicking anywhere within the program window. By moving the mouse, the user can change the attack, decay, filter, and modulation as the sounds are played.

With this program, the player can design variations on sounds on the fly. By drawing different patterns in the interface, the player can make gestures that correspond to specific sound modifications. I made a list of sounds to be used and the gestures that should be applied, then recorded a take that became *Wordless*.

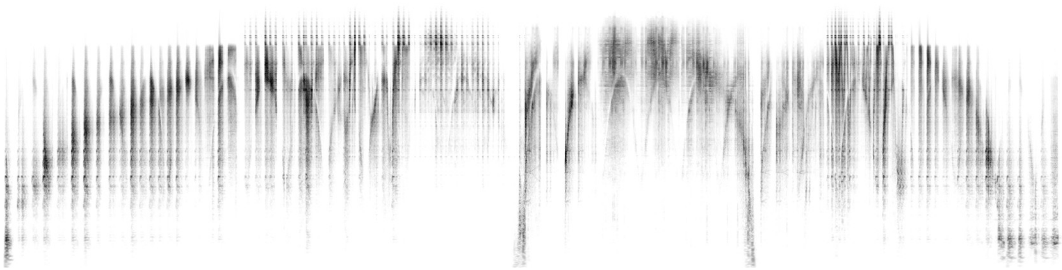


Illustration 24: Sonogram of Wordless

5.5.2. Interpretation

Wordless unfolds in an arch form.

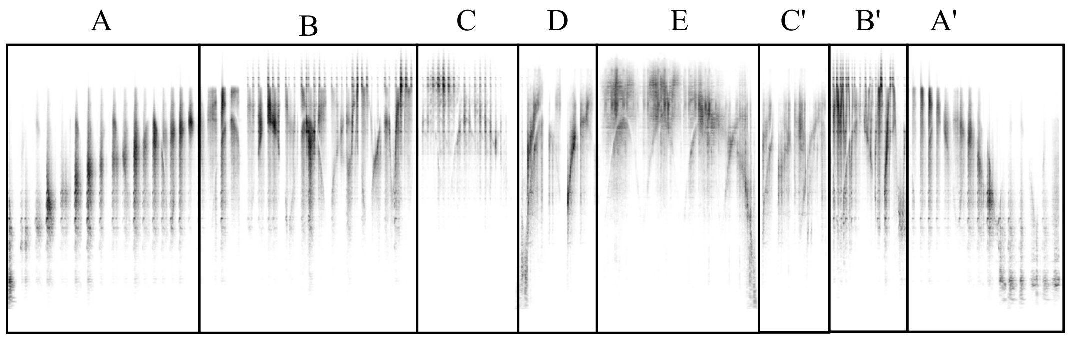


Illustration 25: The form of Wordless

Wordless is an exploration of timbre. The piece explores the sounds of whispers, coughs, wheezes, and gasps. This is partly accomplished through sound selection, and partly through modification of those sounds.

The source sounds are mostly unrecognizable, although distinct wheezing can be heard in the section marked C in the form diagram. All of the sounds share similar spectral content. The sprectra are filled with broadband noise with strong components in the range of the human vocal formants.

The sounds are modified to bring out the rhythm of breathing. The introduction almost sounds like a long series of inhalations and exhalations. In later sections, ring modulation is used to speed up or slow down the pulse of the sounds.

5.6. Toll

5.6.1. Composition

One of the first sound graphs I created was based on the search term *chimes*. I selected this term because I previously created a piece using wind chime sounds and swarm intelligence. The term *chimes* also has a pedigree in acousmatic music, due to Denis Smalley's *Wind Chimes*. Finally, I used the term *chimes* because I knew that users had uploaded many high quality wind chime recordings to FreeSound.

The graph created for this piece is an example of a situation where my software surprised me. Rather than getting a graph full of wind chimes, I got a graph full of music box sounds. In fact, the graph was made up primarily of music boxes that played Christmas music. This was so far from what I intended that I originally saved the graph and moved on to a new graph.

Graph Details for *Toll*	
Node Count	100
Edge Count	272
Degree Distribution	
Degree	**Occurrences**
1	63
2	13
3	9
5	1
6	1
7	1
8	3
11	2
12	4
13	2
18	1

Figure 5: Graph details for Toll

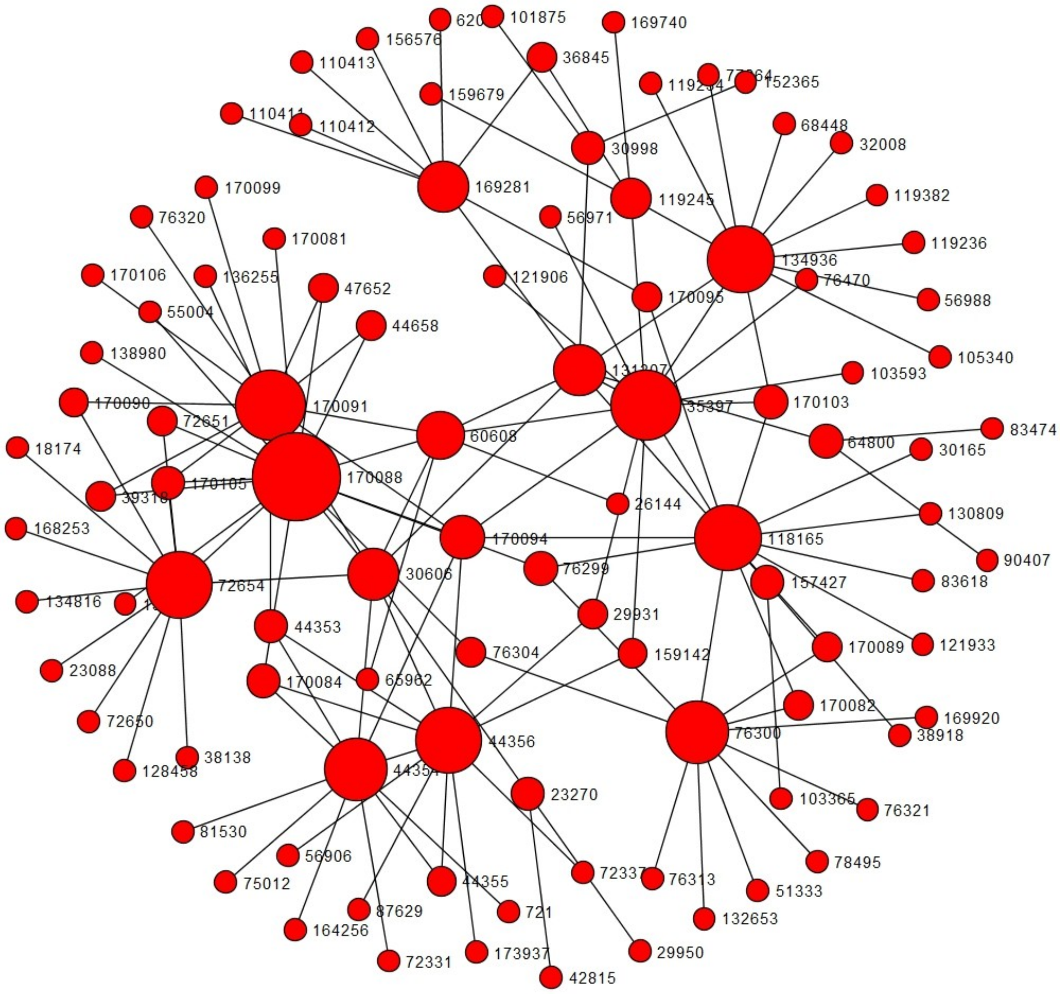


Illustration 26: The complete sound graph for Toll

Later, when I was writing the music for this collection, I thought about all of the young soldiers who died in Iraq and Afghanistan. I wanted to write a piece to lament these soldiers, and thought that the chimes graph would be appropriate.

I selected the king automaton strategy to activate the graph. Recall that this strategy starts by activating the initial sound, then activates each sound at distance one from that node, then each sound at distance two, and so on until all sounds are activated. This piece is short partly because all sounds are activated only once.

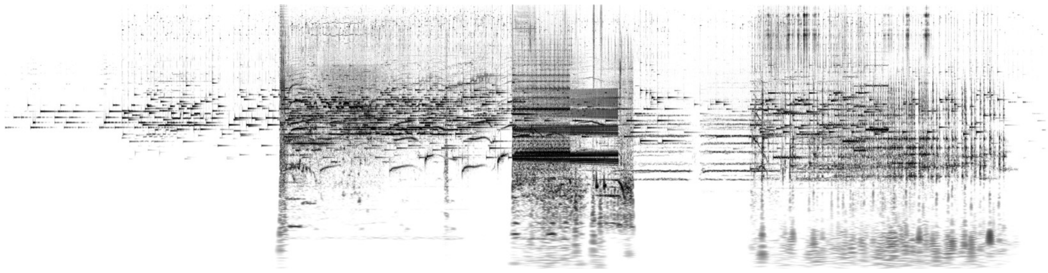


Illustration 27: Sonogram of Toll

5.6.2. Interpretation

Toll is subversive. On the surface, the title refers to the chiming bells that are heard throughout the piece. However, the word toll normally implies that bells are ringing for a negative reason, such as in the title of Hemingway's novel *For Whom the Bell Tolls*, which is drawn from a John Donne meditation on death. If a listener understands the inspiration for this piece, then he might also relate it to the phrase *death toll*, and the daily body counts emanating from Iraq in Afghanistan.

The sound of and associations with music boxes are also subverted in this piece. Music boxes are usually associated with childlike innocence. Those associations are so strong that it would be impossible to use a music box in a piece without drawing out those associations. In this piece the music boxes refer to the ages of the soldiers dying in foreign wars. The sound of music boxes are intended to remind the listener that the soldiers are still children, but in *Toll* the sounds are presented in an anarchic setting. The sounds of music boxes are presented with many other music boxes, synthesized sounds and human shouts and screeches.

6. Conclusion

In this project, I addressed my lifelong compositional goals while exploring my political beliefs through music. These compositions reflect my musical and political values in several ways.

First, I showed that composing with all sound is not entirely out of reach with the infrastructure of the internet and several web API. FreeSound gives programmers access to thousands of sounds, and thorough audio descriptors of those sounds. By using aural and lexical relationships between sounds we can organize them in a way that mimics Schilling's model of creativity.

Second, I achieved my goal of writing new music, as I define the term. To me, new music is music that couldn't have been written yesterday. Generally I define levels of newness as music that couldn't have been written ten, twenty or thirty years ago. My music couldn't have been written twenty years ago in 1993. The required websites didn't exist at that time, and the web infrastucture wasn't programmatically accessible via APIs.

Third, I incorporated my political beliefs into both the process and conception of the compositions. In *Mindphaser Edward Snowden*, for instance, the algorithms that combine sounds from web users is a perfect analog for how the NSA collects data on all Americans. In *Fundamental Disconnect*, I use the technique of detournement, turning Senator Ted Stevens' words against his cause. In *Toll*, I let the sounds speak to create a confusing, chaotic, and childlike atmosphere that is difficult to interpret.

Finally, I suggested a data structure that can be used to explore art on the internet in the future. My *creativity networks* capture the meaning that implicitly connects documents on the internet. These data structures show how aural, lexical, or other quantifiable relationships can be used to combine web documents in a way that is related to how they might be stored in the brain.

7. Appendices

7.1. Technical Details

Sonograms were generated using Photosounder's Save-To-Bitmap feature. This generated sonogram images that were 571 pixels tall and many thousands of pixels wide. The sonograms were then scaled to 500 pixels tall and 2000 pixels wide. Then the colors were inverted to gray-on-white, and the brightness and contrast were adjusted until the contours of the sonogram were clearly visible.

All software created for this project was coded in Java or Processing. The Beads library for audio was used to implement the synthesizers in *Fundamental Disconnect* and *Wordless*. Beads was created by Oliver Bown and is available at beadsproject.net.

8. Supplemental Files

8.1. Music

The music created for this dissertation is available on the included audio CD.

9. Bibliography

About Wordnik. Accessed April 26, 2013. http://www.wordnik.com/about.

Albert, Réka, and Albert-László Barabási. 2002. "Statistical mechanics of complex networks." *Reviews of modern physics* 74 (1): 47.

Baker-Smith, Ben. 2012. *Infinite Glitch*. Accessed July 8, 2013. http://bitsynthesis.com/infiniteglitch/.

Barabasi, Albert-Laszlo. 2002. *Linked: The New Science of Networks*. Cambridge, MA.: Perseus Publishing. Kindle edition.

Barbosa, Alvaro. 2005. "Public Sound Objects: a shared environment for networked music practice on the Web." *Organised Sound* 10(3): 233-242.

Boden, Margaret A. 2004. *The Creative Mind: Myths and Mechanisms*. 2nd ed. London: Routledge.

Cope, David. 2005. *Computer Models of Musical Creativity*. Cambridge, MA: The MIT Press.

Daniel, Sharon and Karen O'Rourke. 2004. "Mapping the Database: Trajectories and Perspectives." *Leonardo* 37 (4): 286-296.

De Bono, Edward. 1970. *Lateral Thinking: Creativity Step by Step*. New York: Harper & Row.

Demers, Joanna. 2010. *Listening through the Noise: The Aesthetics of Experimental Electronic Music*. New York: Oxford University Press. Kindle edition.

Feigenbaum, Lee, and Elias Torres. 2006. "A meaningful Web for humans and machines, Part 1: How humans can share the wealth of the Web." *IBM*

DeveloperWorks. Last modified October 24, 2006. http://www.ibm.com/developerworks/web/library/wa-hmach1/index.html.

Frascina, Francis, Marjorie Perloff and Christine Poggi. 2006. "Collage." *Encyclopedia of Aesthetics*. Oxford Art Online. Oxford University Press, accessed July 11, 2013, http://www.oxfordartonline.com/subscriber/article/opr/t234/e0118.

FreeSound API Documentation. Accessed April 18, 2013. http://www.freesound.org/docs/api/resources.html#sound-search-resource.

Godøy, Rolf Inge. 2006. "Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus." *Organised Sound* 11(2): 149-157.

Gärdenfors, Peter. 2004. "Conceptual Spaces as a Framework for Knowledge Representation." *Mind and Matter* 2 (2): 9-27.

Graham, Renee. 2004. "Jay-Z, the Beatles meet in 'Grey' area." *Boston Globe*. Accessed August 11, 2013. http://www.boston.com/news/globe/living/articles/2004/02/10/jay_z_the_beatles_meet_in_grey_area/.

Harrison, Nate. 2004. *Can I Get An Amen?* Accessed July 23, 2013. http://nkhstudio.com/pages/popup_amen.html.

Hsu, Wun-Ting, and Wen-Shu Lai. 2013. "Readymade and Assemblage in Database Art." *Leonardo* 46 (1): 80-81.

Jenkins, Henry. 2006. *Fans, Bloggers, and Gamers*. New York: New York University Press. Kindle edition.

Kane, Brian. 2007. "L'Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction." *Organised Sound* 12 (1): 15-24.

Katz, Mark. 2010. *Capturing Sound: How Technology Has Changed Music*.Berkeley, CA: University of Californa Press.

Kreidler, Johannes. 2008. *Product Placements (2008)*. Accessed July 8, 2013. http://kreidler-net.de/productplacements-e.html.

Kreidler, Johannes. 2009. *Johannes Kreidler GEMA-Aktion product placements Doku*. Accessed July 8, 2013. http://www.youtube.com/watch?v=EAptRZlwziA.

Marinetti, F. T. 2002. *Selected Poems and Related Prose*. Translated by Elizabeth R. Napier and Barbara R. Studholme. New Haven, CT: Yale University Press.

Martınez, Elena, Oscar Celma, Mohamed Sordo, Bram De Jong, and Xavier Serra. 2009. "Extending the folksonomies of freesound. org using content-based audio analysis." In *Sound and Music Computing Conference, Porto, Portugal*.

Nattiez, Jean-Jacques. 1990. *Music and Discourse: Toward a Semiology of Music*. Translated by Carolyn Abbate. Princeton, NJ: Princeton University Press.

Ostrowski, Matthew. 2007. "The Anxiety of the Client: Dilemmas of the Database as a Compositional Tool." "Unyazi" Special Issue, *Leonardo Electronic Almanac* 15 (1-2). http://leoalmanac.org/journal/vol_15/lea_v15_n01-02/mostrowski.asp.

Payton, Rodney J. 1976. "The Music of Futurism: Concerts and Polemics." *The Musical Quarterly* 62 (1): 25-45.

Pritchett, James. 1996. *The Music of John Cage*. Cambridge: Cambridge University Press.

Rainey, Lawrence, Christine Poggi, and Laura Witman, eds. 2009. *Futurism: An Anthology*. New Haven, CT: Yale University Press.

Redhead, Lauren. 2011. "Relational Aesthetics and the Western Canon of Increasingly Historical Works." *British Postgraduate Musicology* 11. http://bpmonline.org.uk/bpm11/redhead_relational_aesthetics_and_the_western_canon_of_increasingly_historical_works.pdf.

Reed, S. Alexander. 2013. *Assimilate: A Critical History of Industrial Music*. New York: Oxford University Press. Kindle edition.

Ritchie, Graeme. 2007. "Some Empirical Criteria for Attributing Creativity to a Computer Program." *Minds & Machines* 17: 67-99.

Roma, Gerard, Perfecto Herrera, and Xavier Serra. 2009. "Freesound Radio: supporting music creation by exploration of a sound database." In *Workshop on Computational Creativity Support (CHI2009)*.

Roma, Gerard, Perfecto Herrera, Massimiliano Zanin, Sergio L. Toral, Frederic Font, and Xavier Serra. 2012. "Small world networks and creativity in audio clip sharing." *International Journal of Social Network Mining* 1 (1): 112-127.

Russolo, Luigi. 1913. "The Art of Noises: A Futurist Manifesto." In *Futurism: An Anthology*, edited by Lawrence Rainey, Christine Poggi and Laura Witman, 133-139. New Haven, CT: Yale University Press.

Schilling, Melissa A. 2005. "A 'Small-World' Network Model of Cognitive

Insight." *Creativity Research Journal* 17 (2-3): 131-154.

Schwitters, Kurt. 1922. *Ursonate*. Ecart performance group, 1979.

Stoll, Thomas M. 2013. "CorpusDB: Software for Analysis, Storage, and Manipulation of Sound Corpora." In *Musical Metacreation: Papers from the 2013 AIIDE Workshop*.

Thorogood, Miles, Phillippe Pasquier and Arne Eigenfeldt. 2012. "Audio Metaphor: Audio Information Retrieval for Soundscape Composition." In *Proceedings of the 9th Sound and Music Computing Conference*.

Wiggins, Geraint A. 2006. "Searching for Computational Creativity." *New Generation Computing* 24 (3): 209-222.