

Evolving electroacoustic music: the application of genetic algorithms to time-domain waveforms

Cristyn Magnus

CRCA, Cal(IT)², Department of Music, University of California, San Diego
La Jolla, Ca. 92093-0326
cmagnus@ucsd.edu

Abstract

This paper describes a modified genetic algorithm that works directly on time-domain waveforms to produce genetically evolved music, the formal structure of which is derived from the evolutionary process. Genetic algorithms are briefly characterized, a modified genetic algorithm that works directly on waveforms is defined, and the compositional results are described.

1 Introduction

1.1 Goals

My goal is to use genetic algorithms to produce a series of electroacoustic pieces in which the ecological process of evolution is transparent in the form of the piece. Although I am using a machine learning technique, my goal is not to get my machine to learn, to model the creative process, or to create a general-use compositional tool.

Genetic algorithms are flexible enough to be applied to a wide array of problems. The form that this application takes will strongly reflect the designer's intuitions about the domain in question. For a review of the use of genetic algorithms in music, see Burton and Vladimirova (Burton and Vladimirova 1999). At each stage of programming, choices must be made that introduce designer bias into the system (Magnus 2003). By embracing the inevitable design bias, I can create a process that will explore new aesthetic possibilities.

1.2 Overview of Genetic Algorithms

Genetic algorithms use the biological metaphor of an evolving population to explore large-dimensional problem spaces (Holland 1992). Each parameter of a solution is mapped onto a *gene*; a string of genes that contains the entire mapping of a solution is called a *chromosome*. The *error* of each solution is measured. This is used to calculate a *fitness function*, which

defines the probability each individual has of reproducing. Fitter solutions have a higher probability of reproducing than unfit solutions. Members of the population of potential solutions *sexually reproduce*. The chromosomes of both parents are cut at some *crossover point* and spliced together to form a new individual. During reproduction, genes have a slight probability of *mutating*, resulting in offspring that combine traits of both parents but also contain new genetic material. Because unfit individuals are unlikely to reproduce, harmful mutations will be lost but beneficial mutations will become increasingly prevalent. Over time, the entire population will become increasingly fit until it finds some local minima in the error function.

2 Algorithm

2.1 Representation

Genes are an abstract representation of some parameter of the space; each individual is a chromosome. Here, chromosomes are time-domain waveforms. Using instantaneous samples as genes would be a bad idea: sexual reproduction will introduce clicks; mutation will introduce noise. To allow genetic operations to function meaningfully, I define a gene as a segment of a waveform between two zero crossings (figure 1).

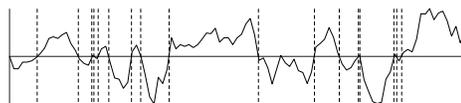


Figure 1: A waveform with dotted lines delineating genes.

An individual's life-span is the duration of its playback. Generations of output are concatenated onto a single audio file that represents the evolutionary process over time. All individuals in the population are multiplied by their fitness

and played simultaneously. Chromosomes can have arbitrary lengths, with some fixed upper bound. When an individual waveform dies (i.e. finishes playing in output), two parents are selected from the population to produce an offspring to take that individual's place in the population.

2.2 Fitness

Fitness is based on similarity to a target waveform. An individual's error is calculated by summing the difference between the desired amplitude and the actual amplitude for each instantaneous sample in the waveform (figure 2). Although the population will retain much of its original character, waveforms with similar lengths, frequencies, and amplitudes to the target waveform will become more prominent. Biodiversity can be encouraged by incorporating general features of the population into the fitness criteria. An individual's fitness drops slightly each time it reproduces, decreasing the chances that the offspring from a handful of individuals will dominate the next generation.

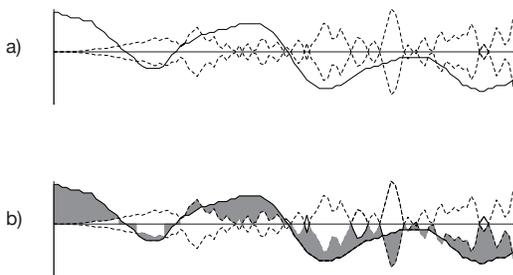


Figure 2: a) A waveform from the population (solid line) shown with the target waveform (dotted line). b) The waveform's error is shown in grey.

2.3 Reproduction

Reproduction is the process by which a generation is derived from the previous generation. Sexual reproduction is carried out by splicing genetic material from two individuals to produce one individual in the next generation. For each offspring, two parents are selected from the population. The probability that an individual will be selected as a parent is based on its fitness. Each parent is divided at some randomly selected crossover point. The location of the crossover point is adjusted to make sure it falls on a zero crossing. The first part of one parent is attached to the last part of the other parent (figure 3). The crossover point can be different on each parent, allowing offspring to have different lengths from their parents. Offspring can potentially be as long as the combined lengths of both parents.

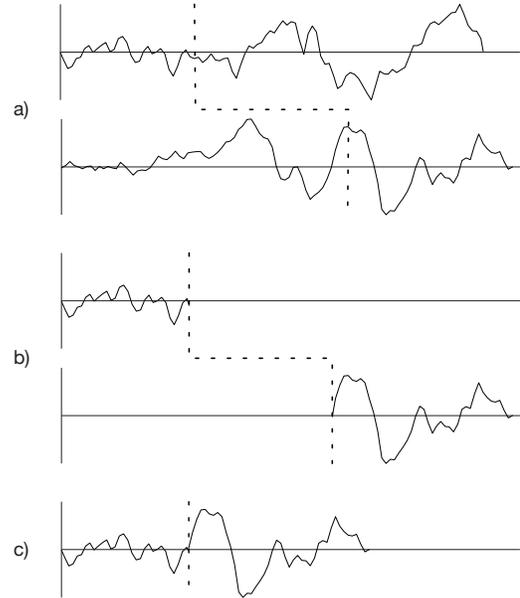


Figure 3: a) Two parent waveforms. The dotted line represents the randomly selected crossover point. b) The crossover point adjusted to fall on gene boundaries (zero crossings). c) The offspring waveform.

2.4 Mutation

Mutation introduces changes into the population. Mutation occurs immediately after the parent genes are combined to produce the offspring. Each gene has a slight probability of mutating. When a gene is selected for mutation, it and a string of genes that immediately follows it undergo the same mutation.

A typical mutation function adds a random number to a gene. An extension of this concept adjusts the amplitude of a gene (figure 4 a). This is done by multiplying genes by a number randomly selected between predetermined bounds. A further extension raises genes to a power that is randomly selected between predetermined bounds (figure 4 b). To prevent the exponentiation from severely amplifying or attenuating the segment being mutated, each segment is normalized after exponentiation so that it retains its original maximum amplitude.

Because mutation is being applied to strings of genes, rather than single genes, we can draw inspiration from the types of errors that happen in actual gene transcription. Mutation functions can reverse a string of genes (figure 5 a), remove a string of genes entirely (figure 5 b), repeat a string of genes a random number of times (figure 5 c), or swap neighboring strings of genes (figure 5 d).

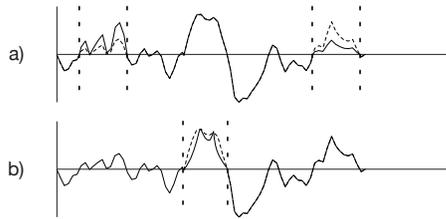


Figure 4: a) Mutation by amplification. b) Mutation by exponentiation. The original waveform is represented by a dashed line.

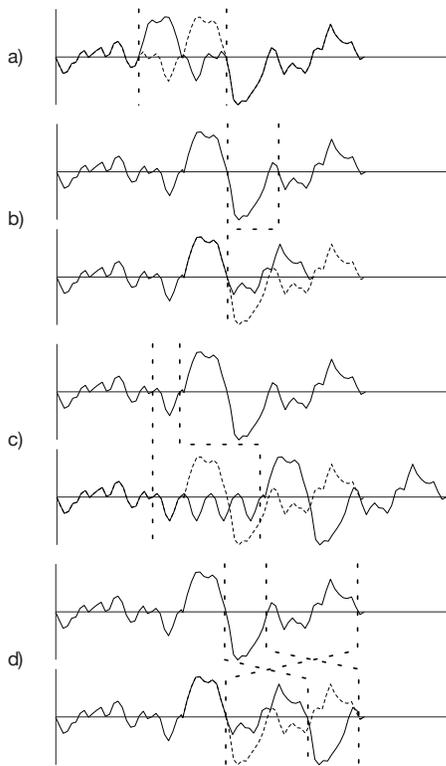


Figure 5: Mutation operations: a) reverse. b) remove. c) repeat. d) swap. The original waveform is represented by a dashed line.

3 Compositional Framework

In a constant environment, the algorithm described above would eventually converge to a local minimum where all individuals would have roughly the same length as the target waveform and would have acquired some of its amplitude envelope and frequency characteristics. However, evolution does not occur in a vacuum—the world is constantly changing. To create global compositional structure, I define a *world* in which the waveforms evolve.

For a given piece, the world will be characterized by some number of locations. These locations may be mapped spatially onto speakers. The *environment* at these locations will initially be defined by some target waveform and some set of mutation probabilities. Individuals within the population will have some probability of migration. In this way, sounds with new characteristics will enter each location, enhancing biodiversity.

The world will be characterized by probability of change. Both target waveform and mutation probabilities can change. There are two sorts of changes that environments can undergo. One is the slow drift that is seen in ice ages: these take place over an enormous amount of time from the perspective of individuals but happen many times over the evolution of a species. This is simulated by slowly cross-fading between two target waveforms. The other is the drastic change that results from catastrophic events, such as fire decimating a forest, causing it to be replaced by grassland. This is achieved by replacing the target waveform with a completely different waveform.

4 Results

4.1 General Description of Output

As evolution occurs, all of the waveforms in the population are written to a single sound file with each individual waveform weighted by its fitness. This weighting causes fit individuals to rise to prominence. Each time a waveform ends, a new individual is generated from the population. The new individual's playback begins immediately at the end of the waveform it replaces. Because the initial biodiversity is very high, the beginning of the output file is a wash of textures reminiscent of the timbres of the initial population. Within a few generations, a few fit individuals dominate the mix, causing a sound in which particular features of the initial population can be identified. At this point, the type of mutations permitted significantly impacts the resulting sound.

As evolution progresses, qualities of the initial population are preserved but are increasingly transformed through reproduction and mutation as the population takes on properties of

the target waveform. The similarity to the target waveform depends on the type of mutation used, on the probability of mutation, and on the amount of time over which evolution occurs.

4.2 Effects of Mutation on Output

Each type of mutation has a characteristic sound that can be readily heard if a population evolves with only that type of mutation. Amplification changes the population in two ways. The amplitude envelopes of individuals in the population tend towards the amplitude envelope of the target environment. Portions of individuals that are in phase with the waveform will be amplified while portions that are out of phase will be attenuated. Exponentiation is very similar to amplification in its behavior, but it is much more invasive; it significantly alters the timbre of the waveform.

The quality of the editing mutations depends largely on the number of neighboring genes grouped for mutation. Application to large segments of the waveform leaves the waveform more intact and recognizable but is less likely to add significantly to the fitness of the population. Given a population of individuals that are several seconds long, only one or two lengthy mutations may propagate to future generations. When very small segments of a waveform are mutated, the sounds are not as recognizable (although they retain many of their original properties), but they are more likely to have a positive effect on fitness and be incorporated in the population. This demonstrates the known property of genetic algorithms—high amounts of mutation tend to have a negative impact on the population while small amounts of mutation over longer periods of time tend to contribute positively to fitness.

When the biologically inspired mutations are applied to perceptibly large segments of a waveform, the function itself can be clearly identified. That is, the listener can tell that a segment of a waveform has been reversed, removed, swapped with another waveform, or repeated. When the grain size is fairly small, portions of the waveform tend to get shuffled around to more closely resemble the target waveform. Portions of a waveform that have been reversed tend to retain some quality that tells the listener that reversal is taking place, but the only biologically inspired mutation that has a significant fingerprint when applied to small segments of a waveform is repetition. Repetition creates pitch out of noisy segments of a waveform. When the grain size is small and the probability of mutation is high, repetition is the most effective mutation at getting the population to denature itself to the point where the target environment can be identified (e.g. a listener unfamiliar with the target environment can identify the environment as a bell when listening to the evolution of

a population of waveforms evolving with a bell as the target environment).¹

4.3 Achieving Musical Results

Higher probability of mutation leads to quicker convergence towards the target environment. If the probability is too high, however, the population can overshoot minima in the error function and fall into regions of the space in which all resemblance to the initial population is lost, the population begins to sound *less* like the target waveform, and the graininess resulting from overlapping mutations dominates.

Because the goal here is to make interesting music, rather than attain a duplicate of some target sound file, I chose fairly small mutation probabilities, and I chose to apply mutations to fairly large segments of waveforms. This allows the sounds to be quite recognizable, despite several minutes² of evolution. The migration of individual waveforms from one environment to another and the ability of environments to change over time significantly contributes to the musicality of the output. It was important to choose probabilities for both migration and environmental change that caused the trajectory of the piece to change every couple of minutes. This kept the piece from losing its biodiversity by staying in one place long enough to lose all sounds but the offspring of a few individuals.

Over the course of a four-minute piece,³ sounds from the initial population slowly evolve. Rhythms change gradually; different sounds from the initial population rise to prominence at different points; and the piece, although slowly changing, has clear directionality.

References

- Burton, A. R. and T. Vladimirova (1999, Winter). Generation of musical sequences with genetic technique. *Computer Music Journal* 23(4), 59–73.
- Holland, J. H. (1992). *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, Mass.: MIT Press.
- Magnus, C. (2003). Evolving waveforms with genetic algorithms. Master's thesis, University of California, San Diego. <http://crca.ucsd.edu/~cmagnus/research.html>.

¹See http://crca.ucsd.edu/~cmagnus/ga_results.html for sample output.

²Several minutes of output—this is a machine learning algorithm so getting several minutes of output takes several hours of computation.

³<http://crca.ucsd.edu/~cmagnus/audio/run5.mp3>