Using Statistical Models and Evolutionary Algorithms in Algorithmic Music Composition

Ritesh Ajoodha
School of Computer Science, University of the Witwatersrand, South Africa

Richard Klein
School of Computer Science, University of the Witwatersrand, South Africa

Marija Jakovljevic
School of Computing, University of South Africa, Pretoria, South Africa

INTRODUCTION

In composition theory, many choices that are drawn from the artist's creativity. These choices are influenced by the rules of composition theory and by the artist's personal perceptions or inspirations. For this reason, a computer would not be able to satisfy every artist's perception into one musical piece as each composer's perceptions are secret and unique. The algorithmic components, however, can be achieved by satisfying components of composition theory.

Composition theory is a branch of music theory which deals with the architecture and construction of music. The composer is commonly a human-agent following the rules of harmony, structure, style, articulation and dynamics. To compose music artificially, one needs to use these rules from composition theory as well as integrate the emotional intelligence that the human composer provides. Currently the human-intelligence that produces emotion cannot be achieved in a computer-agent (Beukes, 2011) which makes the latter impossible. This leads us to believe that a computer-agent might be able to mimic the procedural aspects of composition but will fail when it tries to include the emotional aspects. This research does not attempt to mimic the emotional traits found in human composers.

If one is to ignore emotional traits, it becomes easy to develop music strings through randomization. A way to achieve this is by a statistical model: A statistical model is a structure that produces a sample based on a *probability of occurrence* (PO) in each instance of production. The PO can be attached to each production in a context-free grammar (CFG). By doing this, any string can be derived from the CFG with a user defined occurrence. This CFG will be able to produce a very extravagant music string selection which is why it has become necessary to channel a workable sample from the selection statistically.

It was found that, while the statistical model followed the rules of composition correctly, the CFG primarily produced music strings that sounded computer generated rather than natural. A Genetic algorithm (GA) can be used to encode the softer rules of composition to allow for a more natural sound. A GA is a search heuristic that mimics the process of natural evolution. The heuristic is routinely used to generate useful solutions to optimization and search problems (Melanie, 1996). This work uses a GA to refine the statistical sample into a more natural, aesthetically pleasing sound. After implementation of the described system the results were assessed using a Turing test.

There are five sections in this article: The Background, where a literature review is presented outlining significant accomplishments in the field; The Research Design and Methodology, where the scientific methodology is outlined; Results and Recommendations, where the implemented system and generated music is reviewed; Future Research Directions, where recommended studies are presented; and finally, the Conclusion, where the significant aspects of the proposed methodologies are evaluated.

DOI: 10.4018/978-1-4666-5888-2.ch597

BACKGROUND

The Statistical Model

Conklin (2004) reviewed the process of music generation and equated it with the problem of sampling from a statistical model. One can represent a piece of music as a chain of events, which consist of music objects (e.g. notes) together with a duration and an onset time. A statistical model captures the probabilities of different musical features in a piece, given data such the genre and style. For example, in a Rock song, one is likely to see a 4/4 time signature, 5th intervals and it is rare to see notes that aren't in the current scale. In a Jazz song however, one is likely to see the notes appearing from various modes with a variety of different time-signatures and chromatic runs. To generate music from a statistical model, one samples these different features with frequencies appropriate to the desired style. Conklin (2004) pointed out that statistical models can be beneficial but only a few sampling methods have been explored in the music generation literature.

The Evolutionary Model

Matic (2009) reviewed that composing, as well as, any other artistic activity includes free choice by which a composer express their feelings, moods, intentions or inspiration. He maintains that these choices are seen as a series of instructions that can be relatively easy to interpret. Most composers apply certain rules and instructions when composing and thus any composing process in some way can be considered an algorithm. On the other hand, the absence of human factors in the automatic composition will lead to the appearance of large amounts of objectively bad and useless music as a result of bad computational selected choices. This is evident in music produced by our statistical model alone. The combination of genetic operators such as mutations, selections and crossovers in some way simulates this creative process. As in human-composing these operations allow for continuous "refinement" to attempt to humanise the resulting music. The evaluation function in Equation 1 was used by Matic (2009).

$$\text{Total Fitness} = f + g = \alpha\sum_{i=1}^{m}\vartheta_i\left(\mu_i - a_i\right) + \beta\sum_{i=0}^{m}\rho_i\left(\sigma_i^2 - b_i^2\right) + \gamma\,\frac{1}{bl} \quad (1)$$

where $\sum_{i=1}^{m}\vartheta_i\left(\mu_i - a_i\right)$ and $\sum_{i=1}^{m}\rho_i\left(\sigma_i^2 - b_i^2\right)$ represent the similarity of the two music strings; $\frac{1}{bl}$ represent the collection of good tones; and α, β, and γ served as global weight factors. The experimental results that were produced by Matic (2009) meet some objective criteria of "good" compositions: According to Matic (2009), they contain intervals that are pleasant to the human ear, the rhythm is meaningful and, with a slight adjustment to the appropriate arrangement, the compositions sound unusual but pleasant. Most of the work by Matic (2009) include work on modified genetic operations that can be traced back to George and Wiggins (1998); Horowits (1994); Burton (1996); Brown (2002) and Moroni et al. (2000).

George and Wiggins (1998) reviewed that genetic algorithms have proven to be very efficient search methods, especially when dealing with problems that have large search spaces. This, coupled with their ability to provide multiple solutions, which is often what is needed in creative domains, makes them a good choice for a search engine in a musical application. George and Wiggins [1998] indicate that their GA exhibits the following three significant characteristics which are uncommon in GA applications to music:

- An algorithmic fitness function for objective evaluation of the GA results;
- Problem-dependent genetic operators; and
- A symbolic representation of the structures and the data which helped them solve their problems.

The main limitation of George and Wiggins (1998) is that the music pieces produced had no music structure and as a result there was no room for time signatures which are important in music analysis (Alison, 2006).

Alfonseca, Cebrian, & Ortega (2007) proposed interval distance as a fitness function to generate music in a given style. The main limitation of Alfonseca et al. (2007) is that their genetic algorithm still needed defined parameters to code for the proposed genre and although the authors introduced some information about note duration in the genetic process, it had been ignored by the evaluation function. Some of the pieces of music thus generated recall the style of well-known