Expert Systems with Applications 42 (2015) 7424–7435

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Generating structured music for bagana using quality metrics based on Markov models

D. Herremans [a,*], S. Weisser [b], K. Sörensen [a], D. Conklin [c,d]

[a] ANT/OR, University of Antwerp Operations Research Group, Antwerp, Belgium
[b] Université Libre de Bruxelles, Brussels, Belgium
[c] Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, San Sebastián, Spain
[d] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

ARTICLE INFO

Article history:
Available online 9 June 2015

Keywords:
Markov models
Markov processes
Metaheuristics
Music
Bagana
Computer aided composition (CAC)
Variable neighborhood search (VNS)
Combinatorial optimization

ABSTRACT

In this research, a system is built that generates bagana music, a traditional lyre from Ethiopia, based on a first order Markov model. Due to the size of many datasets it is often only possible to get rich and reliable statistics for low order models, yet these do not handle structure very well and their output is often very repetitive. A first contribution of this paper is to propose a method that allows the enforcement of structure and repetition within music, thus handling long term coherence with a first order model. The second goal of this research is to explain and propose different ways in which low order Markov models can be used to build quality assessment metrics for an optimization algorithm. These are then implemented in a variable neighborhood search algorithm that generates bagana music. The results are examined and thoroughly evaluated.

1. Introduction

Music generation systems can be categorized into two main groups. On the one hand are the *probabilistic methods* (Xenakis, 1992; Conklin & Witten, 1995; Allan & Williams, 2005), and on the other hand are *optimization methods* such as constraint satisfaction (Truchet & Codognet, 2004) and metaheuristics such as evolutionary algorithms (Horner & Goldberg, 1991; Towsey, Brown, Wright, & Diederich, 2001), ant colony optimization (Geis & Middendorf, 2007) and variable neighborhood search (VNS) (Herremans & Sorensen, 2013). The first group considers the solution space as a probability distribution, while the latter optimizes an objective function on a solution space. In this paper, we aim to bridge the gap between those approaches that consider music generation as an optimization system and those that generate based on a statistical model.

The advantage of composing music with optimization techniques is that they offer a way to impose structural constraints. The problem of automatically detecting structure and patterns in music has gained some attention, but remains a difficult task to solve (Conklin & Anagnostopoulou, 2001; Meredith, Lemström, & Wiggins, 2002; Collins, 2011). In this research we start from a given template structure and develop an efficient way of enforcing this structure with a variable neighborhood search algorithm (see Section 2).

The main challenge when using an optimization system to compose music is how to determine the quality of the generated music. Some systems let a *human listener* specify how "good" the solution is on each iteration (Horowitz, 1994). GenJam, a system that composes monophonic jazz fragments given a chord progression, uses this approach (Biles, 2003). This type of objective function considerably slows down the algorithms (Tokui & Iba, 2000) and is known in the literature as the human fitness bottleneck.

Most automatic composition systems avoid this bottleneck by implementing an automatically calculated objective function based on either existing *rules from music theory* or by *learning from a corpus* of existing music. The first strategy has been used in compositional systems such as those of Geis and Middendorf (2007), Assayag, Rueda, Laurson, Agon, and Delerue (1999) and Herremans and Sörensen (2013). Although every musical genre has its own rules, these are usually not explicitly available, which imposes huge limits on the applicability of this approach (Moore, 2001). This problem is overcome when style rules can be learned automatically from existing music, as is done in this research. This approach is more robust and expandable to other styles.

Markov models have been applied in a musical context, for learning from a corpus, for a long time. The string quartet called

---

* Corresponding author.

E-mail addresses: dorien.herremans@uantwerpen.be (D. Herremans), stephanie.Weisser@ulb.ac.be (S. Weisser), kenneth.sorensen@uantwerpen.be (K. Sörensen), darrell.conklin@ehu.es (D. Conklin).

http://dx.doi.org/10.1016/j.eswa.2015.05.043
0957-4174/© 2015 Elsevier Ltd. All rights reserved.

the Illiac Suite was composed by Hiller and Isaacson in 1957 by using a rule based system that included probability distributions and Markov processes (see Sandred, Laurson, & Kuuskankare, 2009, for a recent overview of this work). Pinkerton (1956) learned first order Markov models based on pitches from a corpus of 39 simple nursery rhyme melodies, and used them to generate new melodies using a random walk method. Fred and Carolyn Attneave generated two perfectly convincing cowboy songs by performing a backward random walk on a first order transition matrix, as reported by Cohen (1962). Brooks, Hopkins, Neumann, and Wright (1957) learned models up to order 8 from a corpus of 37 hymn tunes. A random process was used to synthesise new melodies from these models.

An interesting conclusion from this early work is that high order models tend to repeat a large part of the original corpus and that low order models seem very random. This conclusion was later supported by other researchers such as Moorer (1972), who states: "When higher order methods are used, we get back fragments of the pieces that were put in, even entire exact repetitions. When lower orders are used, we get little meaningful information out". These conclusions are based on a heuristic method whereby the pitch is still chosen based on its probability, but only accepted or not based on several heuristics which filter out, for instance, long sequences of non-tonic chords that might otherwise sound dull. Music compositions systems based on Markov chains need to find a balance in the order to use.

Other music generation research with Markov includes the work of Tipei (1975), who integrates Markov models in a larger compositional model. Xenakis (1992) uses Markov models to control the order of musical sections in his composition "Analogique A". Conklin and Witten (1995) present the multiple viewpoint method and apply it to probabilistic generation of chorale melodies. Markov models also form the basis for some real-time improvisation systems (Dubnov, Assayag, Lartillot, & Bejerano, 2003; Pachet, 2003; Assayag & Dubnov, 2004). Some more recent work involves the use of constraints for music generation using Markov models (Pachet & Roy, 2011). Allan and Williams (2005) trained hidden Markov models for harmonizing Bach chorales. Whorley, Wiggins, Rhodes, and Pearce (2013), Whorley and Conklin (2015) applied a Markov model based on the multiple viewpoint method to generate four-part harmonisations with random walk. A more complete overview of Markov models for music composition is given by Fernández and Vico (2013).

In an early survey of statistical models for music generation, Conklin (2003) highlighted the need for approaches to conserve structural patterns during generation, in order to effectively ensure intra-opus repetition. Collins, Laney, Willis, and Garthwaite (2015) implemented this idea. They did not consider optimization to generate a solution, but used only a single random walk. The contributions of this research are twofold. First, we propose a method to generate music while conserving structural patterns during generation. Secondly, we propose and evaluate different ways in which machine learned models can be used to build quality evaluation metrics. To this end, a first order Markov model is built that quantifies note transition probabilities from a corpus of bagana music, a traditional lyre from Ethiopia. This model is then used to evaluate music with a certain repetition structure, generated by an optimization procedure previously developed by the authors (Herremans & Sörensen, 2012). Due to the size of many available corpora of music, including the bagana corpus used in this research, rich and reliable statistics are often only available for low order Markov models. Since these models do not handle structure and can produce very repetitive output, a method is proposed for handling long term coherence with a first order model. This method will also allow us to efficiently calculate the objective function, by using the minimal number of necessary note intervals as possible while still containing all information about the piece. Secondly, this paper will critically evaluate how Markov models can be used to construct evaluation metrics in an optimization context. In the next section more information is given about bagana music, followed by an explanation of the technique employed to generate repeated and cyclic patterns. An overview of the different methods by which a Markov model can be converted into an objective function are discussed in Section 3. Variable neighborhood search, the optimization method used to generate bagana music, is then explained. An experiment is set up and the different evaluation metrics are compared in Section 5.

2. Structure and repetition in bagana music

The bagana is a ten-stringed box-lyre played by the Amhara, inhabitants of the Central and Northern part of Ethiopia. It is an intimate instrument, only accompanied by a singing voice, which is used to perform spiritual music. It is the only melodic instrument played exclusively for religious purposes (Weisser, 2012). The bagana melody and singing voice are quasi homophonic, meaning that the voice and bagana usually follow each other in unison (Weisser, 2005). In this research the focus is on analyzing and generating the instrumental part.

The bagana is made of wooden pillars and soundbox, equipped with ten cattle gut strings. The strings are plucked with the left hand and four strings are used as finger rests. It is tuned to an Amhara traditional pentatonic scale. Each finger of the left hand is assigned to one string (see Fig. 1), except in the case of the index finger (referred to as finger 2 and 2′ in the figure), which plays two equally tuned strings. This allows us to make abstraction from the actual pitch and work with the corpus made by Conklin and Weisser (2014) based on finger numbers (see Section 5).

Bagana songs are typically very repetitive with a recognisable overall structure (Weisser, 2006). This repetition is intentional since repetitive music has a strong influence on the state of consciousness among musical traditions. Even Western-trained listeners describe the sounds as "becoming meditative objects, relaxing the mind" (Dennis, 1974).

Two example bagana pieces, including finger numberings, are displayed in Fig. 2 and 3. Both pieces consist of two sections, and

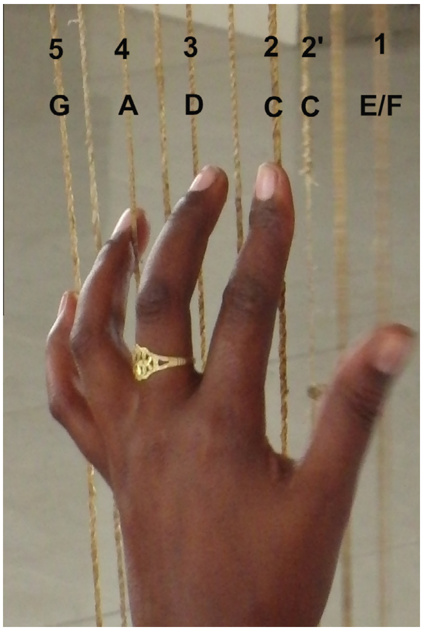

Fig. 1. Assignment of fingers to strings on the bagana.