Music generation with structural constraints: an operations research approach*

Dorien Herremans
Queen Mary University of London, Centre for Digital Music
e-mail: dherremans@qmul.ac.uk

Elaine Chew
Queen Mary University of London, Centre for Digital Music
e-mail: elainechew@qmul.ac.uk

Music generation systems have attracted research attention since the advent of computing. They have become increasingly important, bolstered by rising global expenditure on digital music, which was over 64 billion USD in 2014 alone [2].

Most music generation systems are based on statistical models and rules. A drawback of these early systems is their inability to synthesize music that possess global structure. When music does not have a clear direction or long-term coherence, it fails to hold the listener's attention and can be hard to follow. The computing power available to us today provides us with new opportunities to generate good sounding music with structure.

This talk will review state systems for generating music with structure, and describe an optimisation-based approach to music generation. The problem of structure has recently been tackled using deep learning, with mixed results. Guaranteed success has been achieved using optimisation algorithms that constrain the structure of the generated music. This approach is currently being examined in the authors' EU project MorpheuS[1].

Music generation as an optimization problem

When composing music is redefined as a combinatorial optimization problem, constraints can be used to enforce a larger temporal structure to a piece. For example, a solution could be a list of pitches constrained to a rhythmic template [4].

Measuring music quality An optimization-based music generation system needs a measure of quality of the output. What makes a piece of music sound good? This is a difficult and highly style dependent question. Three main approaches exist. The first one uses human feedback to get an evaluation score. While this might be the most intuitive approach, the human factor produces an enormous bottleneck. Secondly, rules from music theory can be quantified. This approach is only valid for a limited number of music styles that have well-studied

*Presented at the 30th Annual Conference of the Belgian Operational Research (OR) Society (ORBEL30), Louvain-La-Neuve, Jan 27-28, 2016

[1] http://cordis.europa.eu/project/rcn/195685_en.html

formal rules such as counterpoint. A more robust approach is to use machine learning techniques to capture the style of a corpus or individual piece. Multiple ways of converting a Markov model into an objective function have been explored by [4]. This last strategy is most versatile when generating any kind of music.

Local search A proven strategy for solving a monophonic music generation problem with constrained structure is variable neighbourhood search (VNS). Starting from a piece consisting of random pitches, three move types—change one pitch; change two pitches; swap two pitches—together with a perturbation move are used by [3] to navigate out of local optima. The algorithm outperforms a genetic algorithm approach.

Applications involving structure

Defining the music generation problem as one of optimization allows one the freedom to impose both hard and soft constraints. These constraints can be used to define different types of structure.

The tension profile serves as an example of narrative structure. The authors have designed a way to capture aspects of tonal tension based on the spiral array, a three dimensional model for tonality [1]. These tension profiles form one type of long term structure that could be implemented as a hard constraint in the existing algorithm.

Music with a particular tension profile is especially relevant to game music. Currently, music in computer games consists primarily of concatenated audio files joined together by cross-fading. When a user enters a danger/alert situation, tension levels rise and the music switches to another audio file. This could be done more smoothly by generating music on the fly with varying levels of tension.

Film music in another example of music that follows a structural narrative. The tension profile constraint approach can be used to create copyright-free background music for YouTube videos or stock music that follows the emotional narrative of the movie.

OR techniques in other domains of music research

Vast opportunities exist for applying techniques from the field of OR to music related problems far beyond music generation. Potential applications include instrument (for example, piano or guitar) fingering, genre classification, hit song prediction, automatic music transcription, multiple alignment, and many more.

Acknowledgments This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 658914.

References

[1] E. Chew. *Mathematical and Computational Modeling of Tonality: Theory and Applications*, volume 204. Springer, New York, 2014.

[2] McKinsey & Company. Global media report 2015: Global industry overview. 2010.

[3] D. Herremans and K. Sörensen. Composing first species counterpoint with a variable neighbourhood search algorithm. *Journal of Mathematics and the Arts*, 6(4):169–189, 2012.

[4] D. Herremans, S. Weisser, K. Sörensen, and D. Conklin. Generating structured music for bagana using quality metrics based on markov models. *Expert Systems with Applications*, 42(21):7424–7435, 2015.