

Submitted exclusively to the *Journal of Mathematics and Music*
Last compiled on October 18, 2016

EDITORIAL

Machine Learning and Music Generation

José M. Iñesta^{a*}, Darrell Conklin^{b,c}, and Rafael Ramírez^d

^a*Department of Languages and Computer Systems, University of Alicante, Alicante, Spain;*

^b*Department of Computer Science and Artificial Intelligence, University of the Basque Country, San Sebastián, Spain;*

^c*IKERBASQUE, Basque Foundation for Science, Bilbao, Spain;*

^d*Music Technology Group, Pompeu Fabra University, Barcelona, Spain*

(Received 00 Month 20XX; final version received 00 Month 20XX)

This Special Issue on Machine Learning and Music Generation features rigorously refereed papers based on selected extended abstracts presented at the 8th International Workshop on Machine Learning and Music, held in Vancouver, Canada, in August 2015 (MML2015). The workshop was part of the 21st International Symposium on Electronic Art (ISEA2015).

The annual Machine Learning and Music workshops have established themselves as an active forum bringing together researchers in the area of machine learning applied to music, in order to respond to opportunities and challenges arising from the growth of digital musical data and developments in machine learning and data mining. The MML2015 workshop followed the tradition of previous workshops in Helsinki (Finland, MML2008), Bled (Slovenia, MML2009), Florence (Italy, MML2010), Athens (Greece, MML2011), Edinburgh (Scotland, MML2012), Prague (Czech Republic, MML2013), and Barcelona (Spain, MML2014). The latest edition was organized in conjunction with the International Symposium on Electronic Art (ISEA), an interdisciplinary event dedicated to creative applications of new technologies in art, interactivity, and electronic and digital media. Therefore MML2015 particularly welcomed contributions on machine learning for music generation and creation.

Computational approaches to music composition and style imitation have engaged musicians, music scholars, and computer scientists since the early days of computing. Music generation research has generally employed one of two strategies: knowledge-based methods which model style through explicitly formalized rules, and data mining methods which apply machine learning to induce statistical models of musical style.

The five papers in this Special Issue illustrate the range of tasks and design choices in current music generation research applying machine learning techniques and highlighting recurring research issues such as training data, music representation, candidate generation, and evaluation. The papers focus on different aspects of modeling and generating music, including melody (Ponce de León et al. 2016), chord sequences (Conklin 2016), ornamentation (Giraldo and Ramirez 2016), and dynamics (Kosta et al. 2016). Models are induced from audio data (Giraldo and Ramirez 2016; Kosta et al. 2016) or symbolic data (Conklin 2016; Kirilin and Yust 2016; Ponce de León et al. 2016). The papers by

*Corresponding author. Email: inesta@dlsi.ua.es

Giraldo and Ramirez (2016) and Kirlin and Yust (2016) pay particular attention to the relative contribution of different features employed in representing music data. A challenging issue in computational music generation, especially if fully automated, is the determination of concrete musical events which however cohere with certain more abstract or underlying structures, such as metrical regularities (Ponce de León et al. 2016) or both adjacent and distant repetition (Conklin 2016). Machine learning is employed in generating events and event sequences, for example through model-based prediction (Giraldo and Ramirez 2016; Kosta et al. 2016) or sampling from a statistical model (Conklin 2016), and in evaluating candidate event sequences (Ponce de León et al. 2016) or ranking generated sequences (Conklin 2016) based on a statistical model. The contribution by Kirlin and Yust (2016) presents a step towards learning middleground and background regularities, which will further enable future inductive methods to achieve both stylistic and structural coherence of generated music pieces based on models learned from data.

We Guest Editors welcome you to this Special Issue and hope that you find the papers interesting and relevant to your own research.

Chord sequence generation with semiotic patterns

Music derives meaning through repetition. Modeling musical coherence created by internal repetition and reference to earlier material in a piece is a persistent issue in computational music generation, and presents a challenge especially for machine learning methods. The paper by Conklin studies repetition in music, and develops a way to represent and generate repetition using statistical models.

The paper develops a pattern representation which explicitly describes relations between events using variables which are instantiated during generation. A *semiotic pattern* describes a formal language of sequences, and sequences satisfying the pattern are generated by sampling from a statistical model. The statistical model is learned on a corpus of chord sequences and is an instance of the *viewpoint* modeling approach used previously with success for prediction, classification, and music generation.

There are two ways to see this work: as a semiotic pattern representation with pattern instances ranked by a statistical model, or as a statistical modeling approach with patterns used to constrain the space of permissible sequences. Whatever the view, the semiotic patterns, which express long-range dependencies, are essential to achieving coherence in generated sequences.

The method is illustrated by generating chord sequences for trance songs. The results show that semiotic patterns and machine learning of statistical models can work together to support music generation which is both internally and stylistically coherent.

A machine learning approach to ornamentation modeling and synthesis in Jazz guitar

The paper by Giraldo and Ramírez presents a method to render expressive jazz guitar performances of jazz tunes. A machine learning approach is used to learn when a note in a score is ornamented by a guitarist. Ornaments are applied to the notes of a given jazz tune based on the predictions of such an ornamentation classifier. The symbolic performance produced is then rendered using concatenative synthesis.

The paper describes a system that synthesizes a notated jazz melody using ornamentations of notes that are characteristic for a natural musical interpretation. It relies on

a classifier that learns from a dataset of recorded performances which notes to ornament and which notes not to ornament. The recorded performances are transcribed and aligned to their respective scores, resulting in a set of data instances where each note in the score is either played as a single note, or as multiple notes. The latter case is referred to as ornamentation. Each of the score notes is represented by a vector of features conveying information such as pitch and duration, information about neighboring notes, and about more global context (key, tempo, etc.). A binary classifier predicts from the feature vectors whether a score note should be ornamented or not. In the case of a predicted ornamentation, an ornamentation is instantiated by adopting the ornamentation of the nearest ornamented neighbor of the note to be ornamented, in the feature space defined by the training data. In case no ornamentation is predicted, the same procedure is applied, but selecting the nearest non-ornamented neighbor. The results are concatenated and transformed in pitch and time, in order to match the pitch and duration of the current note. The result is an audio file of the performed melody.

The experimental part of the paper runs two feature selection algorithms and tests several classifiers for predicting ornamentation. The results show that using more than the best three features does not improve classification accuracy. The authors also conclude that overfitting is minimal using a subset of the training data, at least for some classifiers.

Analysis of analysis: Using machine learning to evaluate the importance of music parameters for Schenkerian analysis

The paper by **Kirlin and Yust** reports on experiments with learning probabilities for melodic reductions within a system intended to model Schenkerian analysis, directed at discovering the importance of different features for learning probabilities leading to analyses which most closely match the ground truth.

Several methods to approach Schenkerian analysis from a computer can be found in the literature. In general, they try to systematize expert practice in order to make it computable. Those works are usually outlined as a proof of concept and accordingly evaluated with small corpora or even individual short pieces. One of the most difficult impediments to progress in the research of Schenkerian analysis from a computational point of view is the lack of available curated corpora. In previous work, Kirlin introduced a new corpus based on Yust's Maximal Outerplanar Graphs (MOP) representation that begins to fill this gap.

Using that corpus, the authors have recently introduced a technique that tries to solve the analysis of a musical work using a Schenkerian approach. Musical pieces encoded in a symbolic format are described in terms of melodic, harmonic, and metrical features. These eighteen features model note triples, and cover aspects such as the scale degree of notes, the harmony present at the onset of the center note, the interval from the left to the right note, and the metrical strength of the left note compared to the center note.

The main contribution of the article presented in this Special Issue is the use of a machine learning approach to objectively study the influence of the melodic, harmonic, and metrical features for solving a Schenkerian analysis task. Random forests and probabilistic context-free grammars are used in a parsing algorithm that determines the most probable analyses of works in the corpus, which are compared to the manually analyzed works in the corpus.

The result of the study is an ordering of the features by what the authors name *expendability*: how useful is each feature in the context of others. From the results the authors conclude that though all features are significant and necessary, melodic intervals and

melodic orientation to a key are the most essential factors for performing a Schenkerian analysis.

Mapping between dynamic markings and performed loudness: A machine learning approach

Musicians manipulate sound properties such as pitch, timing, loudness, and timbre in an attempt to communicate emotions, clarify musical structure, or shape the music according to her own intentions. Among these properties, loudness manipulation is one of the main actions for expressivity in music performance. In the context of classical music, loudness specifications are frequently notated in music scores as dynamic markings. However, the meaning of such markings is in general ambiguous and depends on the context in which the markings appear.

The paper by **Kosta, Ramirez, Bandtlow, and Chew** investigates the relationship between dynamic markings in music scores and performed loudness by applying machine-learning techniques to induce predictive models of loudness levels corresponding to dynamic markings, and to classify dynamic markings given loudness values. The predictive models are obtained by mining a data set consisting of forty-four recordings of performances of Chopin's Mazurkas each by eight pianists. The results reported in the paper show that loudness values and markings can be predicted when trained across recordings of the same piece, but this is not the case when trained across the pianist's recordings of other pieces. This demonstrates that score features seem to be more important than individual style when modeling loudness choices. The results also show that for some Mazurkas in the data set loudness values can be predicted more easily than for others. This result is related to the range of the different markings that appear in a piece as well as their position and function in the score in relation to structurally important elements.

Data-based melody generation through multi-objective genetic algorithms

The paper by **Ponce de León, Iñesta, Calvo-Zaragoza, and Rizo** studies the role of statistical and structural descriptors in a genetic system for melody composition. Genetic algorithms are an optimization technique that mimics the biological evolution of living beings and natural selection. A population of individuals, representing different possible solutions, are subjected to crossovers and mutations, and a selection stage decides which individuals might best solve the problem. Those individuals are allowed to procreate a new generation of, supposedly, better individuals until convergence. There are two main difficulties here: the choice of a proper representation of the problem as a "chromosome", and the implementation of the selection process.

In this work, genetic operators are adapted to the chromosomic representation of melodies as *melodic trees*, based on the rhythm structure as a tree. Each bar is represented as a sub-tree, depending on its meter. The level of a node determines its duration: the root represents the duration of the whole melody, the nodes of the next level represent a division of the upper node that, together, sum up to its duration. Pitch codes are found in the leaves of the tree and any kind of absolute or relative pitch can be used.

The other issue to solve is how to compute the "fitness" of each individual: how every solution is able to actually *solve* the problem at hand, in this case to have an acceptable music composition. The paper uses a combination of different fitness functions by means

of a multi-objective optimization method. This is able to combine multiple and diverse values in order to rank the individual in a single fitness space, based on statistical and structural descriptors of the melodies represented by the trees. A number of machine learning techniques are used to rate evolutionary generated music according to the parameters tuned from melodies that supposedly share some common properties such as genre, style, and mood. The authors present a graphical interface prototype that allows one to explore the creative capabilities of the proposed system.

Acknowledgements

The Guest Editors point out that this Special Issue would not have been possible without the help and dedication of many experts in the areas of music, mathematics, and machine learning, who provided detailed anonymous reviews of the manuscripts submitted to this Special Issue. We thank the anonymous referees for their rigorous work, and we thank all authors for their very professional work during the intense writing and revision processes. Special thanks are given to Thomas Fiore who enthusiastically supported this Special Issue since its inception and provided detailed editorial comments on all of the manuscripts. We would also like to thank Moreno Andreatta for supporting the Special Issue when it was first proposed, and Kerstin Neubarth for assistance with this editorial article.

Funding

The Guest Editors acknowledge the financial support of the following funding sources: the Ministerio de Economía y Competitividad project TIMuL (No. TIN2013-48152-C2-1-R, and No. TIN2013-48152-C2-2-R, supported by UE FEDER funds); the project Lrn2Cre8 which is funded by the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission under FET grant number 610859; and the European Union Horizon 2020 research and innovation programme under grant agreement number 688269.

References

- Conklin, Darrell. 2016. "Chord sequence generation with semiotic patterns." *Journal of Mathematics and Music* 10 (3): XX-XX.
- Giraldo, Sergio, and Rafael Ramirez. 2016. "A machine learning approach to ornamentation modeling and synthesis in Jazz guitar." *Journal of Mathematics and Music* 10 (3): XX-XX.
- Kirlin, Phillip B., and Jason Yust. 2016. "Analysis of analysis: Using machine learning to evaluate the importance of music parameters for Schenkerian analysis." *Journal of Mathematics and Music* 10 (3): XX-XX.
- Kosta, Katerina, Rafael Ramirez, Oscar F. Bandtlow, and Elaine Chew. 2016. "Mapping between dynamic markings and performed loudness: A machine learning approach." *Journal of Mathematics and Music* 10 (3): XX-XX.
- Ponce de León, Pedro J., José M. Iñesta, Jorge Calvo-Zaragoza, and David Rizo. 2016. "Data-based melody generation through multi-objective evolutionary computation." *Journal of Mathematics and Music* 10 (3): XX-XX.