Cognitively motivated representations of symbolic music

Johanna Devaney
Ohio State University

CogMIR™

Introduction

- This talk will discuss the early stages of a project that uses a top-down approach to develop a robust representation(s) of symbolic music data
- Such a representation would be useful for a number of analytic applications, including
  - harmonic analysis
  - provenance assignment
  - similarity analysis

Goal and Motivation

Goal and Motivation

(Some) Existing Approaches

Goal and Motivation

(Some) Existing Approaches

Predicting Phrase Labels

Goal and Motivation

(Some) Existing Approaches

Predicting Phrase Labels

Next Steps

Goal and Motivation

(Some) Existing Approaches

Predicting Phrase Labels

Next Steps

Conclusions

Ultimate Goal

- To develop a representation for western art music scores that is
  - cognitively meaningful
  - computationally accessible
  - able to capture temporal relationships
  - able to facilitate the analysis of multiple levels of musical structure

Motivation

- Existing representations have shown limited success for harmonic and higher-level musical analysis
- Multi-level model employed in music theory pedagogy may be useful in developing new representations

Motivation

A pedagogical approach

Formal Structure	↑	*Higher*
Phrase-level Function		
Roman Numerals (Harmonic Analysis)		
Notes (Musical Surface)	↓	*Lower*

Motivation

Speech recognition	↔	**Higher-level music analysis**
“Language model”	↔	Functional analysis
Phonemes	↔	Harmonic analysis
MFCCs and representations	↔	*Representation*
Acoustic signal	↔	Musical surface

(Some) Existing Approaches

- Piano-roll notation
  - David Temperley's preference rules (2001)
- Vertical slices of polyphonic music to form "chords"
  - Quinn and Mavromatis (2011) with extension to incorporate voice leading
  - Chris White's chord-based function model
- Martin Rohmeier's scale degree, function, and phrase-level representations (2011)

(Some) Existing Approaches

- Large-scale analysis approaches are heavily influenced by text retrieval methods, namely N-grams (sequences of N contiguous symbols)
- N-grams work well for
  - melody retrieval in monophonic contexts (Pickens 2001)
  - chord retrieval in polyphonic contexts when the chords occur as distinct units (Scholz et al. 2009)
    - e.g., peachnote.com's N-gram viewer (Viro 2011)

(Some) Existing Approaches

- N-gram representations encounter problems with more complex textures
  - e.g., where the notes of chords are not played simultaneously
- N-grams do not distinguish between what is structurally significant and what is not
  - e.g., between what is a chord-tone and what is a non-chord tone/ornamentation.

“Language Model” for Music

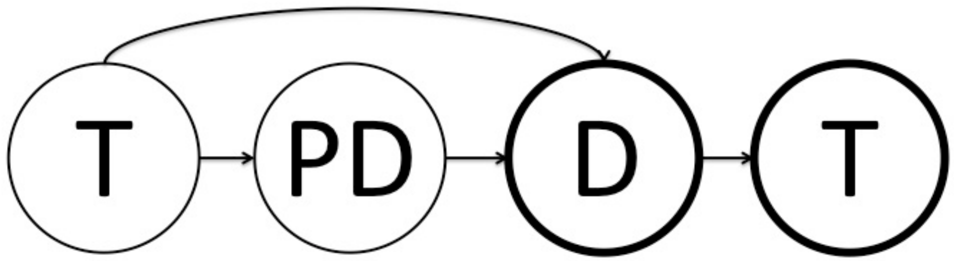
- Harmonic/Roman numeral analysis is a hard task when working from the musical surface
- A model of phrase-level function and its relationship to roman numeral labels (the “language model”) facilitates a top-down approach
    - This approach also allows for
        - the evaluation of existing representations
        - the development of learned representations

Musical Phrases

- Complete musical statements built from the ordered presentation of three harmonic functions
  - tonic, pre-dominant, and dominant functions
- End with a cadence
  - remaining on the dominant function for a half cadence
  - returning to the tonic function for an authentic or a deceptive cadence

Predicting Phrase Labels

- Hidden Markov model (HMM)
  - State space



- Encoded chord label and function for all examples from Laitz's textbook The Complete Musician (2011) into digital representation
  - used to train HMM transition probabilities and evaluate model

This part of the project was undertaken with Daniel Shananan

Predicting Phrase Labels

- Evaluation
  - 80/20 split of textbook data: 94.3% overall accuracy
    - tonic: 93%, pre-dominant: 93%, dominant: 89%
  - Workbook exercises: 87.5% overall accuracy
    - tonic: 96%, pre-dominant: 83%, *dominant: 52%***

*** 36% of phrases ended with a dominant function in the workbook versus 17% in the training set*

This part of the project was undertaken with Daniel Shananan

Next Steps

- Experiment with both hierarchical hidden Markov models (HHMM) and conditional random fields (CRF)
- HHMM implementation
  - 3-state phrase-level function HMM will sit on top of a lower-level roman numeral HMM
- CRF implementation
  - each state will have a roman numeral and a function
  - advantage over HHMMs in this context: ability to utilize temporal context, including metrical position and features from previous time frames

Next Steps

- Once the "language model" has been implemented different existing representations will be evaluated along with learned representations
- A successful representation must be able to capture
  - information about note duration and metrical strength across different time scales
  - temporal relationships between pitch classes and chords

Conclusions

- This talk has described an ongoing project for developing a multi-level model for western art music analysis
- The model is inspired by
  - music theory pedagogy
  - speech recognition
- The ultimate goal of this project is to develop a representation of the musical surface that is both utilitarian and related to human musical information processing

References

Laitz, S. 2011. The Complete Musician, Oxford University Press, Oxford

Pickins, J. 2001. A survey of feature selection techniques for music information retrieval. In *Proceedings of the International Symposium for Music Information Retrieval.*

Quinn, I. and P. Mavromatis. 2011. Voice Leading and Harmonic Function in Two Chorale Corpora. In *Proceedings of Mathematics and Computation in Music Conference*.

Rohrmeier, M. 2011. Towards a generative syntax of tonal harmony. Journal of Mathematics and Music. 5 (1): 35–53.

Scholz, R., E. Vincent, and F. Bimbot. 2009. Robust modeling of musical chord sequences using probabilistic N-grams. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing.*

Temperley, D. 2001. The Cognition of Basic Musical Structures. Cambridge, MA, MIT Press.

Viro, V. 2011. Peachnote: Music score search and analysis platform. In *Proceedings of the International Society for Music Information Retrieval Conference*.