Melody Generation: Rule Based vs Hidden Markov Models

Michael Finley, Gavin Baker
ECE477: Computer Audition
14 December 2019

Introduction

- Melody generation is a field of research that hasn't explored certain models fully, e.g. Hidden Markov Models (HMMs)

Motivation

- We believe the success of the tri-gram language model can be explored in the context of melody generation while also considering a rule based model

Rule Based Model

- Progressive Embellishment Approach
- Iterates over inputted melody, outputting more intricate melody each iteration
- Melody notes outputted as tuples [MIDI Number, Note Length]
- Parameters:
  - Intricacy
    - Number of embellishing iterations
  - Chordal Complexity
    - Number of chords in base chord progression
  - Scale/Tonal Center
    - Facilitates conversion of scale degrees produced by the model to actual to actual notes

Implementation Details

For 1:Complexity:
  Generate whole note s.t. no repeated notes, last in sequence is IV or V

For 1:Intricacy - 1:
  for each note in current sequence
    with some probability P:
      replace note with two evenly divided notes of different pitches

Replace outputted notes with MIDI values based on Key and Scale

Hidden Markov Models

1st-order matrix

Note	A	C♯	E♭
A	0.1	0.6	0.3
C♯	0.25	0.05	0.7
E♭	0.7	0.3	0

2nd-order matrix

Notes	A	D	G
AA	0.18	0.6	0.22
AD	0.5	0.5	0
AG	0.15	0.75	0.1
DD	0	0	1
DA	0.25	0	0.75
DG	0.9	0.1	0
GG	0.4	0.4	0.2
GA	0.5	0.25	0.25
GD	1	0	0

Figure: 1^{st} *order and* 2^{nd} *order transition matrices.*

- The Markov property is detailed below in Eqn (1)

$$Pr(X_{n+1} = x | X_1 = x_1, ..., X_n = x_n) = Pr(X_{n+1} = x | X_{n-3} = x_{n-3}, ..., X_{n-1} = x_{n-1}) \quad (1)$$

Implementation Details

- Tri-gram melody model accurately captures melodic phrasing
- Trained on 18,000 midi songs, from the Meertens Tune Collections
- Done in Python for simplicity and access to PyTorch library

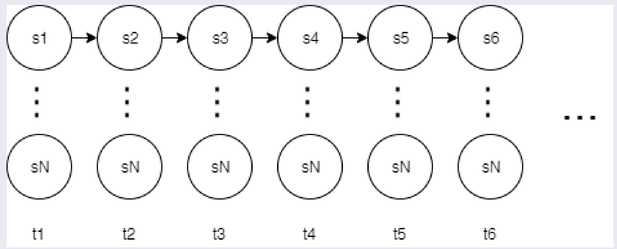

Figure: Concept of sequences used for melody generation.

Dataset- The Meertens Tune Collection (2018)

- 18,618 MIDI tunes
- Sourced from Dutch folk songs, written sources
- Monophonic melody files
- Limited scope of data sources limits model training

Summary

- The tri-gram language model was successfully adapted for use to train a model for melody generation
- Subjective analysis of the results is difficult and requires more time for data collection
- Hybridization of the rule based model and the HMM model requires more sophistication and testing

Acknowledgements

- Thank you to Dr. Duan for your help in this process
- Thank you Narges Mohammadi, Shadi Artipi, Haiqin Yin, and You Zhang for your advice

UNIVERSITY of ROCHESTER