# ChordAL: A Chord-Based Approach for Music Generation using Bi-LSTMs

**Tan Hao Hao**
School of Computer Science and Engineering
Nanyang Technological University, Singapore
tanh0207@e.ntu.edu.sg

## Abstract

In this paper we propose ChordAL, a chord-based generation system which is capable of composing melodies based on a few starting chords. It consists of 3 components: a *chord generator*, a *chord-to-note generator*, and a *music styler*. In our chord generator, our learnt chord embeddings surprisingly unveil the Circle of Fifths, which shows that our model is able to learn hidden musical structure through a simple chord generation task. The chord-to-note generation is treated as a *seq2seq* task like machine translation, and we make use of the chord embeddings learnt previously for melody generation. Evaluation results show that ChordAL performs well in generating fluent harmony due to its harmony-based nature, although its performance on rhythm and structure needs further improvement.

## Introduction

Algorithmic composition of music using deep learning has gained increasing attention in recent years, with a large number of proposed music generation systems in the field. However, not much work has been done on analyzing *chord-based approaches* - particularly, how to compose melodies based on a given chord progression. Yet, it seems to be a very intuitive approach for human composers to write songs based on chord progressions, especially in Western music genres like pop and jazz music.

Here we propose ChordAL, a chord-based generation system which is capable of composing melodies based on a few starting chords. The motivation is to investigate the capability of chord-based approaches in "translating" chord progressions into fluent, harmonic melodies.

## Overview

ChordAL's pipeline can be divided into 3-step:

- **Chord Generator**: First, a chord progression is generated from scratch given a few chords as starting seed.
- **Chord-to-Note Generator**: Next, the sequence of chords is fed into the Chord-to-Note Generator, that generates the melody line based on the given chords.
- **Music Styler**: Finally, both the generated chord and melody are fed into the Music Styler component for post-processing and styling, in order to combine both parts into a presentable piece.

### Chord Generator

Chords are first encoded as 24 different indices, which represents 12 different pitches in an octave with *major* and *minor* chord each. During preprocessing, we decide to remove the repeated chords to ensure that all adjacent chords are different. Generation results show that this greatly helps to increase the variation and quality of the generated chord progressions.

For the model, we pass the chord vectors into a 32-dimension *embedding layer*. The embeddings are then fed into a stacked Bi-LSTM of 2 layers, with 64 hidden neurons each. A dense layer is added for the output with a softmax activation function for multi-class classification to predict the output chord index.

The concept of using *embeddings* is borrowed from the natural language processing domain. Word embeddings are expected to contain meaningful representation of the words in the vocabulary, such as the *word2vec* model. Here, we hope that the embedding layer could learn meaningful representations about the relationships between chords, which will be useful for the chord-to-note generation later.

### Chord-to-Note Generator

We use *piano-roll* representation for the melody, which can be interpreted as a matrix of shape (128, $t$), where $t$ is the length of the piano-roll.

For the model, the chord indices are converted to their corresponding embeddings learnt in Chord Generator, and then fed into a stacked Bi-LSTM of 2 layers with 64 hidden neurons each. Dropout layers of probability 0.2 are added after each Bi-LSTM layer. A *tanh* activation is added after the first dropout layer. A time-distributed dense layer is added for the output, with a softmax activation function for multi-class classification to predict the output note.

### Music Styler

The Music Styler implements additional tuning on the generated melody line, such as removing random short notes to improve the fluency of the song. Then, it applies styling as defined by the user for both the chord and melody parts. We find that *sustaining instruments* (strings, organ, brass, etc.)

sound best on the compositions, as they are mainly chordal in nature.

## Dataset

The following parallel chord-to-note datasets are used:

- **Nottingham dataset**: We use the cleaned version pre-processed by Jukedeck, with separated chord and melody parts for each piece.
- **McGill-Billboard Chord Annotations**: It contains chord annotations for around 1000 Billboard chart songs.
- **CSV leadsheet database**: A leadsheet database that contains around 2200 Western music pieces across different genres including rock, pop, jazz, etc.

All entries are indexed into a *chord database* and a *melody database*. We use the chord database to train the Chord Generator, and both databases to train the Chord-to-Note Generator. Each entry is transposed to all 12 different pitches in the octave for normalization.

## Results and Discussion

**Resemblance of major Circle of Fifths** We extract the chord embeddings for each chord index and visualize them after applying Principal Component Analysis (PCA) to reduce the 32 dimensions into 2. As seen in Figure 1 below, the chords almost form the exact same circle as the major Circle of Fifths, This shows that our embedding layer is capable to learn the underlying structure of music, even with a simple chord generation task.

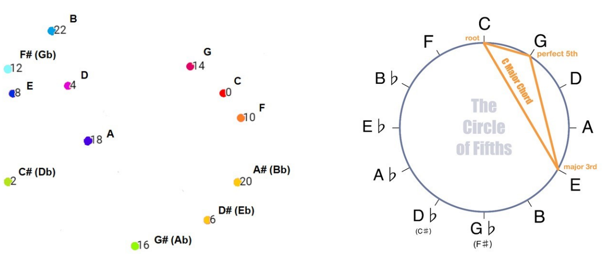


Figure 1: The learnt chord embeddings (left) is found to highly resemble the major Circle of Fifths (right).

**Evaluating generated melodies** Overall, the pieces generated by ChordAL sound pleasant and harmonic. This is accredited to the harmony-based nature of the approach itself, as the network is aimed to learn about the relationship between the notes and the chords. However in general, each melody note sustains for a long duration. One reason may be that the note representation in the piano roll is highly repetitive - sustained notes are encoded as the same one-hot vector across a long duration. This may cause our network to be biased against generating highly repetitive values that result in long notes.

**Subjective Evaluation** A comprehensive subjective evaluation is conducted on 5 of the songs generated by ChordAL, which prompts the respondents to rate its performance on a 5-point Likert scale based on harmony, rhythm and structure. A total of 15 respondents with a higher music proficiency level, which include professional piano players, band performers and music teachers, took part in this comprehensive evaluation. Figure 2 below shows the results of the evaluation.

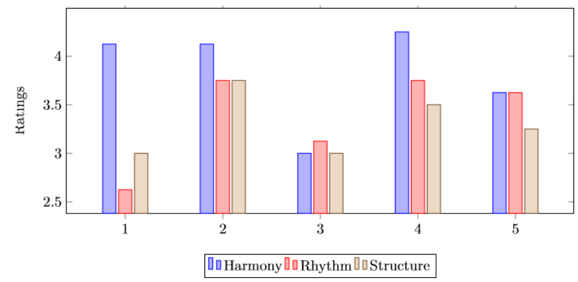


Figure 2: Results for subjective evaluation.

In general, ChordAL's compositions score highest in terms of harmony, with an average rating of 3.825. This further shows that a chord-based generation framework can guarantee harmonic, pleasant-sounding generation. However, ChordALs compositions lack rhythm and structure, with an average rating of 3.375 and 3.3 respectively. We think that it is possible to build additional components on top of the chord-based framework that handle the aspects of rhythm and structure, while preserving the strength of this framework in composing fluent harmony.

## Conclusion and Future Work

We propose a 3-step chord-based framework for melody generation, which is able to ensure fluent harmony in the generated pieces. We also show that chord embeddings could be learnt by a simple chord generation task, which they highly resemble the Circle of Fifths.

ChordAL is most suitable for generating chamber music and strings/woodwind ensemble pieces, as these types of music are chordal in nature. Future work will be focusing on increasing the vocabulary of chords used, as well as improving on the rhythm and structure in the compositions.

ChordAL is open-source and welcomes contribution at: `https://github.com/gudgud96/ChordAL`. All compositions by ChordAL can also be heard on Soundcloud at: `https://bit.ly/2HzighM`.

## Acknowledgement

I would like to thank my supervisor, Prof Bo An for providing guidance and support for this project. I would also wish to acknowledge the funding support for this project from Nanyang Technological University under the Undergraduate Research Experience on CAmpus (URECA) programme.