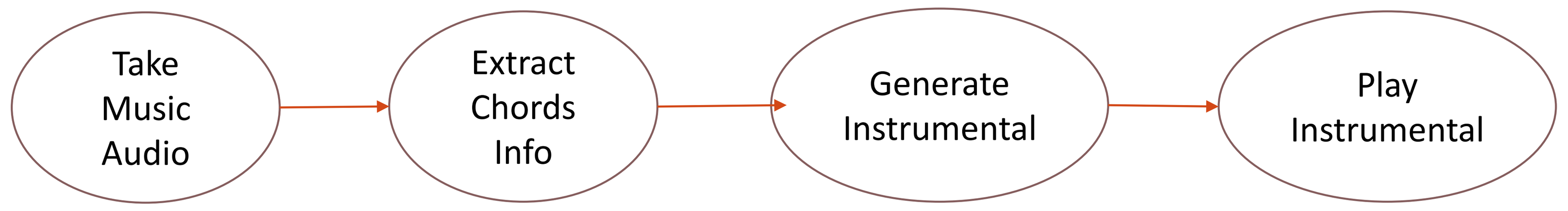
Music Generator

Kelly Ndombe (ndombe@stanford.edu)
Stanford University

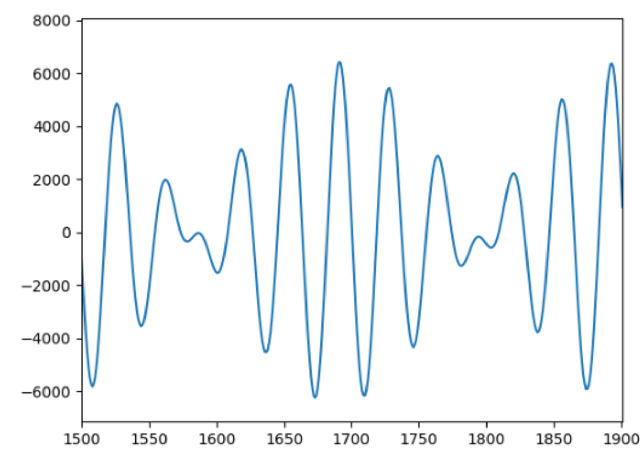
Objective



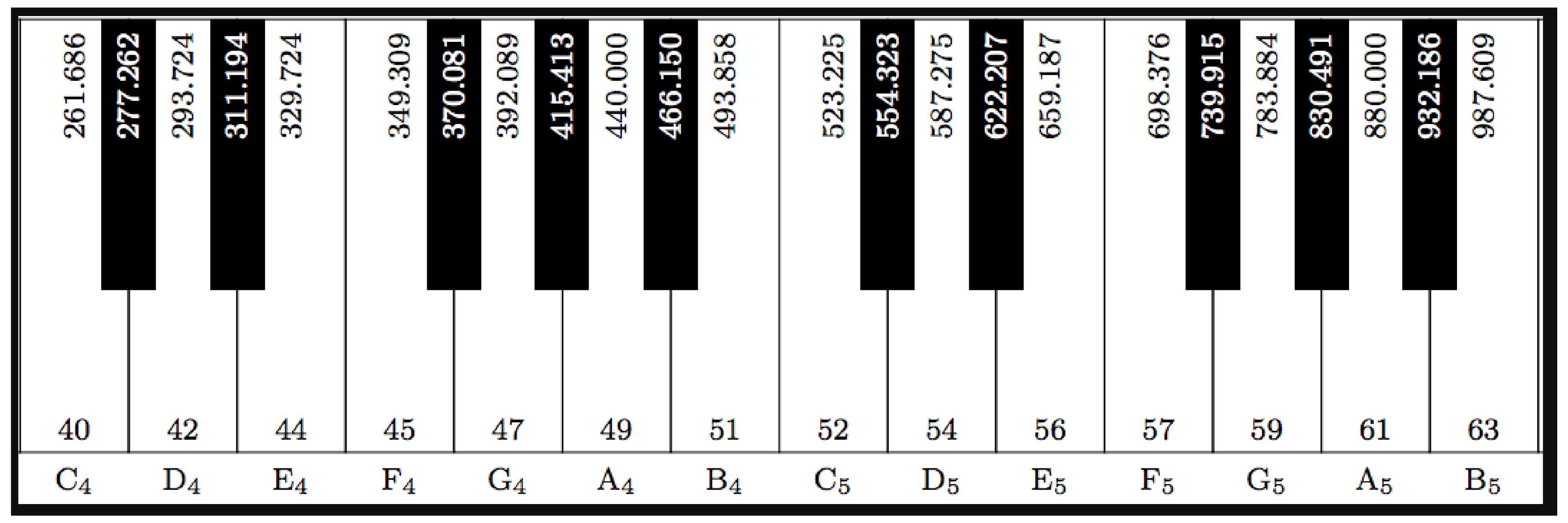
We want to be able to take an audio file as input, decipher the chords that are played in the audio, come up with an internal representation of those chords and play back an auto-generated instrumental of the music.

Chord Extraction (Multiple Notes)

When multiple notes are combined, we still get the observation as one audio signal containing the all the signals mixed together.
It is possible to get the **beat frequency** when the number of notes combined is small, but it is not as easy a task when there are numerous signals mixed together.
To the right is the signal for an A4 (440 Hz) and a C5 (523.25 Hz) produced by a sine wave instrument.

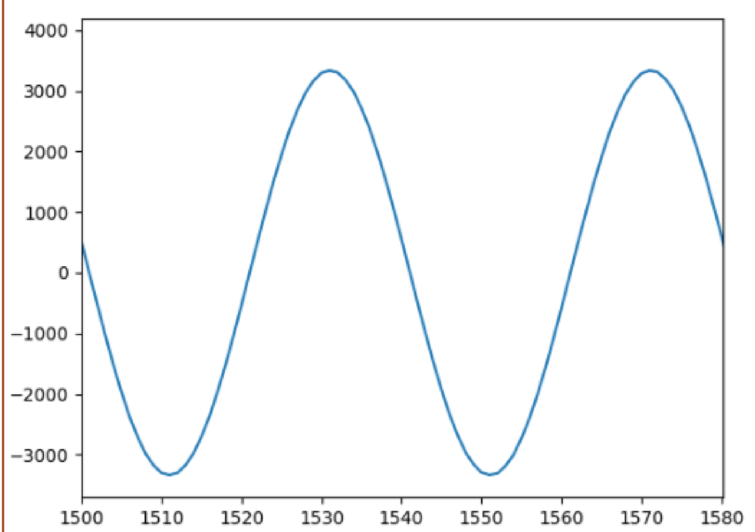


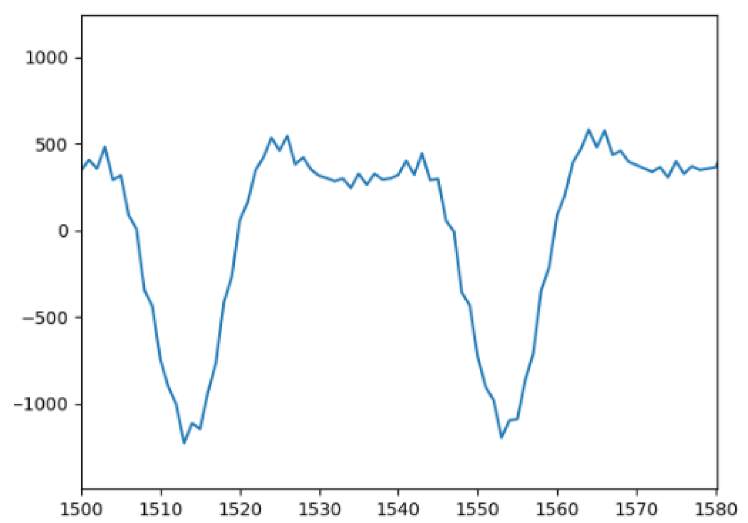
NOTICE: The frequency of the lowest note will always stand out from a sound wave of combined notes.

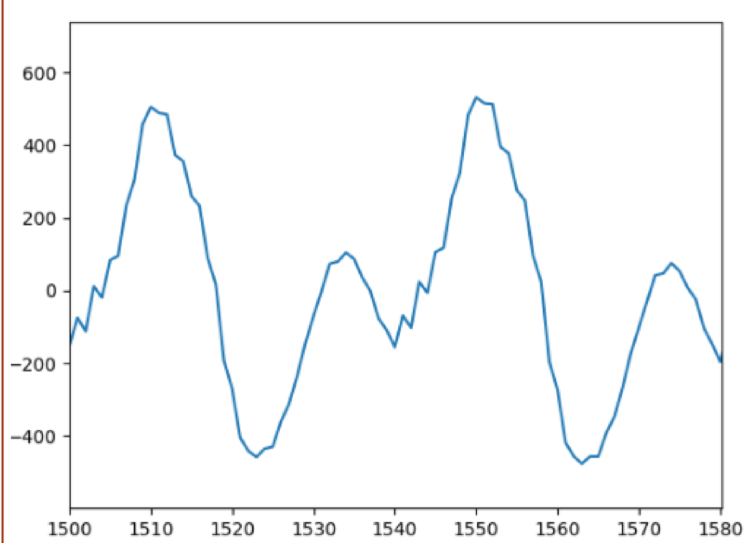


Chord Extraction

Successfully extracting the chords necessitates understanding what separates different chords and what groups similar chords (Frequencies). Below are the sound waves for the note A4 from a sine wave instrument (top left), a piano (top right) and a violin (bottom), sampled over the same window of 2/440 seconds (aka, 2 times the period of a A4)







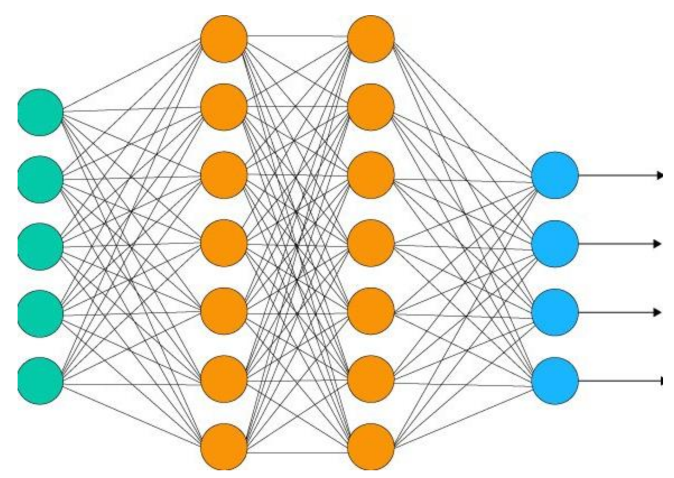
$$Period = 1/frequency$$

$$Beat\ Freq = |Freq1 - Freq2|$$

Similar notes played on different instruments will have similar frequencies/periods. Their sound waves may look nothing similar, but the frequency of repeating patterns is the same in a given window of time. This is good as it allows transferability of our trained knowledge.

Chord Extraction (Models)

We used a few Neural Networks to help with chord identification.



1. Fully Connected NN: Train on all chords that we wish to identify. When predicting chords in a song, define a 'window width'. The associated with the window is the highest on average in that window.
2. Multiple Fully Connected NN: Train 12 models to identify the presence of each of the 12 musical notes. This allows for far better precision in defining the 'chords'.
3. [Multiple/Single] Recurrent NN: Allow for the processing of long short term memory and a bidirectional connection between points of the sound wave, especially if we want to train on identifying periods and not overfitting to the sound waves used in training.

Prediction/Representation

For a given audio file, we predict the chords by scanning along a fixed window width the most probable chords and generalizing for the entire window.
We use various hyperparameter such as '**precision**' and '**window**' that define how deeply we want to identify chords variations. **Higher precisions** often lead to unrealistic chord variations because they pick up subtle change in prediction, while **lower precisions** may miss to identify some variations.
Tempo: we identify the tempo from the average shortest variation. If the tempo is correctly identified, we are able to quantize the playing of the music better.

```
('C_M', 17.68, 19.67, 0.4511111111111113)
('D_m', 19.67, 21.24, 0.4577777777777778)
('E_m', 21.24, 23.0, 0.4088888888888889)
('F_M_base_G', 23.0, 24.7, 0.4466666666666666
('C_M', 24.7, 26.43, 0.4288888888888889)
('D_m', 26.43, 28.13, 0.4333333333333335)
('E_m', 28.13, 29.84, 0.3955555555555555)
('F_M_base_G', 29.84, 31.45, 0.331111111111111
('D_m', 31.45, 33.25, 0.4355555542310079)
('E_m', 33.25, 34.92, 0.4622222222222222)
('F_M', 34.92, 36.69, 0.4777777777777778)
('E_m', 36.69, 38.53, 0.3777777777777777)
('D_m', 38.53, 41.87, 0.3222222222222224)
('F_M_base_G', 41.87, 45.17, 0.4466666666666666
('C_M', 45.17, 48.91, 0.4666666666666667)
('D_m', 48.91, 52.18, 0.4311111111111111)
('F_M_base_G', 52.18, 55.46, 0.46)
```

Play Instrumental: We are able to play using the Midi library in PyGame.

Results

The most challenging aspect is the chord identification. The Neural Nets are giving pretty satisfactory results when it comes to identifying more general chords such as 3-note chords in a given key.

However, differentiating between more nuanced chords is being quite a challenge. Chords such as A minor 9 and A minor 11 sound very similar but have differences.