Prediction and Harmony

Abstract

Hawkins' model of human learning has been presented in terms of conditional probability of event series in time. We investigate a particular case where time does not matter: chords in musical harmony. We show that, by using a suitable domain, we can again turn chord recognition into a problem of learning conditional probability of event series, in the chosen domain.

Hawkins' Idea

Jeff Hawkins, "The Brain - Prediction and Intelligence", talk 21 march 2005:

- [the brain] learns sequences and stores conditional probabilities
- when you put these together in a hierarchy, which is what the cortex is, you have a tree of conditional probability functions

The idea sounds reasonable in a wide variety of experiences.We concentrate on music and improvisation.

- We learn music by listening troughout our whole life.
- By learning, we build and renew our music taste.
- In music improvisation, the improviser tries to match our music taste, while giving out someting new, fresh, unexpected.
- Our music taste: an hierarchical product of culture, genre, tonality and harmonic structure of the piece, current part being played (e.g. strophe or refrain), current chord, current phrase of the solo...
- Melodic improvisation seems to conrm Hawkins' model.

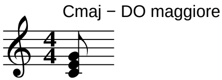
Concepts that do not Depend on Time

Claim:

There are concepts that we can learn, and that are not time dependent!

Examples:

- the tactile characteristics of different surfaces
- the scent of a flower
- the taste of our favourite food
- the sound of different kinds of chords!



DO MI SOL

DO MIb SOL

Do we really learn Chords? An experiment

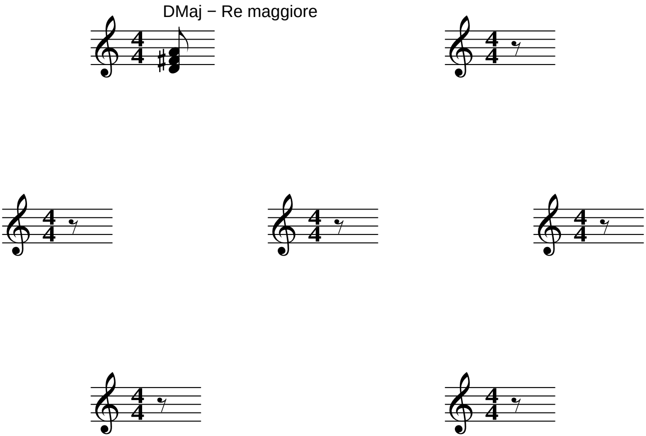
Let's see if we can learn the difference between major and minor chords.





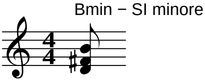
Do we really learn Chords? An experiment

Let's see if we can learn the difference between major and minor chords.



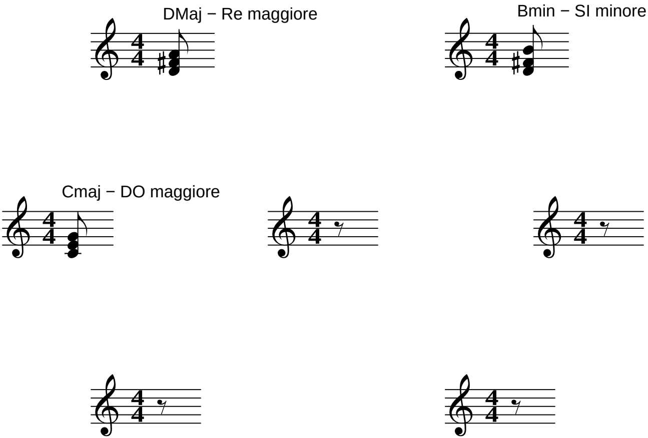
Do we really learn Chords? An experiment

Let's see if we can learn the difference between major and minor chords.



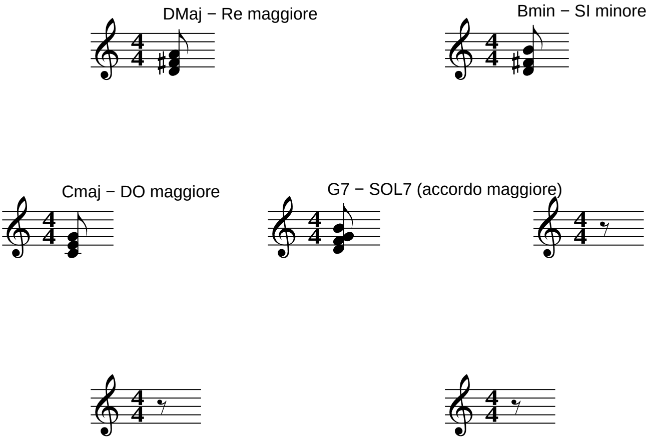
Do we really learn Chords? An experiment

Let's see if we can learn the difference between major and minor chords.



Do we really learn Chords? An experiment

Let’s see if we can learn the difference between major and minor chords.



Do we really learn Chords? An experiment

Let’s see if we can learn the difference between major and minor chords.

DMaj − Re maggiore

Bmin − SI minore

Cmaj − DO maggiore

G7 − SOL7 (accordo maggiore)

Emin

Do we really learn Chords? An experiment

Let's see if we can learn the difference between major and minor chords.

DMaj − Re maggiore

Bmin − SI minore

Cmaj − DO maggiore

G7 − SOL7 (accordo maggiore)

Emin

Cmin7 − DO minore 7

Do we really learn Chords? An experiment

Let's see if we can learn the difference between major and minor chords.

DMaj − Re maggiore

Bmin − SI minore

Cmaj − DO maggiore

G7 − SOL7 (accordo maggiore)

Emin

Cmin7 − DO minore 7

CMaj7 ~ Emin6

Idea: Changing Domain

Is it still possible to view chord recognition as a problem of learning conditional probability?

- In our opinion, yes, by changing the domain from time to frequency.
- Difficult to find such a theory applied in computing, since rule-based approaches work very well, and there's no need to implement it in practice.
- However, we found similar work by Cabral, Pachet and Briot. A comparison between rule-based and machine-learning approaches has been made by Gomoz and Herrera.

The idea: represent a single "moment" of a sound in the frequency domain as a sum of a certain number of base frequencies, using the Fourier Series as a basis.

- Problems: relative pitches, chord voicings.

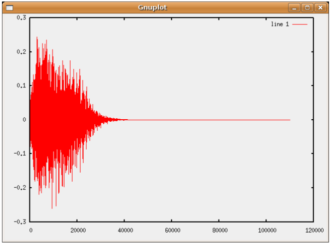
Harmonic Analysis

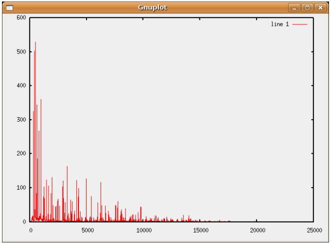
- Any periodic function of period T is decomposable as a sum of infinite sinusoids, the **Fourier Series**:

$$f(x) = \sum_{n=0}^{\infty} a_n \cos(\frac{2\pi nx}{T}) + b_n \sin(\frac{2\pi nx}{T})$$

  a_n and b_n are the amplitude contributions of every harmonic with frequency $\omega_n = \frac{2\pi n}{T}$.

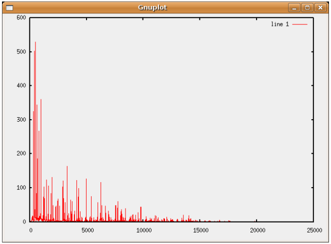
- The *Fourier transform* is a mathematical function that can extract these coefficients from audio data.

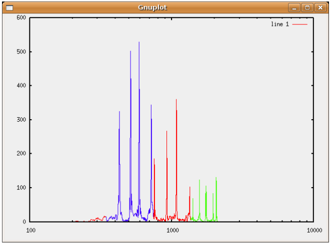




Logaritmic Scale

- The integer multiples of a base frequency, $\omega = \frac{2\pi}{T}$, are called **harmonics** of ω.
- In music, every time that frequency doubles (we go *an octave higher*) we obtain the same note.
- More than the distance between frequency components, it matters their ratio. For this reason, the audio spectrum (i.e. the Fourier coefficients) are represented using a logarithmic scale.





Octaves, voicings and PCP

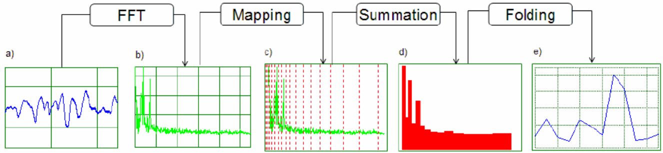
A very important problem to face, if we want to give a satisfying model of how we perceive different kinds of chords, is that two different *voicings* usually sound almost the same to us:

Cmaj − DO maggiore

Cmaj − DO maggiore

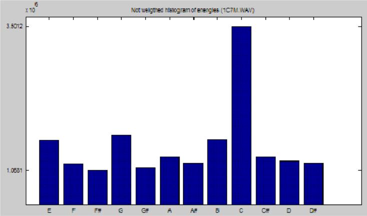
- This is due to the fact that we perceive sounds of different octaves as being similar.

In order to solve this problem, the Fourier coefficients are often divided into octaves and summed, obtaining the so-called PCP (pitch class profiles).

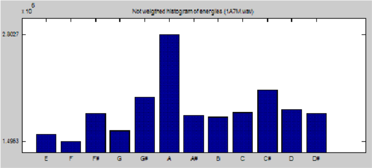


PCP Examples

Some examples taken from the work of Cabral, Pachet, Briot:



Example of the PCP for a Cmaj7

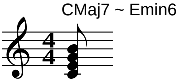


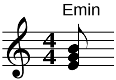
Example of the PCP for a Amaj7. Each column represents the intensity of a note, independently on the octave.

Conditional Probability Recovered

Switching from time to frequency, and using PCPs, we can talk again of a conditional probability of a sequence of events:

- On this representation, we again have events, which are occurrencies of frequencies, starting from the *dominant* of a chord.
- What is the dominant? We don't know!
- Perhaps we have to try them all, and to accept ambiguity. In fact, there are chords that *sound like* they are both minor and major.





Unfortunately, not so many authors dealed with this approach, since rule based approaches work very well in practice, and there's no need to use machine-learning algorithms.

Chord Recognition in EDS (1)

G. Cabral, F. Pachet, J. Briot: “Recognizing Chords with EDS: Part One”.

- They present a genetic approach based on various combinators for features extracted from the PCP

In the paper, there is a comparison between:

- Their genetic approach, on which the EDS tool, implemented at SONY Paris, is based.
- A simple method based on direct comparison between sample PCP's (PCP templates) and the current sample (the “brute force approach”).
- A *k-nearest-neighbour* algorithm, which is interesting for our purposes because, even if it is not a probabilistic approach, it is still a non-symbolic machine-learning algorithm.

Results

The EDS tool is able to extract complex featuers useful for classification, e.g.:

```
EDS1: Power (Log10 (Abs (Range (Integration (Square (Mean (FilterBank
(Normalize (Signal), 5.0))))))), -1.0)

EDS2: Power (Log10 (Abs (Range (Sqrt (Bartlett (Mean (FilterBank
(Normalize (Signal), 9.0))))))), -1.0)

EDS3: Sqrt (Range (Integration (Hanning (Square (Mean (Split (Signal,
3736.0)))))))
```

Selected features for the *Amaj/min* chord recognizer

On the other hand, results clearly shows that the KNN approach gives better results:

Approach / Database	PCP Template	KNN	EDS
Maj/Min (fixed root)	100%	100%	90.91%
Chord Type (fixed root)	89%	90.62%	87.5%
Chord Recognition	53.85%	63.93%	40.31%

This result is attributed to the fact that, in this work, they used EDS as a normal user, not a signal processing expert, would do. They expect better result if EDS is used by an expert of the domain.

Cognitive vs. Machine Learning Strategies

An interesting paper, which is cited in the work above, is "Estimating the tonality of polyphonic audio files: cognitive versus machine learning modelling strategies" - E. Gomez, P. Herrera.

They compare a cognitive, rule-based approach, and various machine-learning approaches, in various problem related to tonality recognition.

Problems:

- learning the tonic (key note) of a piece
- learning the *mode* (e.g. major, minor, dorian...)
- learning both the key note and the mode

Audio database:

- 878 excerpts of classical music (Mozart, Chopin, Scarlatti, Bach. . .) plus some jazz versions of classical pieces
- all the key note and mode annotations were taken from the FreeDB database

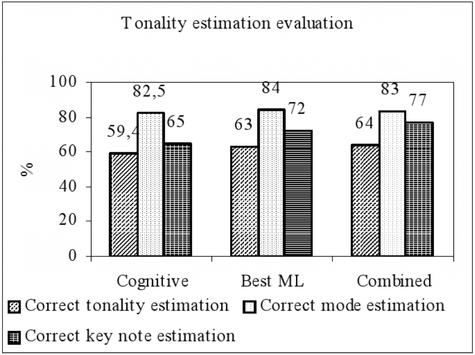
Various kind of learners:

- knn, bayesian, neural networks, SVM. . .

Results

There is *no* single best learner:

- When learning the tonic, a bayesian (stochastic) classifier gave the best results (72% accuracy).
- When learning the mode of a piece, knn scored far above the others (same results in Cabral, Pachet, Briot).
- When trying to solve both problems simultaneously, a multilayer perceptron (a form of neural network) proved the best solution.



Cognitive Dissonance

From Wikipedia, the free encyclopedia:

Cognitive Dissonance (L. Festinger, 1959)

Overview:

According to cognitive dissonance theory, there is a tendency for individuals to seek consistency among their cognitions (i.e., beliefs, opinions). When there is an inconsistency between attitudes or behaviors (dissonance), something must change to eliminate the dissonance.

Basic Theory

The introduction of new cognition that is dissonant with a currently held cognition creates a state of "dissonance". Dissonance can be reduced either by eliminating dissonant cognitions, or by adding new consonant cognitions. Once dissonance reaches a level that overcomes the resistance of one of the cognitions involved, that cognition will be changed or eliminated, and dissonance will be reduced.

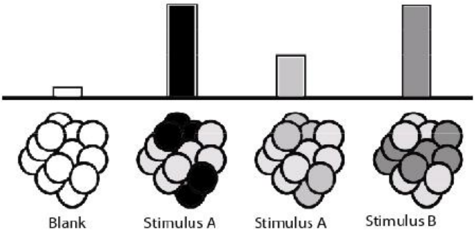
This reminds music listening:

- We know music since we have listened to music.
- We like certain kinds of music, and we tend to focus on listening those.
- When we listen to new, different kinds of music, we experience a “discrepancy between attitudes and behavior”.
- We can either change the attitude (i.e. learn the new kind of music) or the behavior (i.e. say that we’re bored and stop listening).

Adaptation to Repetitive Stimuli

Zoe Kourtzi (Max Planck Institute for Biological Cybernetics), Kalanit Grill-Spector (Stanford University) - "fMRI adaptation: a tool for studying visual representations in the primate brain".

- One of the most fundamental properties of the brain that clearly distinguishes it from artificially constructed computational devices is its ability to continuously update its functional properties based on prior experience.
- Fairly long term changes (on the order of days) in the brain structure were observed after subjects learned to recognize unfamiliar shapes, or when trained to recognize subliminally-presented visual objects and even single presentations of objects.
- Experience-dependent changes are not only evident on long range time scales lasting days, but also in short times scales in the order of seconds.
- A particularly robust phenomenon is repetition-suppression, or adaptation, in which repeated presentation of the same stimulus leads to a consistent and gradual reduction in activation within seconds of the occurrence of the first stimulus presentation.



Cognitive Dissonance and Harmony

Can cognitive dissonance be explained in Hawkins' model?

- Cognitive dissonance would compell a bayesian learner to increase its memory!
- The fact that the brain "prefers" repetition and predictable behaviors is evident in repetition-suppression.
- All this can be seen as the principle of minimizing the quantity of information that one has to store in order to understand the surrounding environment.

By changing the domain from time to frequency, we can see harmonic dissonance as cognitive dissonance in the same way.

- It is well known that, if one is "trained" to listen to harmonic dissonances, he/she can start to "understand", and appreciate dissonance
- *Doesn't this look like "changing the attitude"?*

Conclusions

We showed that, by choosing a suitable domain different than time, a classification problem that does not depend on time can become similar to problems related to event series in time.

An important open question:

- If Hawkins' model is a good model of human learning, and can be extended to other domains, how is this reflected into the phisiology of the brain?

There should be phisiological evidences that

- The brain can decode signals and work in domains different than time
- there is some part of the brain, or of sensor organs, that can convert signals into a frequency spectrum. If so, is it a specialized function or is it "learned" in turn?


LOONEY TUNES
"That's all Folks!"