Harmonic Analysis of Music With CCG

First Year Review

Mark Wilding

School of Informatics
Edinburgh

22nd June 2010

Introduction

- Automatic analysis of tonal music
- Music is structured
  - Hierarchical structure in harmony
  - Hierarchical structure in meter
- We're looking at harmony

Music's Relationship to Language

- Harmonic structure similar to language's syntactic structure
- Common analogy
- Evidence of shared specialized cognitive processing
- Katz and Pesetsky (in draft):

> All formal differences between language and music are a consequence of differences in their **fundamental building blocks**... In all other respects, language and music are identical.

Harmonic Analysis

- Interpretation is hard
- Highly ambiguous surface form
- Needed for many musical tasks: e.g. transcription, harmonization
- We currently handle chord sequences

Background

- Bernstein's lectures, 1973
- Transformational grammars for music
- *A Generative Theory of Tonal Music* (GTTM, Lerdahl & Jackendoff): grammatical rules, transformations, preference rules
- Many approaches use preference rules
- Some others have used NL-style grammars
- Bod applied DOP to music
- Many statistical approaches: HMMs, n-grams

Introduction

Musical Syntax

Tonality and Interpretation

Our Grammar

Corpus

Baseline Models

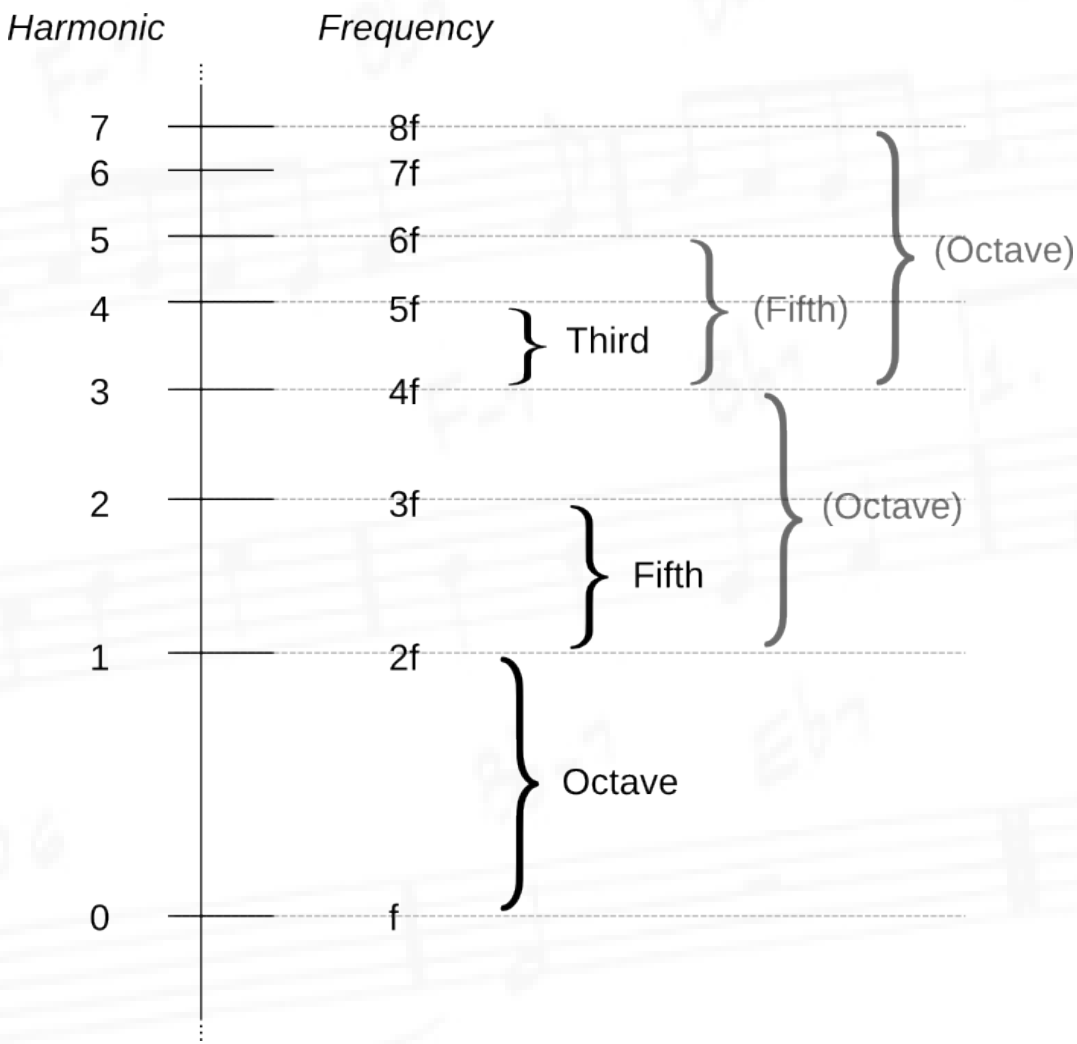
Next Steps

Plan

Structure In Harmony

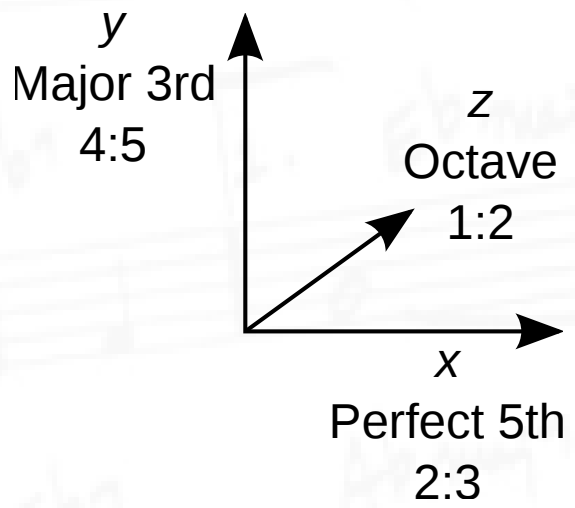
- Chords classified as *tonic*, *dominant* or *subdominant*
- Dominant creates expectation of resolution to tonic ♫
- Building block of harmonic structure
- ♫ Recursive
- ♫ Coordinated
- ♫ Substitutions
- ♫ *Call Me Irresponsible*

Tonality


Harmonic
Frequency
7
6
5
4
3
2
1
0
8f
7f
6f
5f
4f
3f
2f
f
Third
(Fifth)
(Octave)
(Octave)
Fifth
Octave

Tonal Space

- Longuet-Higgins formalized this as a 3D space
- Discrete space
- Project along z-dimension to ignore octaves



Tonal Space

$\sharp G^-$	$\sharp D^-$	$\sharp A$	$\sharp E$	$\sharp B$	$\sharp\sharp F^+$	$\sharp\sharp C^+$	$\sharp\sharp G^+$	$\sharp\sharp D$
E^-	B^-	$\sharp F$	$\sharp C$	$\sharp G$	$\sharp D$	$\sharp A^+$	$\sharp E^+$	$\sharp B$
C^-	G^-	D^-	A	E	B	$\sharp F^+$	$\sharp C^+$	$\sharp G$
$\flat A^-$	$\flat E^-$	$\flat B^-$	F	C	G	D	A^+	E^+
$\flat F^-$	$\flat C^-$	$\flat G^-$	$\flat D^-$	$\flat A$	$\flat E$	$\flat B$	F^+	C^+
$\flat D^{--}$	$\flat\flat A^-$	$\flat\flat E^-$	$\flat\flat B^-$	$\flat F$	$\flat C$	$\flat G$	$\flat D$	$\flat A$

Equal Temperament

$\sharp G^-$	$\sharp D^-$	$\sharp A$	$\sharp E$	$\sharp B$	$\sharp\sharp F^+$	$\sharp\sharp C^+$	$\sharp\sharp G^+$	$\sharp\sharp D$
E^-	B^-	$\sharp F$	$\sharp C$	$\sharp G$	$\sharp D$	$\sharp A^+$	$\sharp E^+$	$\sharp B$
C^-	G^-	D^-	A	E	B	$\sharp F^+$	$\sharp C^+$	$\sharp G$
$\flat A^-$	$\flat E^-$	$\flat B^-$	F	C	G	D	A^+	E^+
$\flat F^-$	$\flat C^-$	$\flat G^-$	$\flat D^-$	$\flat A$	$\flat E$	$\flat B$	F^+	C^+
$\flat D^{--}$	$\flat\flat A^-$	$\flat\flat E^-$	$\flat\flat B^-$	$\flat F$	$\flat C$	$\flat G$	$\flat D$	$\flat A$

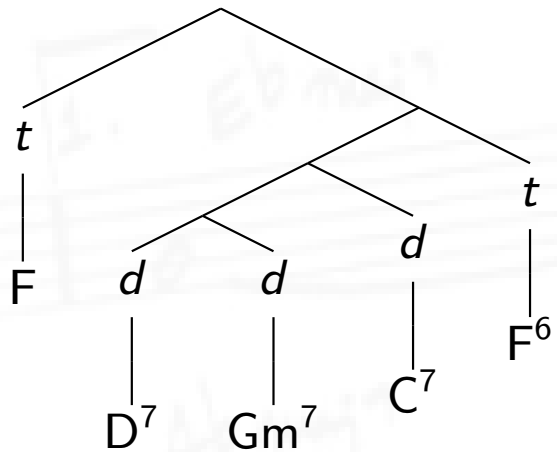
- Spread 12 notes evenly across octave

Tonal Ambiguity

- Harmonic relation between notes is ambiguous under ET
- Musical notation partly distinguishes
- Must disambiguate to transcribe scores
- Harmonic relations dependent on key
- Important for generation tasks
- Corresponds to harmonic analysis done by:
  - musicologists
  - jazz performers (when improvising)
  - you (unconciously)

Syntax

- Need harmonic structure to establish harmonic relations between chords
- Map notes in ET space back onto infinite space
- Syntax $\rightarrow$ harmonic structure
- Semantics $\rightarrow$ tonal space interpretation



Our Grammar: Syntax

- Modification of NL CCG
- Atomic categories of form:

$$X{-}Y$$

- Important syntactic information: start and end chord roots
- E.g. simple tonic chord on C interpreted as:

$$C{-}C$$

- Just one chord – start and end

Our Grammar: Syntax

- Complex categories built up with slashes
- ♫ Simple dominant chord on G expects resolution to C

$$G{-}X \;/\; C{-}X$$

- Next category must begin at C

$$\frac{\dfrac{\mathrm{G}^7 \qquad \mathrm{C}}{G{-}X \;/\; C{-}X \qquad C{-}C}}{G{-}C}\,{>}$$

- Interprets this chord as a movement from G to C

Our Grammar: Semantics

- Semantics: path through tonal space
- λ-calculus terms on lexical entries
- Represented as a list of points in the space
- Tonic: single point in space

  $C{-}C \quad : \quad C :: Nil \qquad [C]$
- Dominant: one point moving to another

  $G{-}X \;/\; C{-}X \quad : \quad \lambda x.G :: x \qquad [G] + x$

  $\lambda x.G :: x \quad C :: Nil \quad \Rightarrow \quad G :: C :: Nil \qquad [G, C]$

Ambiguity

- Huge lexical ambiguity
- Generally use only a few chord types
- Lots of possible structures
- Big search space for parsing
- Sparse data
- Need some good modeling

Jazz Corpus

- Hand-constructed corpus
- ∼100 sequences (∼4,000 chords)
- Annotated with lexical entries and trees
- Only training set

Some Baseline Models

- Two simple parsing models
- Super-naive baseline
  1. PCFG model
  2. C&C supertagger

Aside: Absolute Pitch

- Absolute pitch irrelevant to interpretation
- More or less irrelevant to perception
- Models should be insensitive to absolute pitch

Model 1: The PCFG Model

- Generative model based on Hockenmaier (2001)
- Frequentist model of derivations
- Only naive smoothing
- Beam applied to chart during parsing
- All pitches relative to parent in expansions

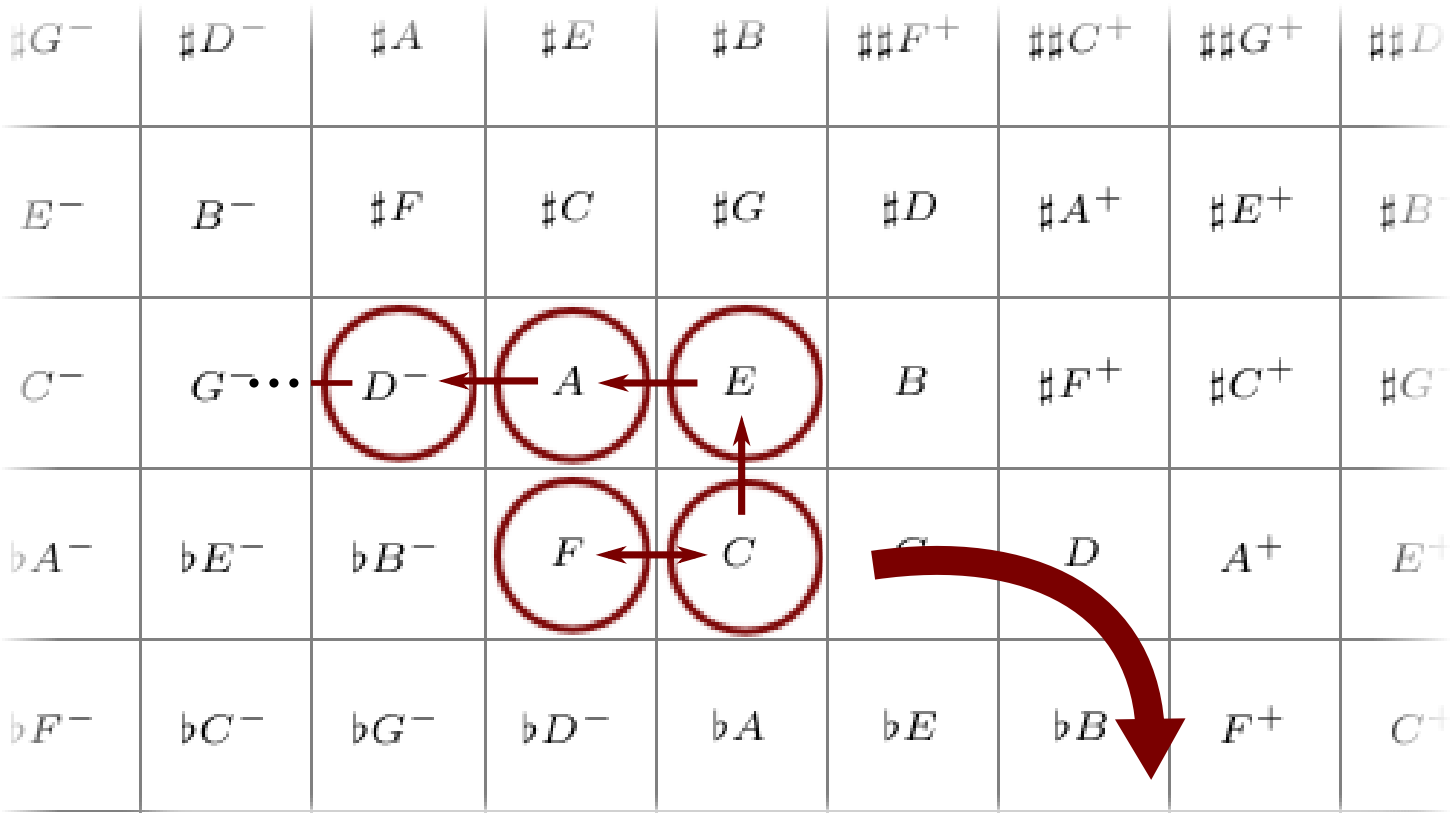
Model 2: Supertagger Model

- Uses supertagger component of the C&C parser (Clark and Curan, 2002)
- Chooses categories for chords using log-linear model
- Parser asks for less probable tags if it can't get a full parse
- Observations: chord types and intervals between chord roots

Evaluation

- 10-fold cross validation
- Only fully annotated sequences (65)
- Compared gold standard tonal space paths to top result
- Report error rate in alignment of paths and accuracy of aligned points

Evaluation: Path Comparison



Evaluation: Path Comparison

GS: $C_T :: F_S :: C_T :: E_D :: A_D :: D_D :: \ldots$

Costs: 0 $?_1$ $?_1$ 0.5 0 0

Result: $C_T :: E_T :: A_D :: D_D :: \ldots$

Total cost: 2.5
Accuracy: 42%

Results

	C&C supertagger	PCFG
Recall	21.4%	100.0%
Mean accuracy	90.8% (10.0%)	59.8% (11.9%)
Mean error rate	0.10 (0.11)	0.41 (0.13)

- PCFG: parses everything, but not that well
- C&C: terrible recall

Next Steps

- Both methods can be improved
- Smoothing is critical
- More annotation to fill in gaps
- Better input to the tagger
- Combined models

Plan

1. Devise evaluation methods
   - use them to do consistent evaluation of baseline methods already implemented
   - evaluate improvement of better models as they are developed

   Months 10-11
2. Enhance the baseline models with more sophisticated smoothing 11-13
3. Consider alternative annotation schemes. Possibly get more data 10-14
4. Try more sophisticated statistical models 14-20
5. Investigate annotator agreement. Experiment with other annotators 14-20
6. Investigate semi-supervised methods 20-24
7. MIDI realization models 20-34
8. Use best full models (tonal space interpretation from MIDI) to implement some applications 25-34
9. Write up 31-36