

# Parsing Jazz

## *Second Year Review*

Mark Granroth-Wilding

School of Informatics  
Edinburgh

12 Oct 2012

# A Song

- *Night and Day*, Cole Porter
  - Fred Astaire
  - Ella Fitzgerald
  - Coleman Hawkins
- Obviously the same song: melody, rhythm, words, ...
- Still recognisable
- Different melody, different rhythm, no words, different harmony
- Harmonic structure

# A Song

- *Night and Day*, Cole Porter

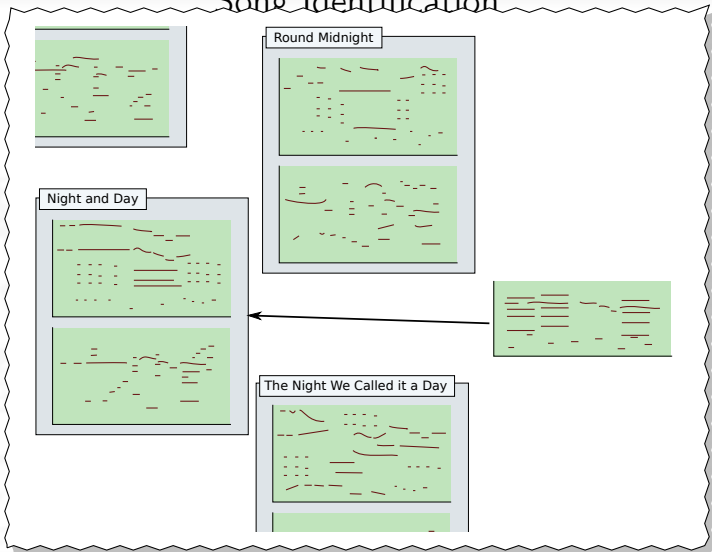
The image shows a musical staff with a treble clef and a key signature of one flat (B-flat). The melody is written in 4/4 time. Above the staff, red arcs indicate harmonic groupings. The first group of four notes (D4, F4, A4, Bb4) is bracketed together. The second group of two notes (G4, A4) is bracketed together. The third group of two notes (Bb4, C5) is bracketed together. The fourth group of two notes (D5, E5) is bracketed together. The fifth group of two notes (F5, G5) is bracketed together. The sixth group of two notes (A5, Bb5) is bracketed together. The seventh group of two notes (C6, Bb5) is bracketed together. The eighth group of two notes (A5, G5) is bracketed together. The ninth group of two notes (F5, E5) is bracketed together. The tenth group of two notes (D5, C5) is bracketed together. The chord symbols D-7 b9, G7, -9-7, Cmaj7, D-7 b9, G7, -9-7, Cmaj7 are written below the staff.

- Harmonic structure

# Song Identification

- *Music information retrieval*
- Song identification from a database
- Difficult task

# Song Identification



# Musical Similarity

- Need metrics of musical similarity
  - Rhythmic patterns
  - Melodic similarity
  - **Harmonic similarity**
- Structured harmonic analysis
- Not used much to date for music similarity

# Musical Similarity

●

Structured harmonic analysis

Not used much to date for music similarity

## Two Problems



### Harmonic parsing:

identify harmonic structures in a stream of notes

The image shows a musical score in 2/4 time. The first two measures are highlighted with a red box and a green box, respectively. Red arcs connect notes across measures, indicating harmonic relationships. The score includes treble and bass staves with triplets and rests.

- Segmentation

## Two Problems

- An easier problem

A<sup>2</sup>

### Chord parsing:

identify harmonic structures in a sequences of chord symbols

A musical score in 2/4 time, key of G major, showing a sequence of chords: G, C, G, C, G, D7, G. The score is written in treble and bass clefs. Red arcs are drawn above the notes, connecting them to the chord symbols below, illustrating the harmonic structure. The arcs show how the notes in the melody and bass line relate to the chords. For example, the first arc connects the G notes in the melody and bass to the G chord. The second arc connects the C notes in the melody and bass to the C chord. The third arc connects the G notes in the melody and bass to the G chord. The fourth arc connects the C notes in the melody and bass to the C chord. The fifth arc connects the G notes in the melody and bass to the G chord. The sixth arc connects the D, F, and A notes in the melody and bass to the D7 chord. The seventh arc connects the G notes in the melody and bass to the G chord.

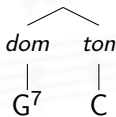
# Easier Problem: Chord Parsing



- Structures of tonal harmony
- Harmonic adaptation of *combinatory categorial grammar*
- Harmonic semantics
- Parse chord sequences

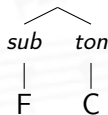
## Harmonic structure

- Chords classified by **function**: *tonic*, *dominant* or *subdominant*
- Tonic creates no tension
- Dominant creates expectation of resolution to tonic
- Subdominant resolves to tonic



## Harmonic structure

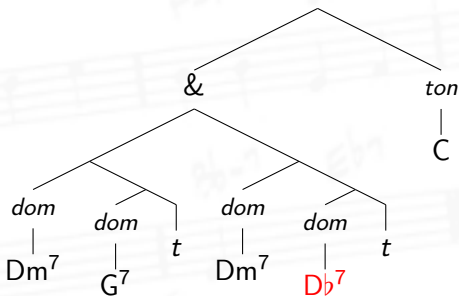
- Chords classified by **function**: *tonic*, *dominant* or *subdominant*
- Tonic creates no tension
- Dominant creates expectation of resolution to tonic
- Subdominant resolves to tonic



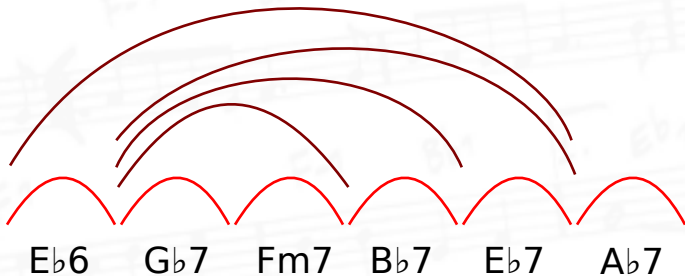
## Harmonic structure

- Chords classified by **function**: *tonic*, *dominant* or *subdominant*
- Tonic creates no tension
- Dominant creates expectation of resolution to tonic
- Subdominant resolves to tonic

*Recursion — Coordination — Substitution*



# Parsing



- HMM supertagger
- Parsing model: Hockenmaier (2001)

# Chord Parsing Results



<i>Model</i>	<i>P (%)</i>	<i>R (%)</i>	<i>F (%)</i>	<i>Cov. (%)</i>
HMM baseline	74.6	82.1	78.2	100
Supertagger/parsing model	<b>87.4</b>	85.7	86.5	95
Supertagger/parsing model, backoff	85.3	<b>91.0</b>	<b>88.1</b>	100

- Grammar model better than HMM baseline
- Using HMM baseline as backoff produces best results

# The Harder Problem



**Harmonic parsing:** identify harmonic structures in a note stream



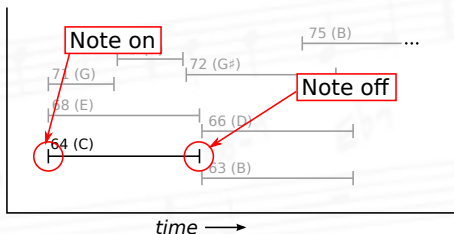
**Chord parsing:** identify harmonic structures in a chord sequence

## Aside: MIDI

- Symbolic musical data format
- For sending instructions between instruments
- MIDI files
- Wide range of musical instructions
- Most importantly: **note-on** and **note-off** events

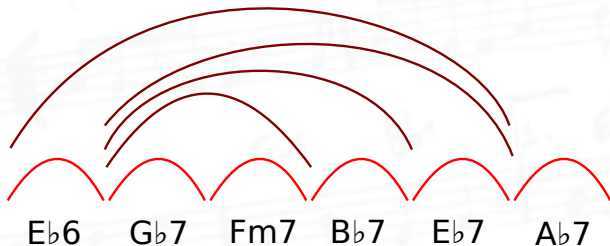
# MIDI

Musical Instrument Digital Interface



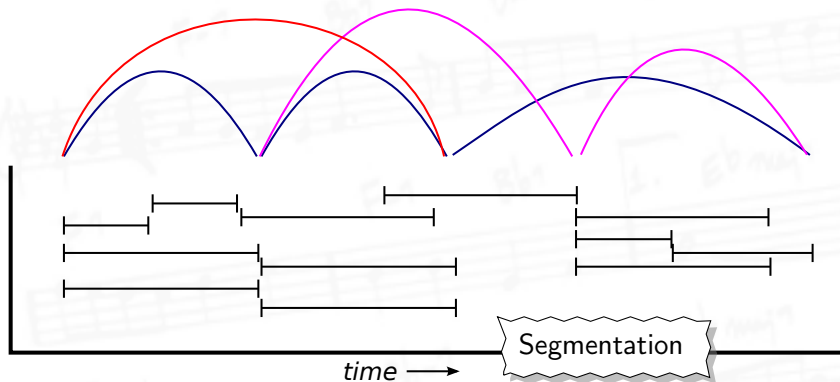
- Lots of MIDI files on the web

# The Harder Problem: One Approach



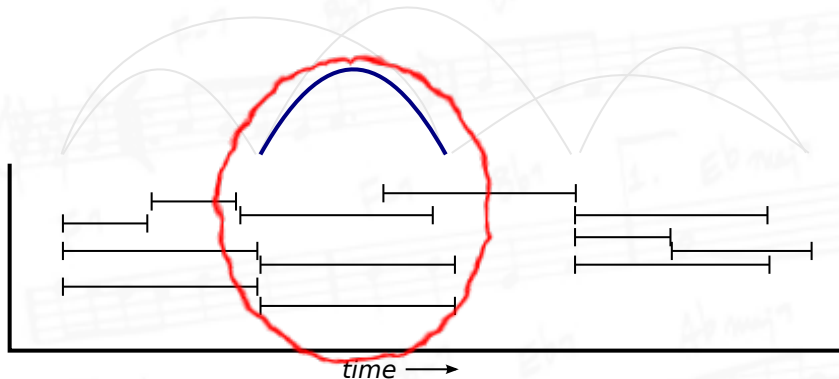
- Replace chord supertagger with MIDI supertagger

# A Music Supertagger



- Spanning edges
- How can we model these complex note streams?

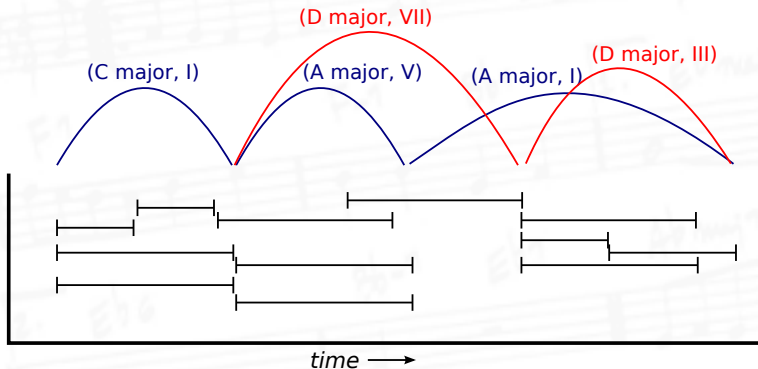
# A Music Supertagger



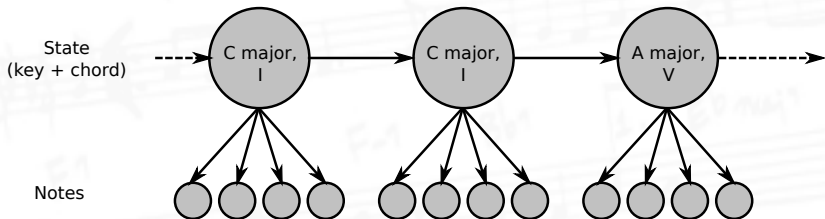
- Spanning edges
- How can we model these complex note streams?

# Raphael and Stoddard's Model

- Raphael & Stoddard (2004) (*R&S*)
- Different kind of harmonic analysis

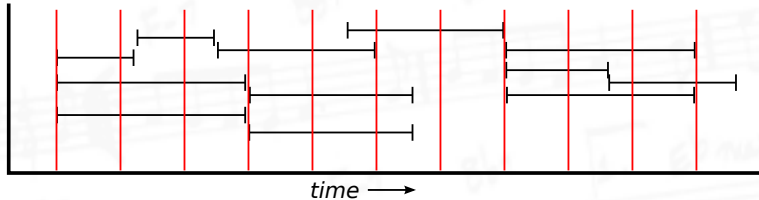


# Raphael and Stoddard's Model



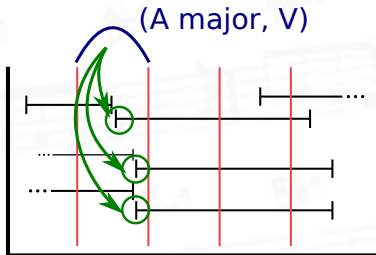
- HMM
- State labels → harmonic analysis
- Emissions → bag of notes

# R&S Model: Assumptions



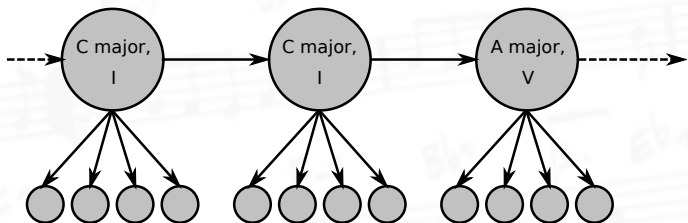
- Notes divided into small time units
- Model decides where new chords begin
- Notes emitted independently given harmonic label
- Notes represented as *pitch classes*
- Simple metrical model

## R&S Model: Emission Distribution



- Emission distribution shared across states
- Probabilities of notes relative to state:
  - triad
  - in-key notes
  - other notes
- Limited extension to tetrads

## R&S Model: Transition Distribution



- Transition distribution parameters tied
- Relative pitch:
  - Relative key changes
  - Chord transitions independent of key

# R&S Model: Training

- Unsupervised training by EM
- Initialize emission distribution



Triad notes most  
probable



Scale notes less probable



Out-of-key notes least probable

- Transition initialization makes no difference

# R&S Model: Training

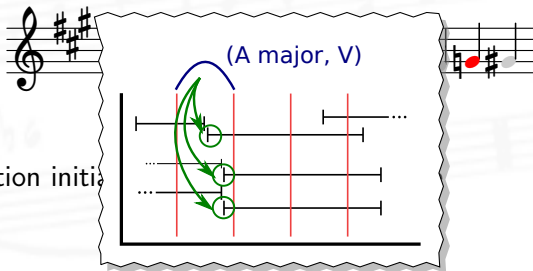
- Unsupervised training by EM
- Initialize emission distribution



Triad notes most probable



Scale notes less probable



- Transition initiali

# Our Experiments with the R&S Model

- Replicated original model
- Trained on data similar to R&S's: **Haydn**
- Trained on jazz data: **Jazz**
- Tried initializing transition distributions: **JazzFun**

# R&S Model: Evaluation

- R&S present no empirical evaluation
- No single correct analysis: many valid analyses
- Doesn't matter for now
- Parameters are easily interpretable

# R&S Model: Evaluation

Allegro con brio

The image shows a musical score for a piece titled "Allegro con brio". It consists of two systems of music. The first system has a treble clef staff with a melody and a bass clef staff with accompaniment. The melody starts with a piano (*p*) dynamic and then moves to a forte (*f*) dynamic. The second system continues the melody with triplets and other rhythmic patterns.

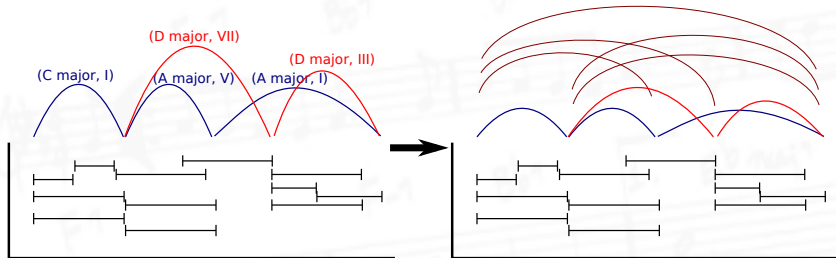
The image shows a musical score for the song "I love my Mister man". It features a treble clef staff with a melody and a bass clef staff with accompaniment. The melody includes the lyrics "I love my Mister man". Above the treble staff, there are three guitar chord diagrams labeled F7, Bb7, and Eb. The score is in a key with two flats and a 4/4 time signature.

- Informal conclusions from the models' output and distributions
- Tried jazz models on jazz MIDI files
- R&S hand-tweaked transition distribution: I haven't

## R&S Model: Conclusions

- Transition distributions learned some aspects of chord function
- Inevitably noisy
- Transition initialization makes no difference
- Metrical model weaker for jazz, but still meaningful
- Tetrads (dominant seventh) helped in some cases, but not all
- Triad model was confused by use of tetrads in jazz
- Worth incorporating tetrads better

# A MIDI Supertagger



- Early ideas for a MIDI supertagger
- Model structure similar to R&S
- First, most direct adaption of R&S to the task
- Then two possible model structures

# A Naive Adaptation of R&S



- States  $\rightarrow$  lexical categories
- Lexicon of chord grammar:

Ton.	$X$	$:= I^T$
Dom.	$X^7$	$:= I^D / IV^D   T$
Dom-tritone.	$X^7$	$:= \flat V^D / VII^D   T$
	$\vdots$	

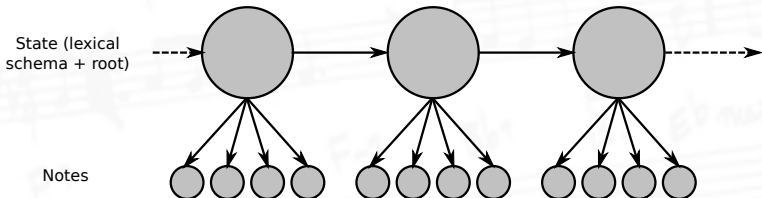
*Schema name    Chord class    Syntactic category*

- State: (*lexical schema, root pitch*)
- E.g. (Dom,  $A\flat$ )  $\Rightarrow$  category  $A\flat^D / D\flat^T | D$



## A Naive Adaptation of R&S

$$(\text{Dom}, Ab) \Rightarrow Ab^D / Db^{T|D}$$



- Relative pitch  $\rightarrow$  only change in chord root
- Emissions conditioned on lexical schema, not root
- Emission distribution over notes relative to state's root
- R&S metrical model less useful for jazz: try with and without

# The First Model

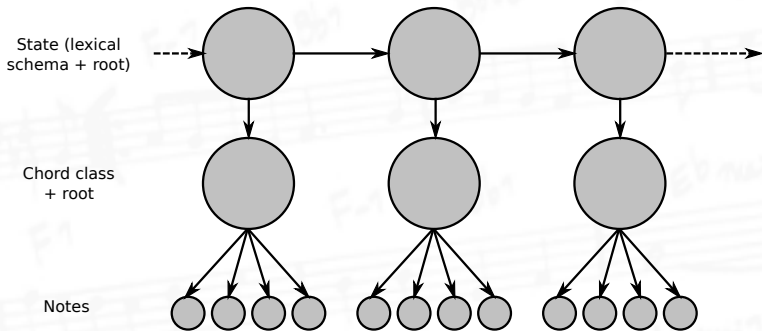


- Emission distribution for every lexical schema
- But we know some interpretations are ambiguous

$$\begin{array}{l}
 \vdots \\
 \text{Dom.} \quad X^7 \quad := I^D / IV^D | T \\
 \text{Dom-tritone.} \quad X^7 \quad := \flat V^D / VII^D | T \\
 \vdots
 \end{array}$$

- These categories should share emission distributions

# The First Model



- Many states share a chord class
- Emissions are conditioned only on this
- 5 emission distributions

# Chord Classes Are Too General



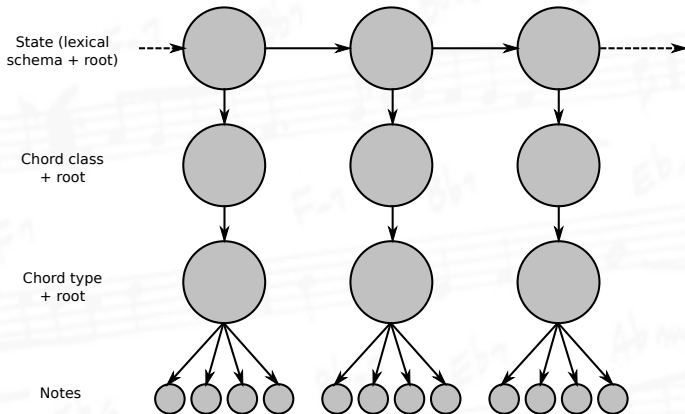
- In chord grammar, a chord class groups together chord types

$$X^7 \text{ with root } C = \{\dots, C^7, C^{aug7}, \dots\}$$

- $C^7$  includes a G
- $C^{aug7}$  includes a G $\sharp$
- G and G $\sharp$  shouldn't be generated independently
- Chord classes are too general to condition emission distributions on
- We should condition on *chord type*



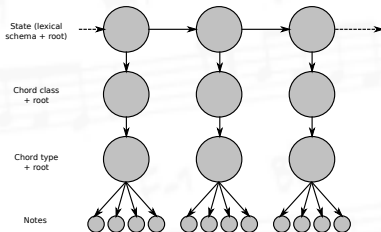
# The Second Model



- Chord types shared by chord classes



## The Second Model

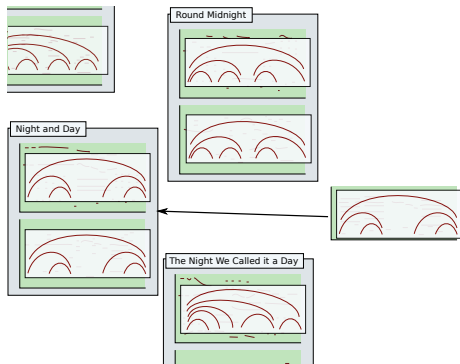


- Emission distribution for each tetrad
- Can initialize as R&S did
- Initialization of transition distributions will become important
- Use labelled corpus

# Song Identification Task

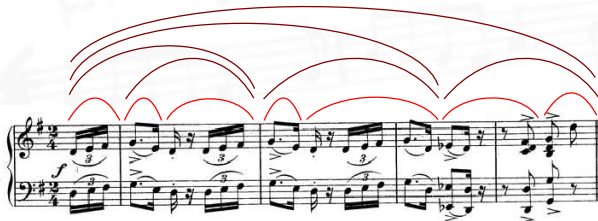
- Need to hear harmonic similarity to recognise songs
- Harmonic analysis can be used as a signal for similarity metric
- Use MIDI corpus: several files per song
- Identify new performance by harmonic similarity
- Compare to using metric over R&S analysis
- Compare to non-harmonic similarity

## Song Identification Task



- Need to hear 1
- Harmonic analysis can be used as a signal for similarity metric
- Use MIDI corpus: several files per song
- Identify new performance by harmonic similarity
- Compare to using metric over R&S analysis
- Compare to non-harmonic similarity

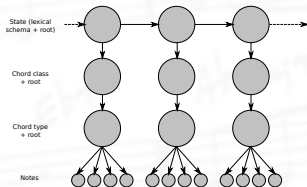
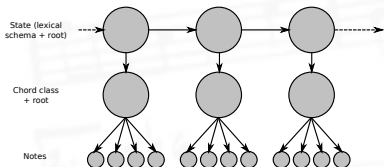
# Conclusion



- Musical similarity depends on harmonic analysis: parsing
- Grammatical model
- Can be extended to the harder, harmonic parsing task
- Ideas for MIDI supertagger, using R&S

## Next steps

- Try out supertagging models
- Combine with parser from chord-parsing model
- MIDI song-identification task
- Baseline model
- Compare system using parser to one using R&S analysis



# Bibliography

- Hockenmaier, J. (2001). Statistical parsing for CCG with simple generative models. In *Association for Computational Linguistics 39th annual meeting and 10th conference of the European Chapter*, vol. 39, (pp. 7–12).
- Raphael, C., & Stoddard, J. (2004). Functional harmonic analysis using probabilistic models. *Computer Music Journal*, 28(3), 45–52.



# Thesis

- Music has a syntax like that of natural language
- Parsing can reveal structures important to musical understanding
- This is useful for music processing tasks

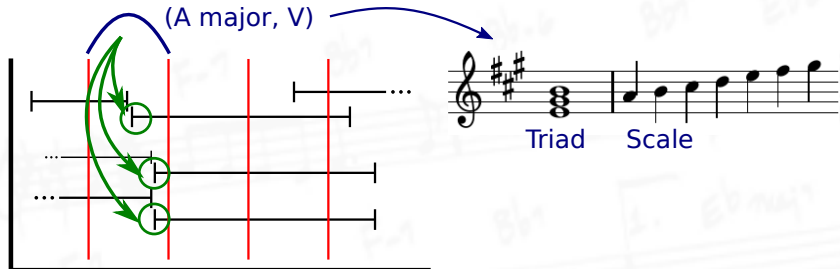
# Grammatical Harmonic Analysis Results

<i>Model</i>	<i>P (%)</i>	<i>R (%)</i>	<i>F (%)</i>	<i>Cov. (%)</i>
St+Grammar	<b>89.9</b>	61.9	73.3	75
HmmPath	74.6	82.1	78.2	100
St+Grammar+HmmPath	81.7	88.0	84.7	100
St+PCFG	87.4	85.7	86.5	95
St+PCFG+HmmPath	85.3	<b>91.0</b>	<b>88.1</b>	100

Evaluation of each model's prediction of tonal space paths using 10-fold cross-validation on the jazz corpus. For each model, we report *precision* (P), *recall* (R), *F-score* (F) and *coverage* (Cov.).

# R&S Model: Emission Distribution

(A major, V)



- Emission distribution shared across states
- Learn probabilities of relationship of a note to the state label
- Probability of:
  - each note of triad
  - any scale note
  - any other notes
- Limited extension to tetrads

## R&S Model: Transition Distribution

(A major, V) → (A major, I) → (D minor, I)  
(B major, V) → (A major, III) → (D minor, I)

- Transition distribution simplified to reduce parameters
- Only relative pitch:
  - Only take relative pitch between keys
  - Chord probabilities not dependent on key
- Ignore previous chord at key changes

# R&S Model: Evaluation

- R&S present no empirical evaluation
- No single correct analysis, but many valid interpretations
- Some ideas:
  - Manually mark output valid or invalid – time consuming
  - Evaluate performance on a task (MIR)

# Jazz Lexicon Extract

Ton.	$X(m)$	$:= I^T$
Ton-III.	$X_m$	$:= bVI^T$
Dom.	$X(m)^7$	$:= I^D/IV^D T$
Dom-tritone.	$X(m)^7$	$:= bV^D/VII^D T$
Subdom.	$X(m)$	$:= I^S/V^S T$
Dim-bVII.	$X_o$	$:= IV^D/bVII^D T$
Colour-IVf.	$X(m)$	$:= V^T/V^T$
Colour-IVb.	$X(m)$	$:= V^T \setminus V^T$



## Tonal Space

$\sharp I^-$	$\sharp V^-$	$\sharp II^-$	$\sharp VI^-$	$\sharp III$	$\sharp VII$	$\sharp\sharp IV$	$\sharp\sharp I$	$\sharp\sharp V^+$	$\sharp\sharp II^+$	$\sharp\sharp VI^+$
$VI^-$	$III^-$	$VII^-$	$\sharp IV^-$	$\sharp I$	$\sharp V$	$\sharp II$	$\sharp VI$	$\sharp III^+$	$\sharp VII^+$	$\sharp\sharp IV^+$
$IV^-$	$I^-$	$V^-$	$II^-$	$VI$	$III$	$VII$	$\sharp IV$	$\sharp I^+$	$\sharp V^+$	$\sharp II^+$
$b II^-$	$b VI^-$	$b III^-$	$b VII^-$	$IV$	$I$	$V$	$II$	$VI^+$	$III^+$	$VII^+$
$bb VII^-$	$b IV^-$	$b I^-$	$b V^-$	$b II$	$b VI$	$b III$	$b VII$	$IV^+$	$I^+$	$V^+$
$bb V^-$	$bb II^-$	$bb VI^-$	$bb III^-$	$bb VII$	$b IV$	$b I$	$b V$	$b II^+$	$b VI^+$	$b III^+$
$bbb III^-$	$bbb VII^-$	$bb IV^-$	$bb I^-$	$bb V$	$bb II$	$bb VI$	$bb III$	$bb VII^+$	$b IV^+$	$b I^+$



# Extra Slides

Thesis

Grammar Results

R&S Distributions

R&S Evaluation

Lexicon

Tonal Space