Christopher Raphael & Joshua Stoddard
Dept. of Math + Stats
Univ. of Massachusetts, Amherst
Oct. 29, 2003

Goal: Given "score-equivalent" data, construct functional harmonic analysis:

$$\underbrace{\mathrm{I}\quad \mathrm{IV}\quad \mathrm{V}\quad \mathrm{I}}_{\text{C maj}}\quad \underbrace{\mathrm{V}\quad \mathrm{I}\quad \mathrm{V}\quad \mathrm{I}}_{\text{F maj}}\cdots$$

Applications:

1. Harmony-based music queries

2. automatically generated one-dimensional music representation (similarity measures)

Music can be partitioned in segments have constant key:

$$k = (t, m) \in \{0, \ldots, 11\} \times \{\text{Major}, \text{Minor}, \ldots\}$$

Key has associated scale

Major	$\underline{0}$	1	$\underline{2}$	3	$\underline{4}$	$\underline{5}$	6	$\underline{7}$	8	$\underline{9}$	10	$\underline{11}$	$\underline{12}$
Minor	$\underline{0}$	1	$\underline{2}$	$\underline{3}$	4	$\underline{5}$	6	$\underline{7}$	$\underline{8}$	9	10	$\underline{11}$	$\underline{12}$

Triad = alternating triple of scale tones

Major Traids:

1	I	0	4	7
2	(ii)	2	5	9
3	(iii)	4	7	11
4	(IV)	5	9	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
7	(vii)	11	2	5

Data: Music is divided into abstract time intervals (beats, measures, etc.)
Let $y_n = y_n^1, \ldots, y_n^K$ be pitches in $\{0, \ldots, 11\}$ in nth time interval.
Object: Find harmonic label for each interval:

$$n\text{th label} = x_n \in \underbrace{\{0, \ldots, 11\}}_{\text{tonic}} \times \underbrace{\{\text{Major}, \text{Minor}\}}_{\text{mode}} \times \underbrace{\{1, \ldots, 7\}}_{\text{chord}}$$

Let $X_1, \ldots, X_N$ be sequence of harmonic labels. Assume X is Markov chain:

$$p(x_n | x_1, \ldots x_{n-1}) = p(x_n | x_{n-1})$$

Is this reasonable?

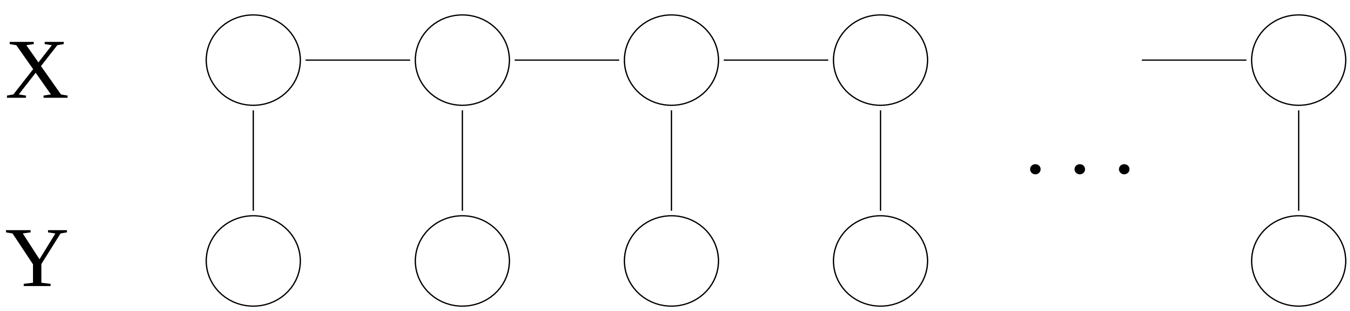
1. keys tend to persist for long periods of time
2. chords tend to persist, but not as long as keys
3. chords tend to move in familar patterns: boogie woogie, Pachelbel

Model data y_n as observation of random variable Y_n. Model

$$p(y_n | x_1, \ldots, x_n, y_1, \ldots y_{n-1}) = p(y_n | x_n)$$

Is this reasonable?

Graphical Depiction of Conditional Indepedencies:



Model has lots of parameters:

$$\begin{aligned} p(\underbrace{x'}_{12\times2\times7} \mid \underbrace{x}_{12\times2\times7}) &= \text{28224 parameters} \\ p(\underbrace{y}_{---} \mid \underbrace{x}_{12\times2\times7}) &= \text{many parameters too} \end{aligned}$$

A complication: Harmonic importance $\leftrightarrow$ measure position

For pitches $y = y^1, \ldots, y^K$ let $r = r^1, \ldots, r^K$ describe the measure positions:

$$r^k \in \{0, \ldots, 5\}$$

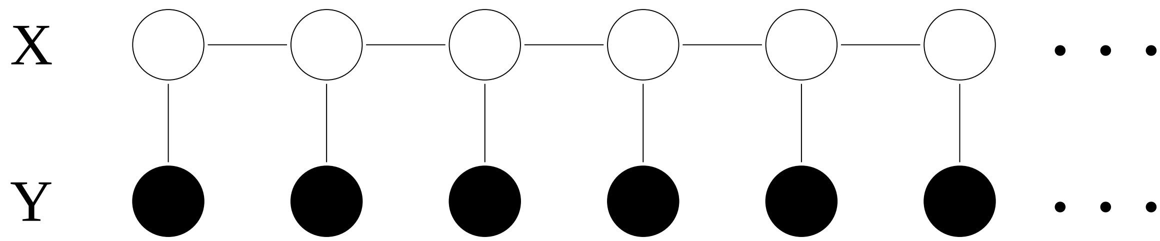
where lower values mean stronger measure positions.

Output probs modeled as $p(y|x,r)$.

$$\begin{aligned} p(y|x,r) &= p(y^1, \ldots y^K | x, r^1 \ldots, r^K) \\ &= \prod_{k=1}^{K} p(y^k | x, r^k) \\ &= \prod_{k=1}^{K} p(y^k | t, m, c, r^k) \\ &= \prod_{k=1}^{K} p(y^k - t | m, c, r^k) \\ &= \prod_{k=1}^{K} p(F(y^k - t, m, c) | r^k) \end{aligned}$$

$$\begin{aligned} p(x'|x) &= p(t', m', c' | t, m, c) \\ &= p(t', m' | t, m) p(c' | t', m', t, m, c) \\ &= p(t' - t, m' | m) \begin{cases} p(c'|c) & \text{if } t' = t, m' = m \\ p(c') & \text{otherwise} \end{cases} \end{aligned}$$

Forward Probabilities: $\alpha_n(x_n) = p(x_n, y_1, \ldots, y_n)$



$$
\begin{aligned}
\alpha_{n+1}(x_{n+1}) &= p(x_{n+1}, y_1, \ldots, y_{n+1}) \\
&= \sum_{x_n} p(x_n, x_{n+1}, y_1, \ldots, y_{n+1}) \\
&= \sum_{x_n} p(x_n, y_1, \ldots, y_n) p(x_{n+1}|x_n) p(y_{n+1}|x_{n+1}) \\
&= \sum_{x_n} \alpha_n(x_n) p(x_{n+1}|x_n) p(y_{n+1}|x_{n+1})
\end{aligned}
$$

Backward Probabilities: $\beta_n(x_n) = p(x_n, y_{n+1}, \ldots, y_N)$ computed analogously.
Then

$$p(x_n|y_1, \ldots, y_N) = \frac{\alpha(x_n)\beta(x_n)}{\Sigma_{x'_n} \alpha(x'_n)\beta(x'_n)}$$

Could restimate $p(y|x)$, for example, by

$$\hat{p}(y|x) = \frac{\Sigma_{y_n=y}\, p(x_n = x|y_1, \ldots, y_N)}{\Sigma_n\, p(x_n = x|y_1, \ldots, y_N)}$$

Forward-Backward algorithm alternates between computing α and β and reestimating output and transition probabilities.

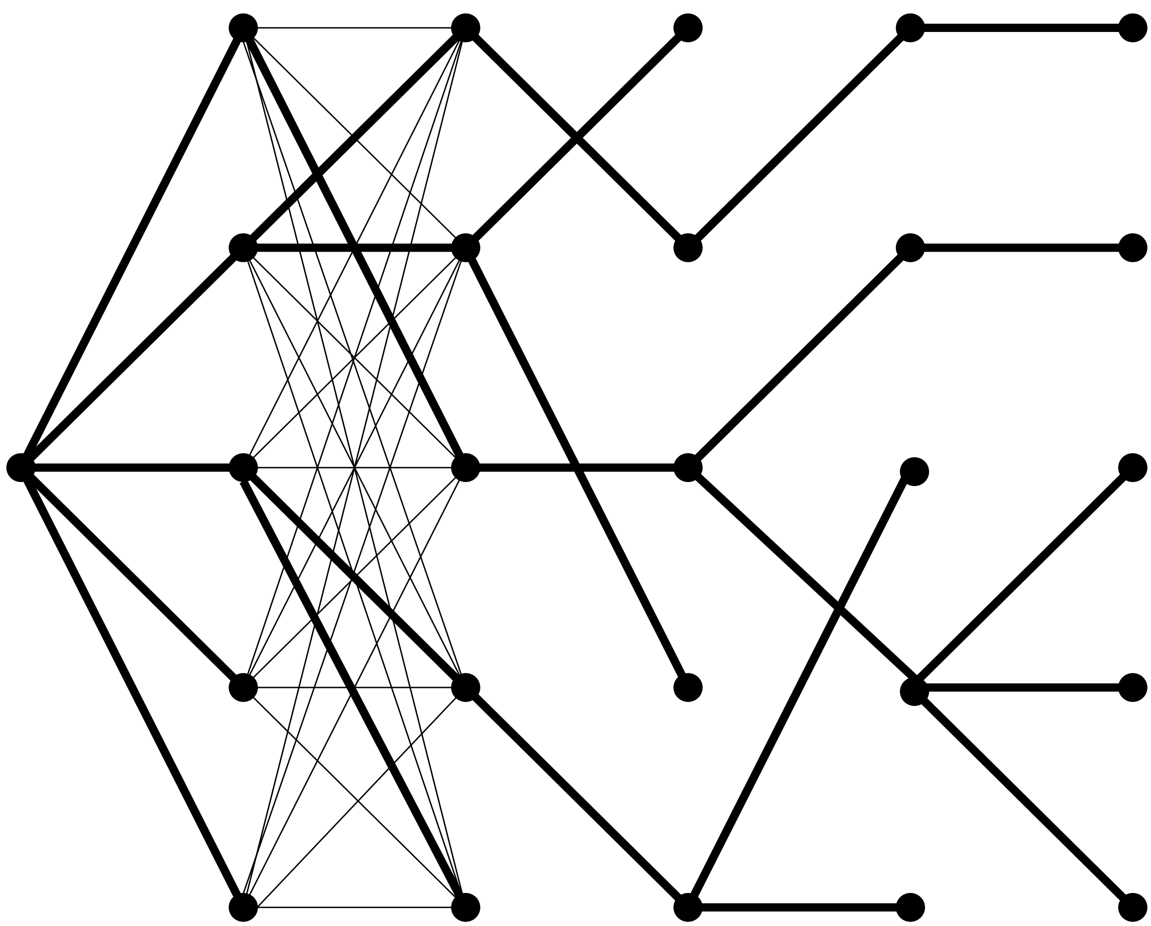
Likely Parse

Seek

$$\arg\max_x p(x|y) = \arg\max_x p(x,y)$$

$p(x,y) = p(x_1, \ldots, x_N, y_1, \ldots, y_N)$ factors as

$$\begin{aligned}
p(x,y) &= p(x)p(y|x) \\
&= p(x_1)p(x_2|x_1)\ldots, p(x_N|x_{N-1})p(y_1|x_1)\ldots p(y_n|x_n) \\
&= \underbrace{p(x_1)p(y_1|x_1)}_{f_1(x_1)} \underbrace{p(x_2|x_1)p(y_2|x_2)}_{f_2(x_1,x_2)} \ldots \underbrace{p(x_N|x_{N-1}p(y_N|x_N)}_{f_N(x_{N-1},x_N)}
\end{aligned}$$



1. Haydn Piano Sonata
2. Harder: Chopin "Raindrop Prelude"
3. Still Harder: Debussy Prelude (*Suite Bergamasque*)

More sophisticated choice: Model $p(y_n|x_n y_{n-1})$