Conditionally Independent Voice Model

Christopher Raphael, Gabi Teodoru, & Eric Nichols
School of Informatics
Indiana Univ., Bloomington
Dec. 7, 2007

"Logjam" in moving from monophonic to polyphonic music




Conditionally Independent Voice Model

Christopher Raphael, Gabi Teodoru, & Eric Nichols
School of Informatics
Indiana Univ., Bloomington

Dec. 7, 2007

The Model

Whereas:

- Music composed of voices
- Each voice has internal logic governing its evolution
- Voices do not evolve independently, but rather through shared attributes: harmony, meter, phrase structure, etc.

1. We model voices as *conditionally independent* Markov chains driven by an additional collection of variables (also a Markov chain).
2. *Everything* governing interaction between voices contained in driving chain.

Examples:

Pitch spellings of each voice	cond. indep. MC's given	functional harmonic analysis
Expressive inflections of each voice	cond. indep. MC's given	overall expressive intent
Notes and rhythms of each voice	cond. indep. MC's given	overall musical plan (harmony, structure, affect, etc.)

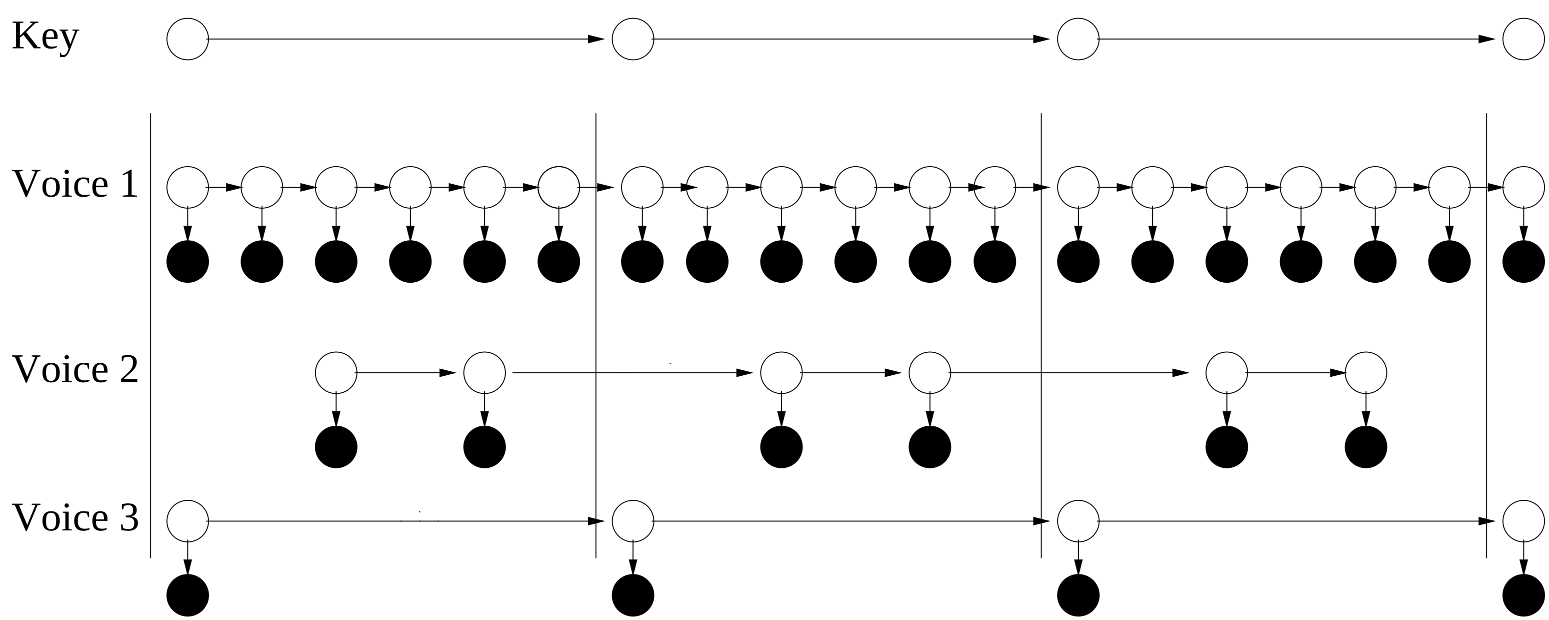
Probabilistic Model for Pitch Spelling

Convert **Midi Pitch #'s ⟶ Pitch Spellings** ($G\#$ vs. $A\flat$) (Why do this?)
(Temperley+Sleator, Meredith, Cambouropoulos, Chew+Chen, Longuet-Higgins ...)

Principles:

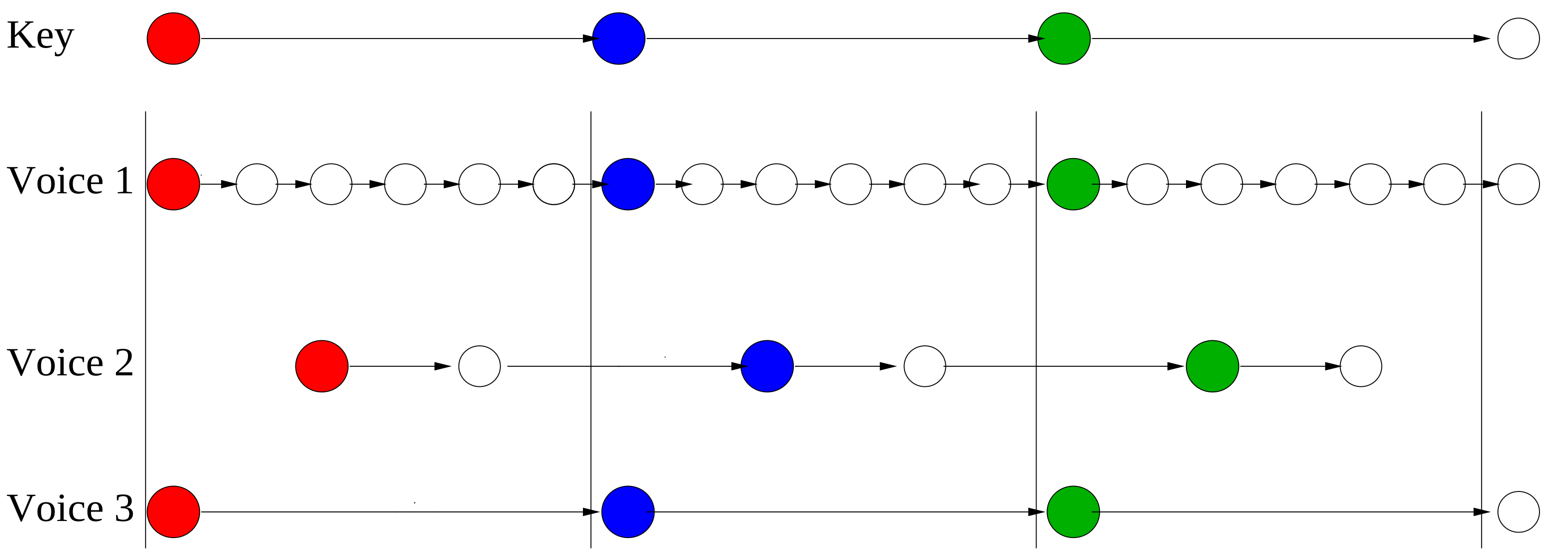
1. Vertical: Spell notes according to ambient key
2. Horizontal: Incorporate voice leading tendencies

- Key sequence modeled as MC of (tonic,mode) pairs
  [keys tend to persist, or move to neighbors]
- Each voice modeled as *solfege* MC: $\{1, \sharp 1, \flat 2, 2, \sharp 2, \flat 3, 3, 4, \sharp 4, \flat 5, 5, \sharp 5, \flat 6, 6, \sharp 6, \flat 7, 7\}$
  [$\sharp 2 \rightarrow 3$, $\flat 3 \rightarrow 2$]
- observable MIDI = f(Solfege, Key)



Computing Most Likely Configuration with Dynamic Programming

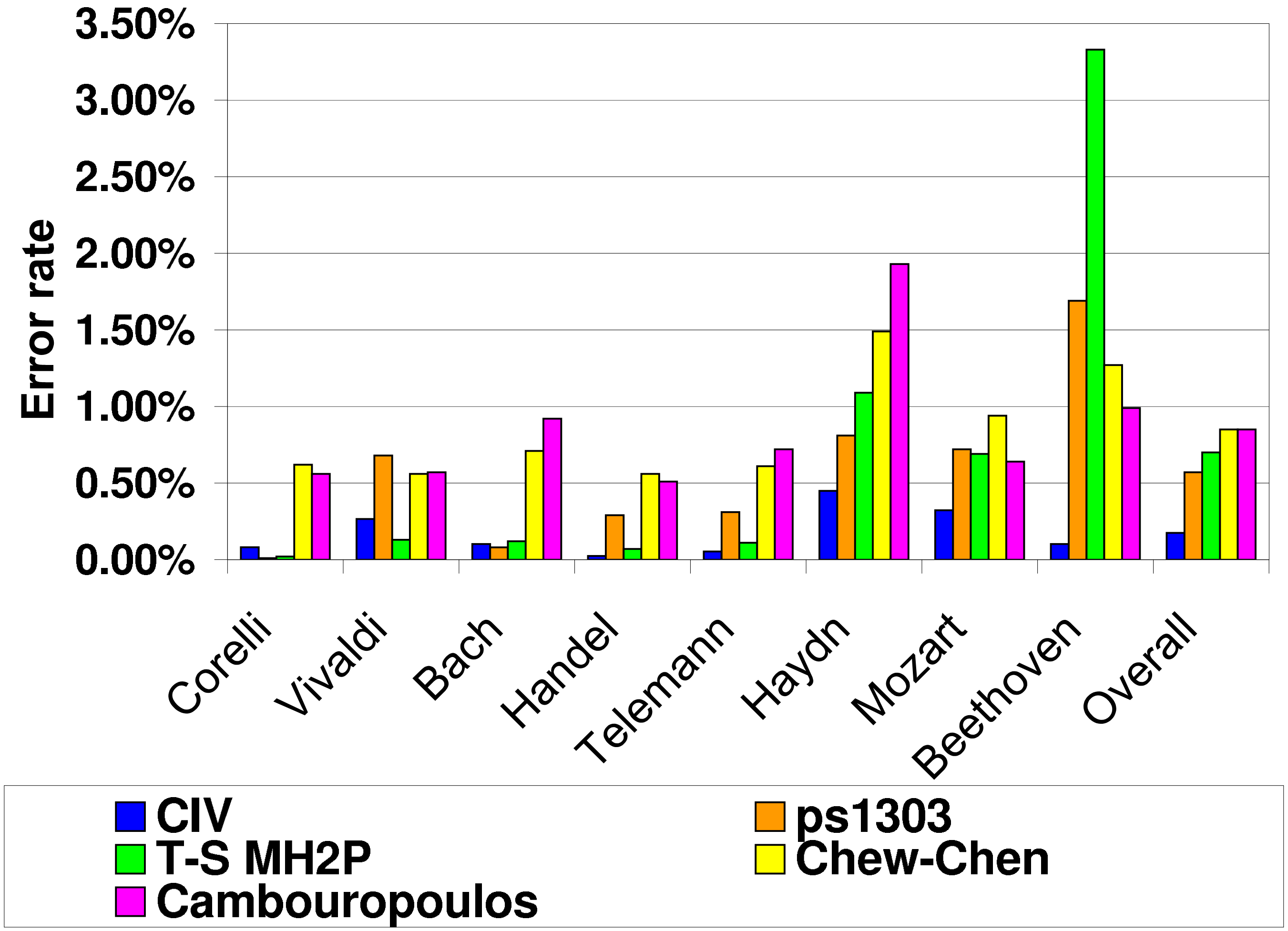
Easy to compute *most likely configuration* of hidden keys and solfege variables given MIDI pitches.



Key = D Maj. + Solfege = 3, then $F\sharp$.

Key = $D\flat$ Maj + Solfege = $\flat 6$, then $B\flat\flat$.

Results on CCARH Corpus



We still have trouble with

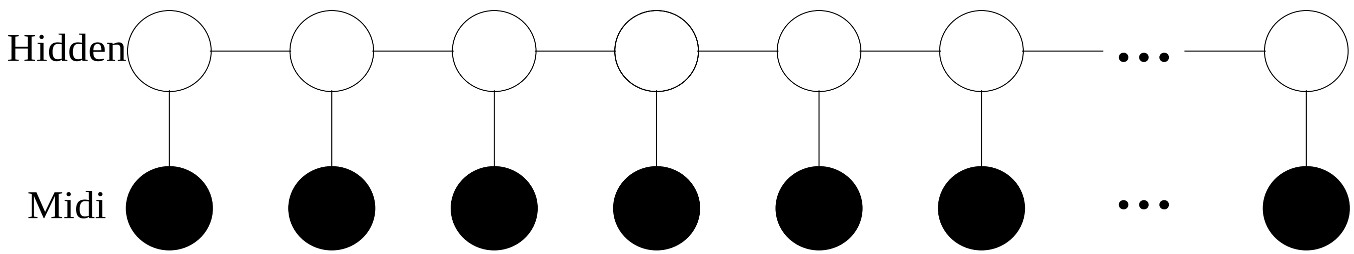
1. delayed resolution: $\#C \to E \to D$

2. spelling requiring knowledge of chord (eg. German Augmented 6th)

3. overly simplified latent harmonic state

Training the Model

- Can express model as HMM using *composite rhythm* as time-grid.
- Then train using the usual Baum-Welch algorithm.



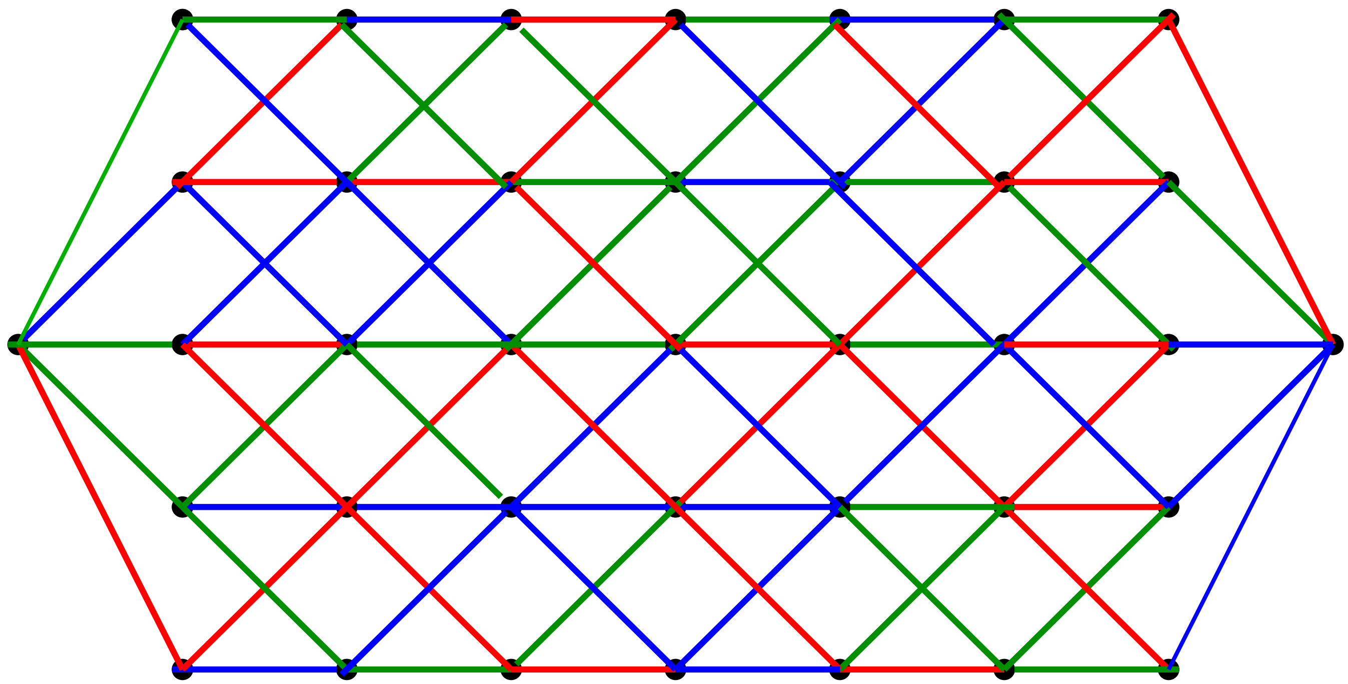
This led to **worse** results on test and *training*!

Why does this happen?

A New Training Paradigm

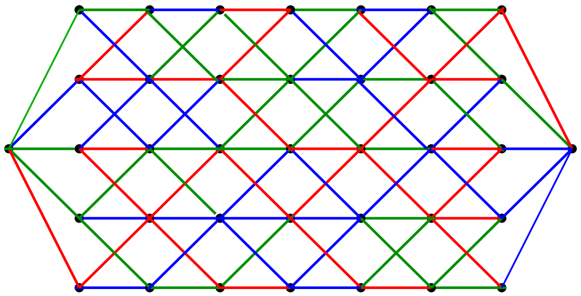
- Can think of note spelling task (or any HMM recognition problem) as search for best path through trellis.

- Now assume each arc score is affine function of several parameters: $\theta = (\theta_1, \ldots, \theta_M)$.



We will perform dynamic programming *simultaneously* over the entire parameter space.

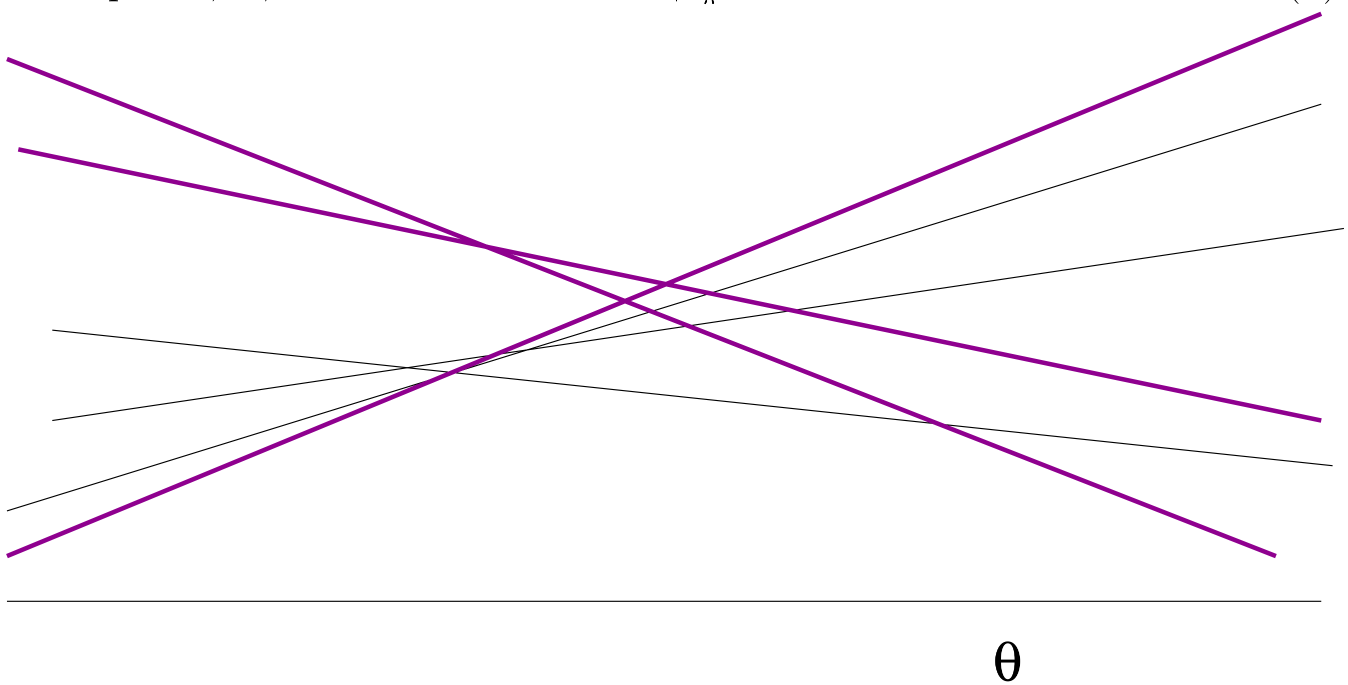
Crucial Observation



Fix a node in the trellis and consider a path π to node.

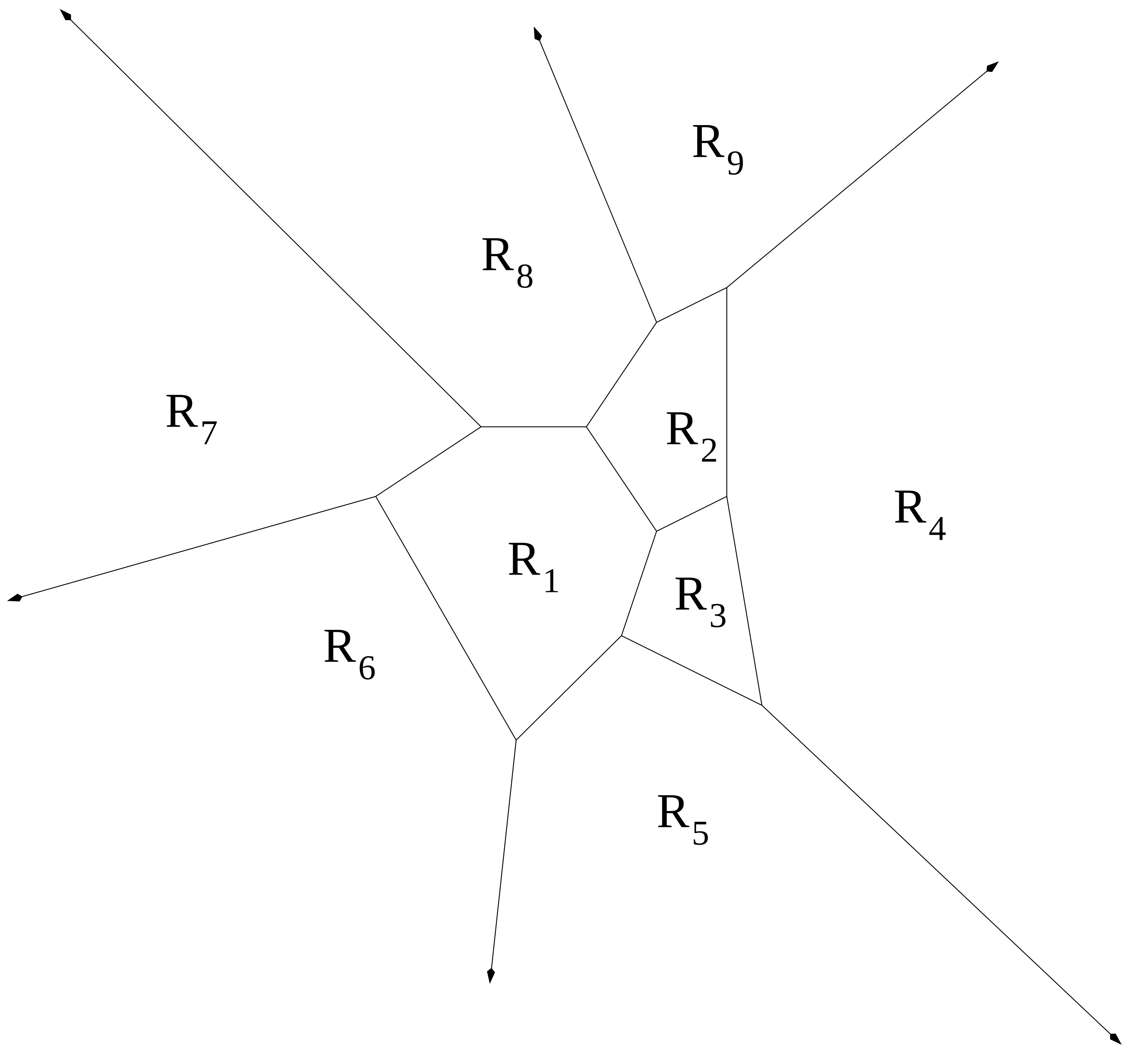
- Note: For path π, $S_\pi(\theta) = \beta_\pi^t \theta + b_\pi$
- The maximal score to node, as function of θ, is $S(\theta) = \max_\pi \beta_\pi^t \theta + b_\pi$

Note that some paths, π, have affine functions $\beta_\pi^t \theta + b_\pi$ don't "contribute to" $S(\theta)$.



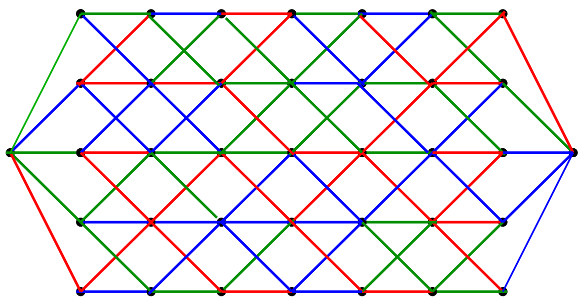
What does $S(\theta)$ look like in 2-D?

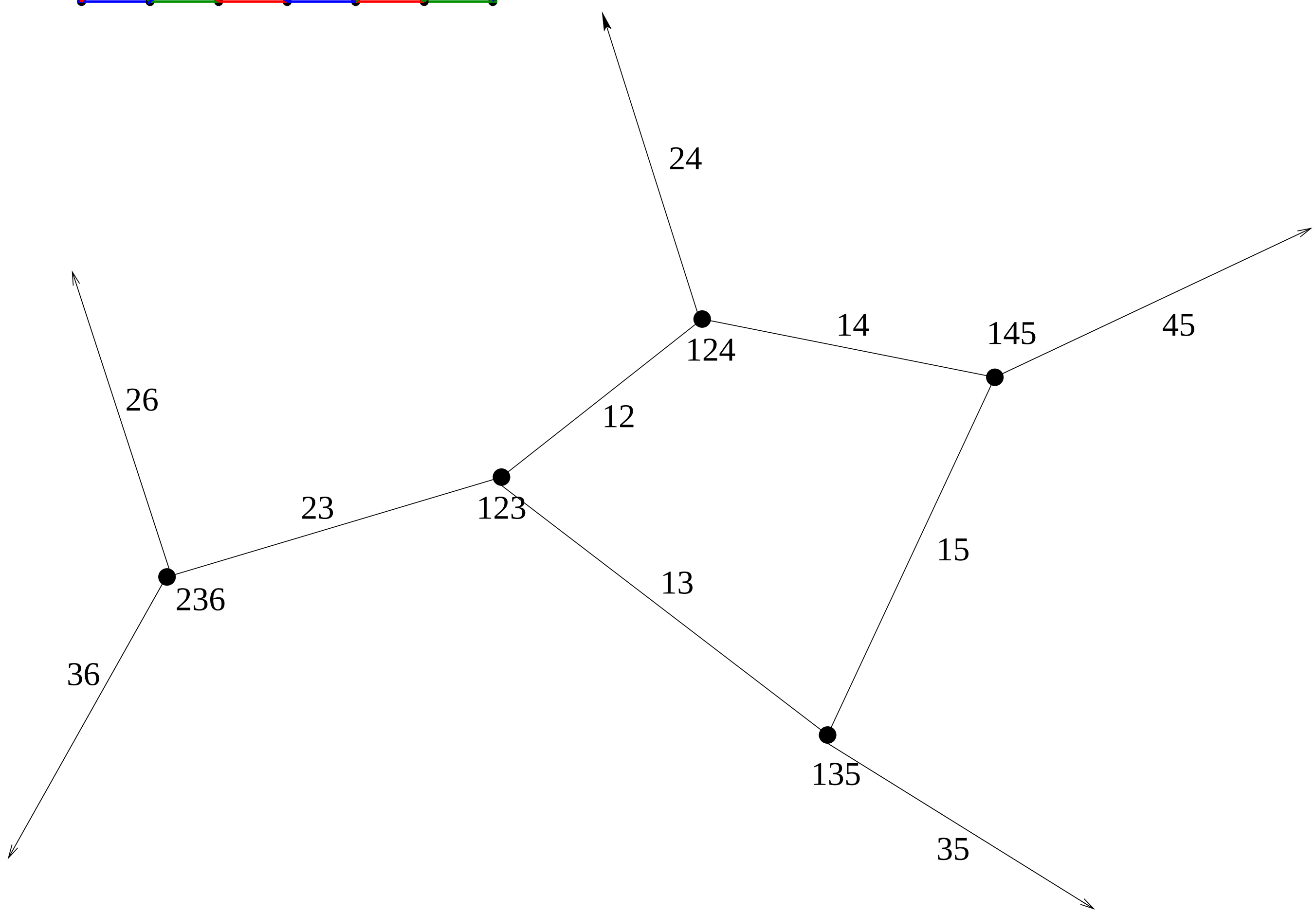
As maximum of affine functions, $S(\theta)$ is convex.



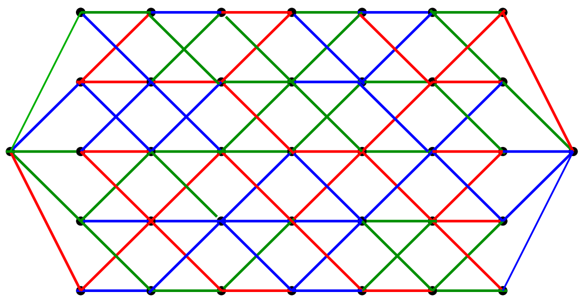
The DP algorithm

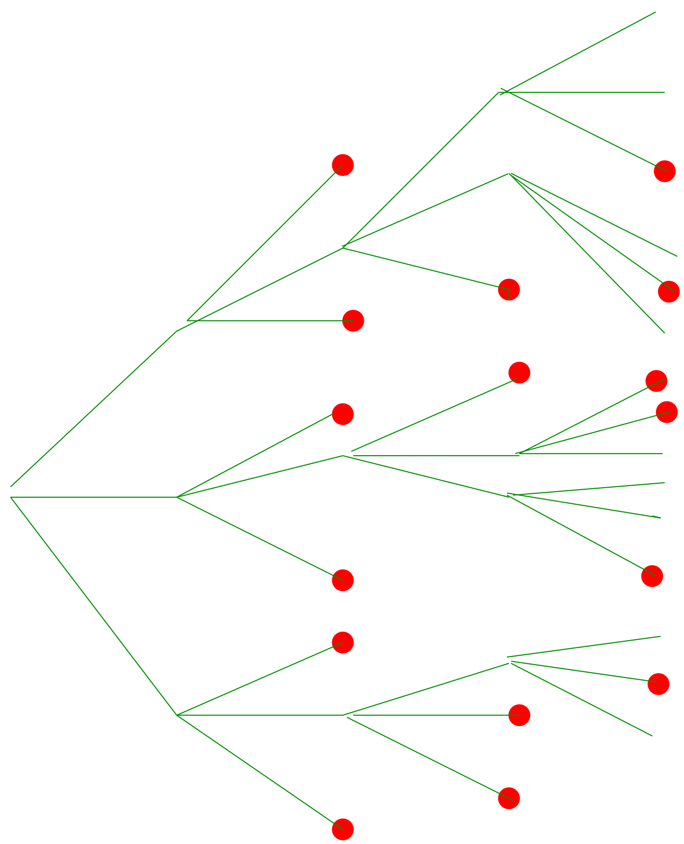
As in POMDPs, can carry through dynamic programming recursion with this functional form.





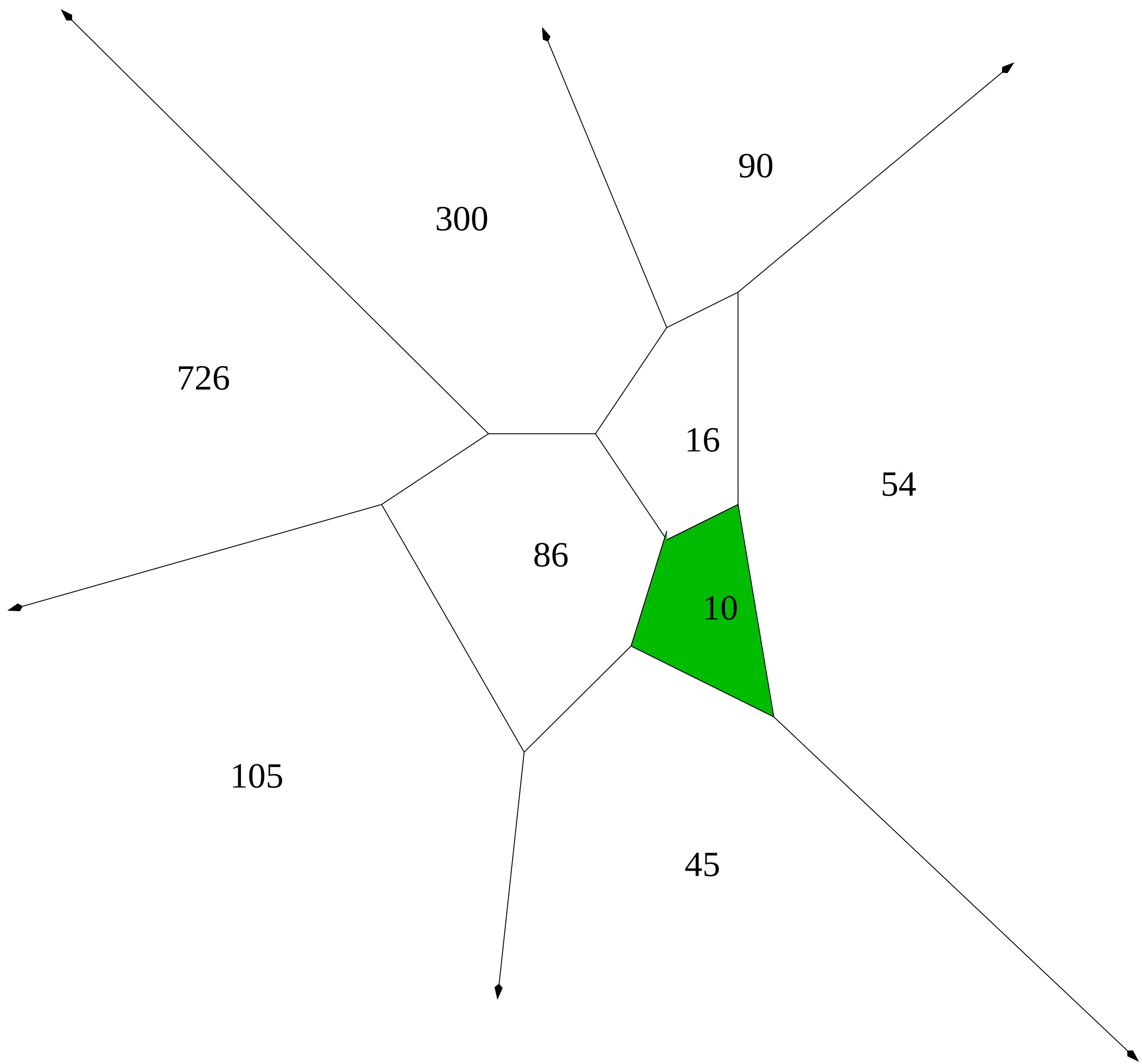
The DP search tree





Estimating θ

Reasonable choices of θ lie in regions corresponding to "attractive" paths, e.g. those making a minimum of "errors."



This Just In . . .

OPEN-RANK FACULTY POSITION IN MUSIC INFORMATICS

The Indiana University School of Informatics seeks to fill an open-rank faculty position in Music Informatics, starting August 2008. Indiana University is home to the internationally recognized Jacobs School of Music, as well as the Variations3 Digital Library project. These resources, combined with the vibrant and interdisciplinary environment of the School of Informatics, provide a unparalleled home for researchers pursuing music informatics.
etc.