Effectiveness of HMM-Based Retrieval on Large Databases

Jonah Shifrin
William Birmingham
University of Michigan, Ann Arbor

ISMIR 2003
Funded by NSF Grant IIS-0085945

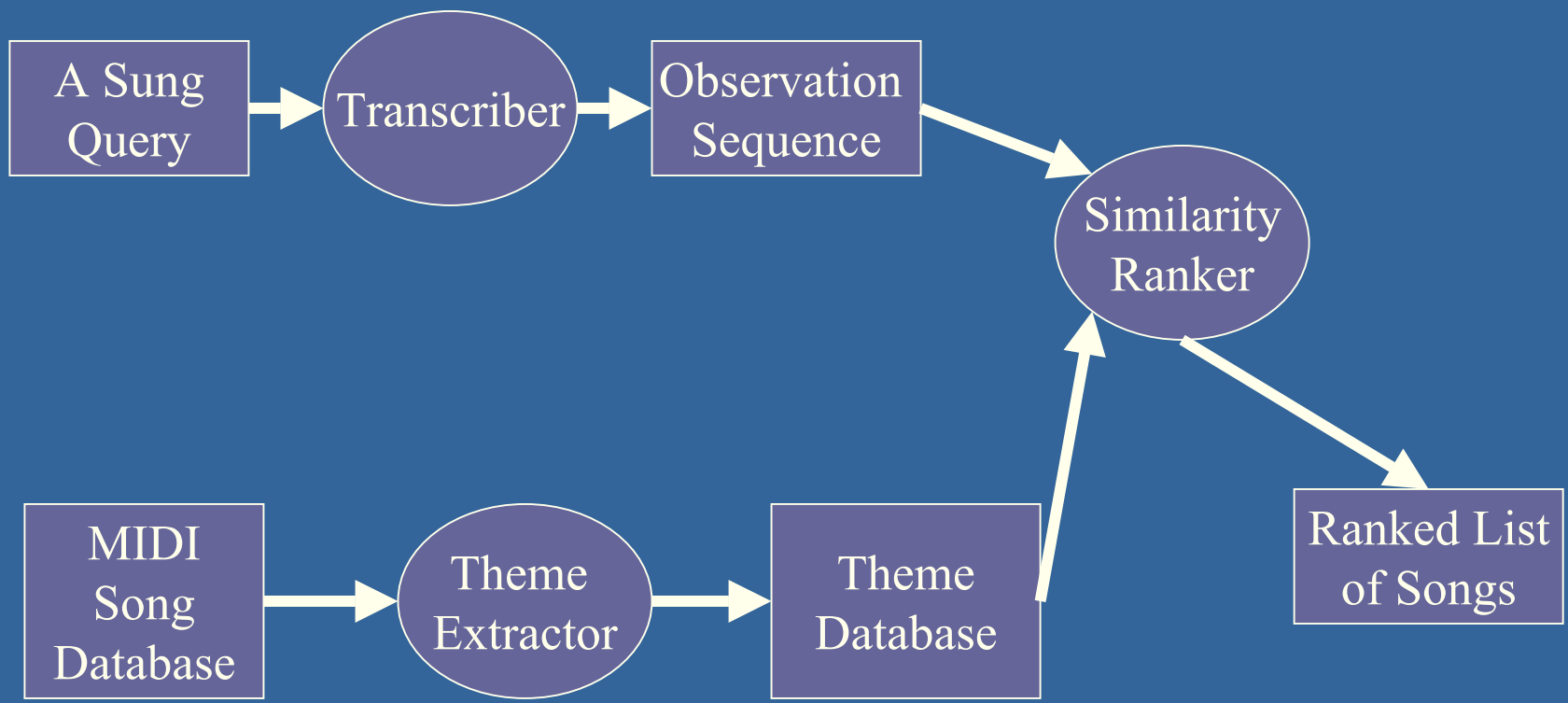
Performance on a large Database

- “real-world” system contain a large database
  - 100-gigabyte hard drive can hold 22,000 mp3s
- Are individual themes distinguishable at all?
- Can queries modeled for our system be retrieved?
- Will “real-world” modeled queries be identified by our system?

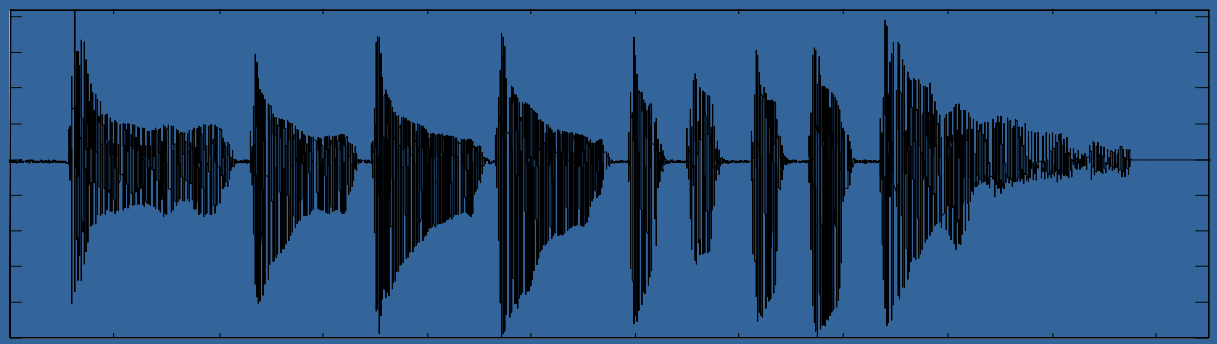
Presentation Outline

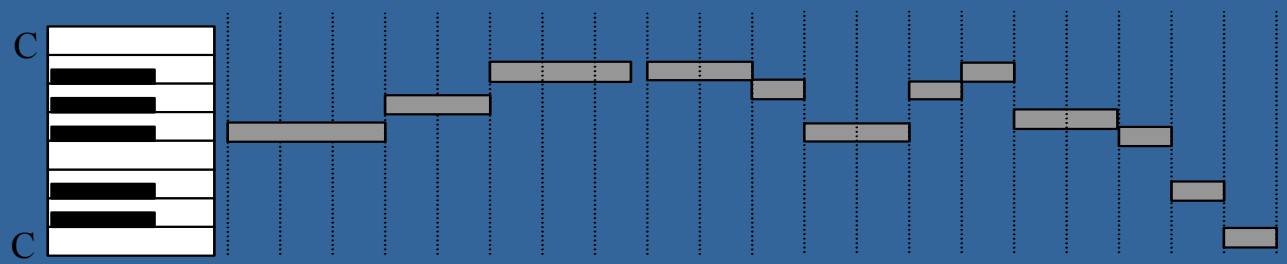
- Overview of System
- Large Database Creation
- Query Creation
- Experimental Results
- Conclusions

Introduction to Existing System



Sung Queries

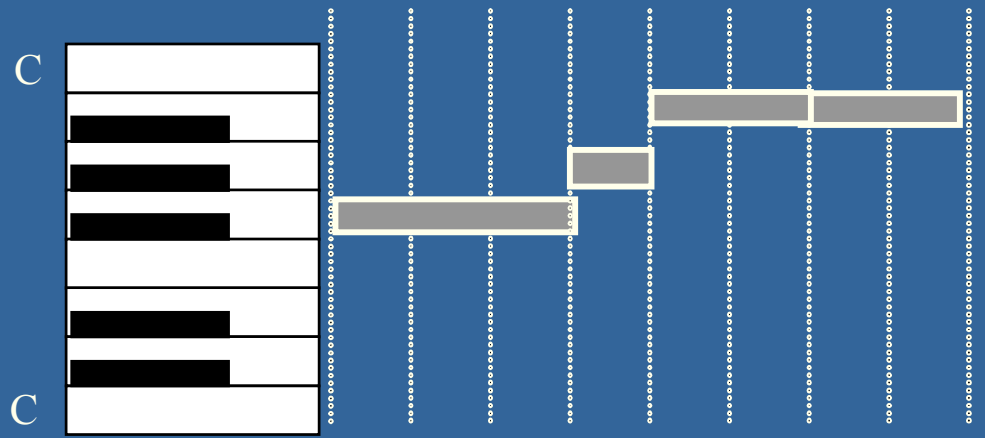




delta pitch	2	2	0	-2	-2	2	2	-4	-1	-3	-3
IOI (100 ms units)	3	2	3	2	1	2	1	1	2	1	1
IOI ratio	1.5	.66	1.5	2	.5	2	1	1	2	1	1

Query Representation

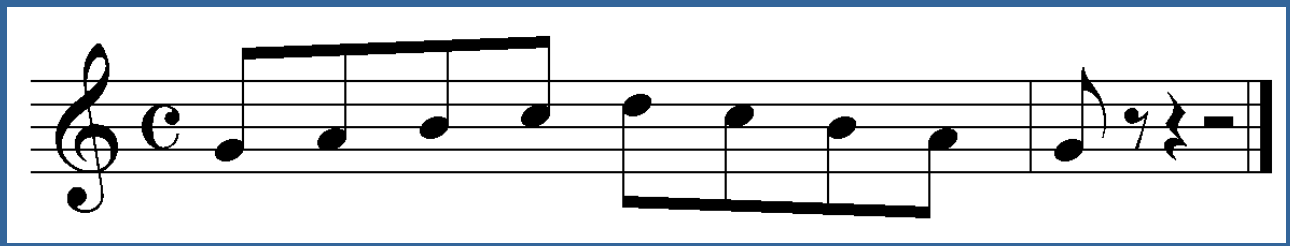
- Sequence of duples
  - Change in Pitch (Delta Pitch)
  - Rhythmic Ratio (Inter Onset Interval Ratio)



Delta Pitch	2	2	0	
IOI	3	1	2	2
IOI Ratio	3	0.5	1	

Targets – Themes

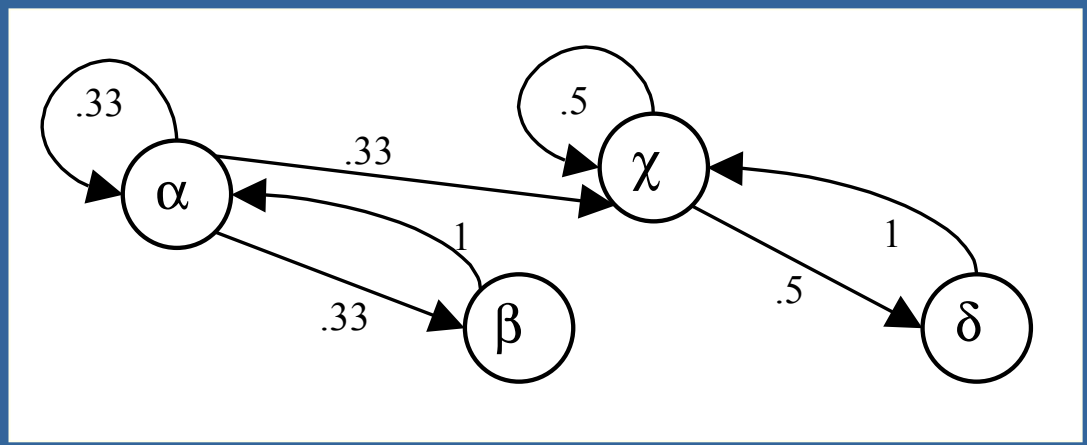
Theme



States

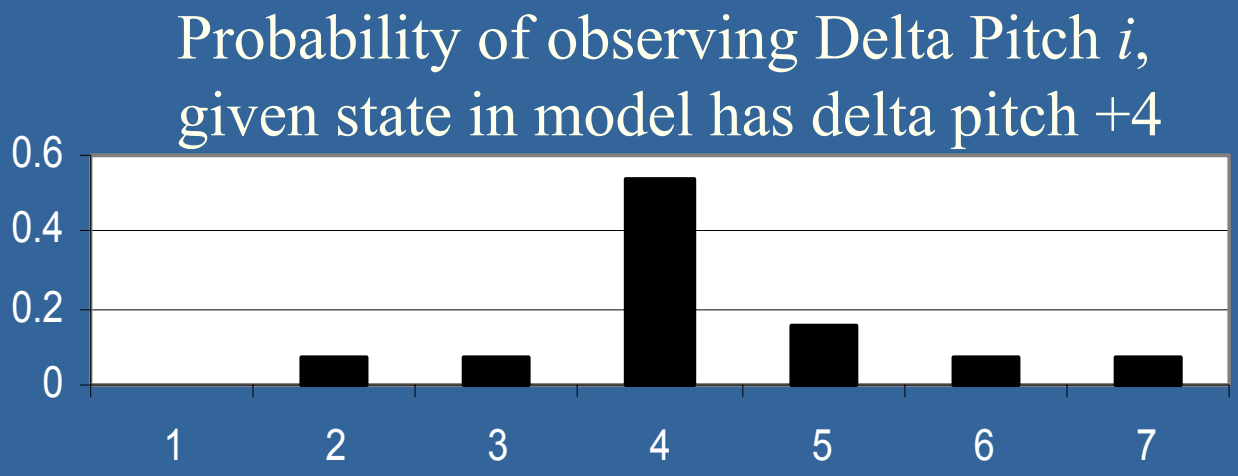
Delta pitch	2	2	1	2	-2	-1	-2	-2
IOI	1	1	1	1	1	1	1	1
IOI ratio	1	1	1	1	1	1	1	1
State	α	α	β	α	χ	δ	χ	χ

Markov Model



Matcher

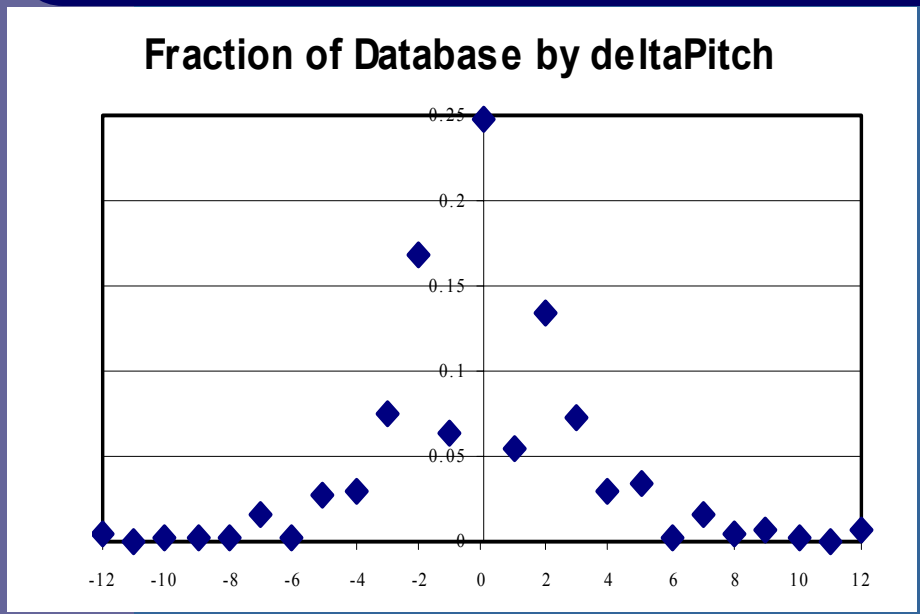
- Forward algorithm
  - determines probability target generated observation
- Pitch and duration assumed conditionally independent in observation tables

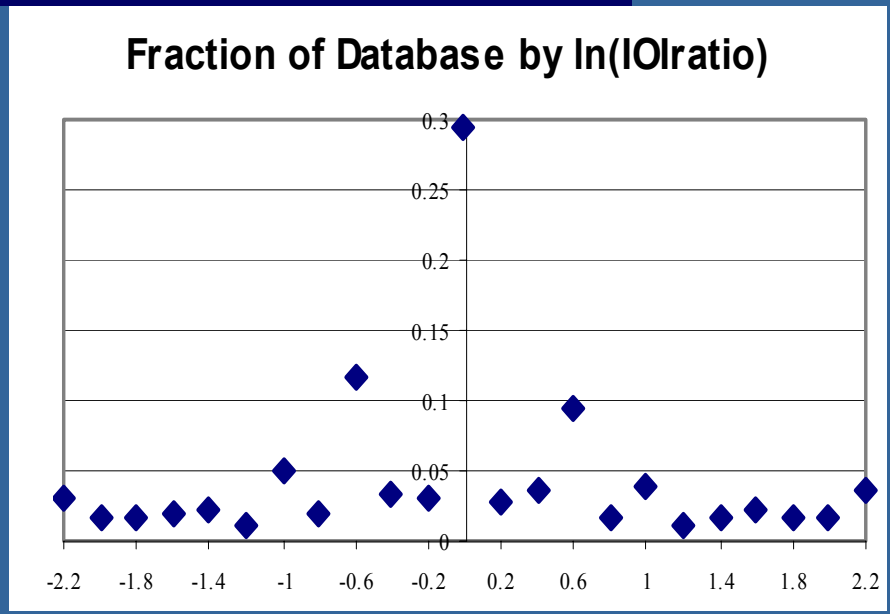


Synthetic Target Database

- Created synthetic database from Beatles database
  - Beatles Database has 284 themes from 260 songs
- Examined *length*, *deltaPitch* and *IOIratio*
- Median theme length is 40 notes
  - standard deviation of 20

Database Generation





- Created 50,000 themes

Synthetic Query Generation

- Three query sets containing 5100 queries
  - Query length ranges from 5 to 55 notes
  - 100 queries of each length
  - How many notes necessary to distinguish target?
- Queries are notes and durations
  - Equivalent to segmented, pitch-tracked sung queries

Query Set 1: Perfect Queries

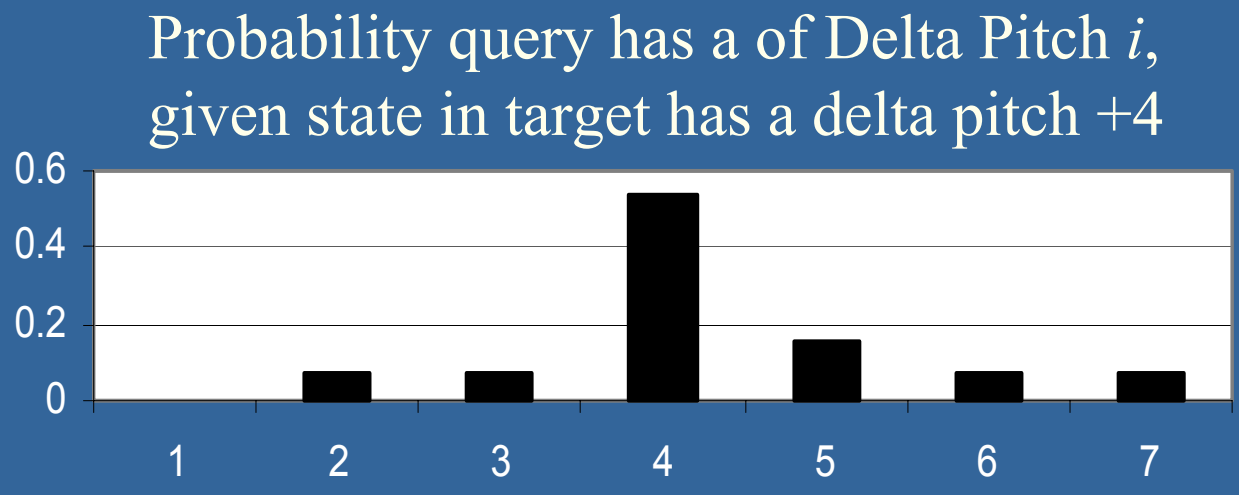
- Are individual themes distinguishable at all?
- Perfect queries created by exacting excerpt of length *n* from target in database
- Provide baseline

Query Set 2: Imperfect Queries

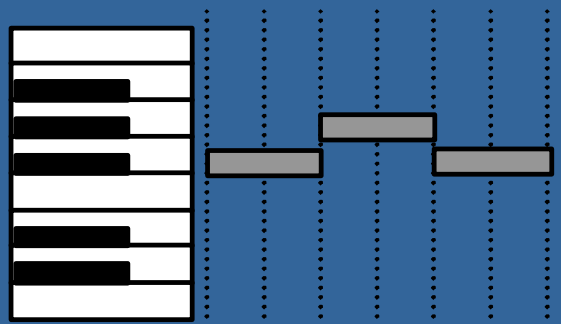
- Can queries modeled for our system be retrieved?
- Assumption: Perfect model of singer and transcription error
- Generate queries based on singer-error statistics
- Represent best-case real world scenario

Imperfect Query Generation

- Select a length-*n* subsequence in target
- Transform selected portion into duples of
  - <*deltaPitch*, *IOIratio*>
- Generate query from observation tables

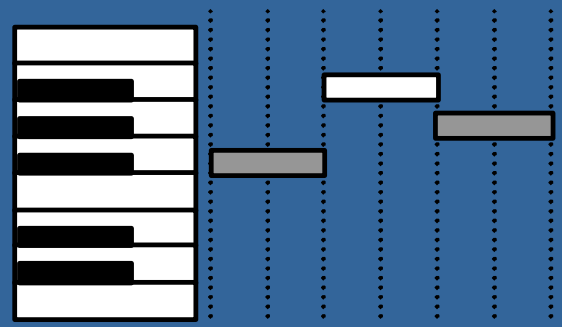


Imperfect Queries: Assume Cumulative Error only



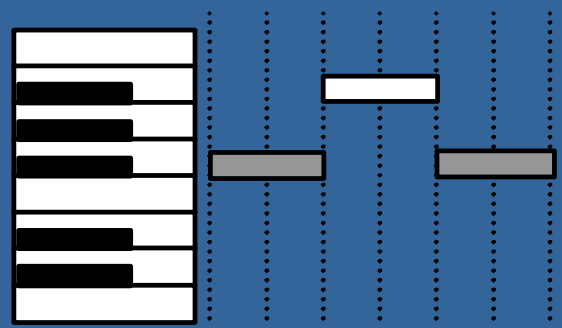
deltaPitch: +2 -2

Target on Piano Roll



deltaPitch: +4 -2

Query with Cumulative Error



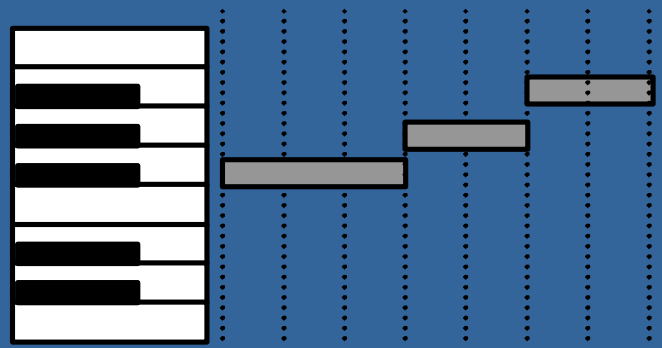
deltaPitch: +4 -4

Query with Local Error

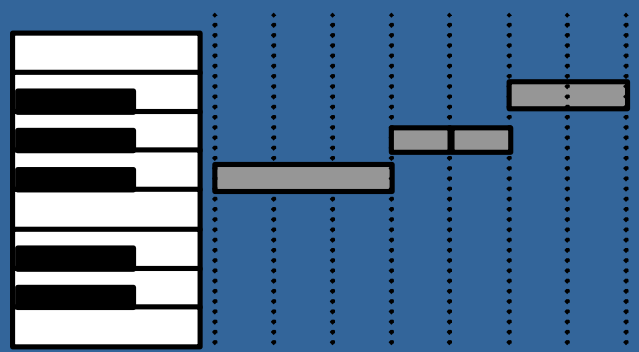
Query Set 3: Imperfect Queries with Insertions and Deletions

- Will “real-world” modeled queries be identified by our system?
- Insertions and Deletions present in real-world queries
- Sources:
  - People
  - pitch-trackers/segmenters

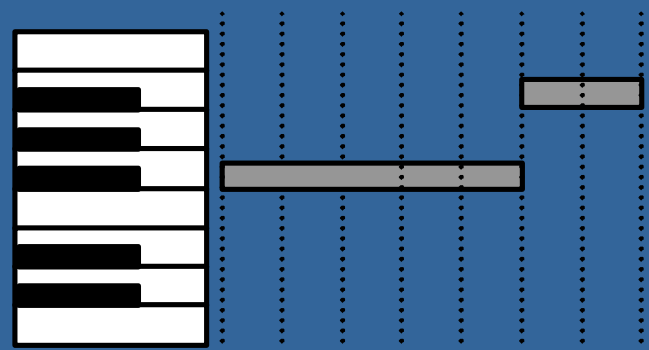
Insertions and Deletions



Query on Piano Roll



Insertion on Second Note



Deletion on Second Note

Insertion and Deletion Generation

- Used same method as Imperfect Queries
  - Insert or delete notes based on probabilities

P(“no edit) = 0.81

P(“insertion”) = 0.06

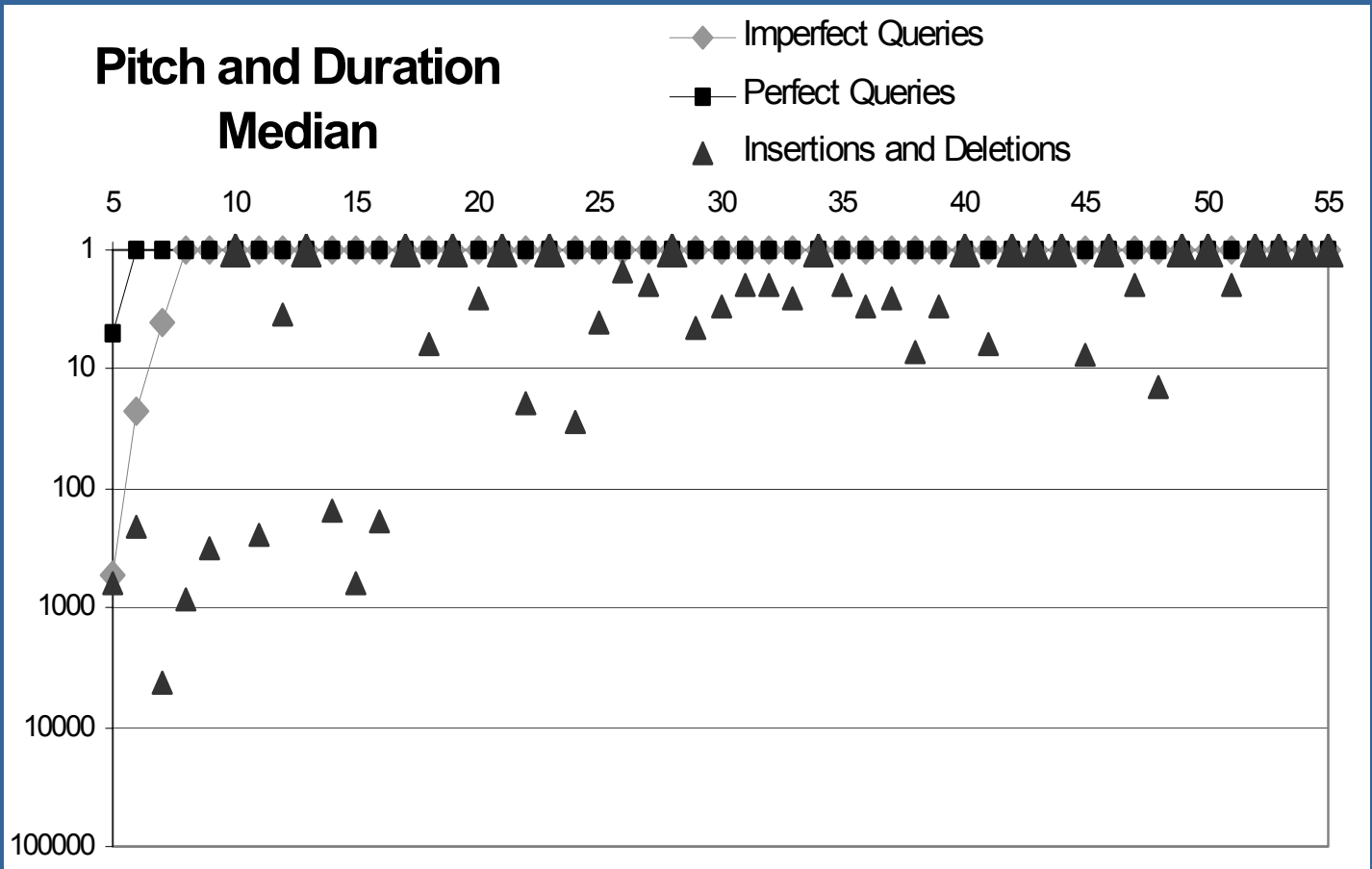
P(“deletion”) = 0.13

Matcher Modifications

- Pitch and Duration Matcher
  - No modifications
- Pitch Only Matcher
  - Only consider deltaPitch in observation table
- Duration Only Matcher
  - Only consider IOIratio in observation table
- Show effectiveness of various representations

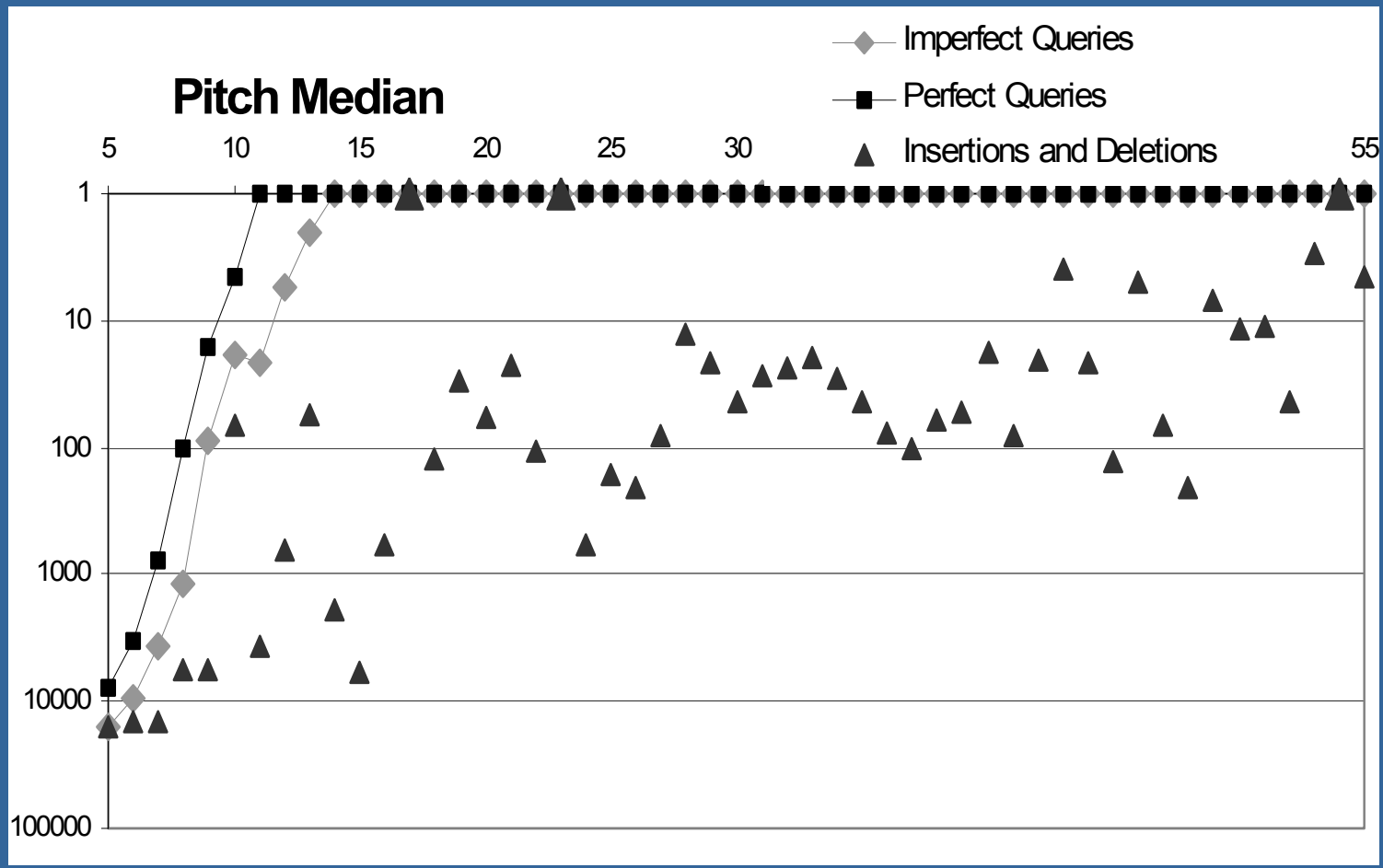
Results:
Pitch and Duration Matcher

Significant Benefit from longer queries
Huge drop-off with insertions and deletions



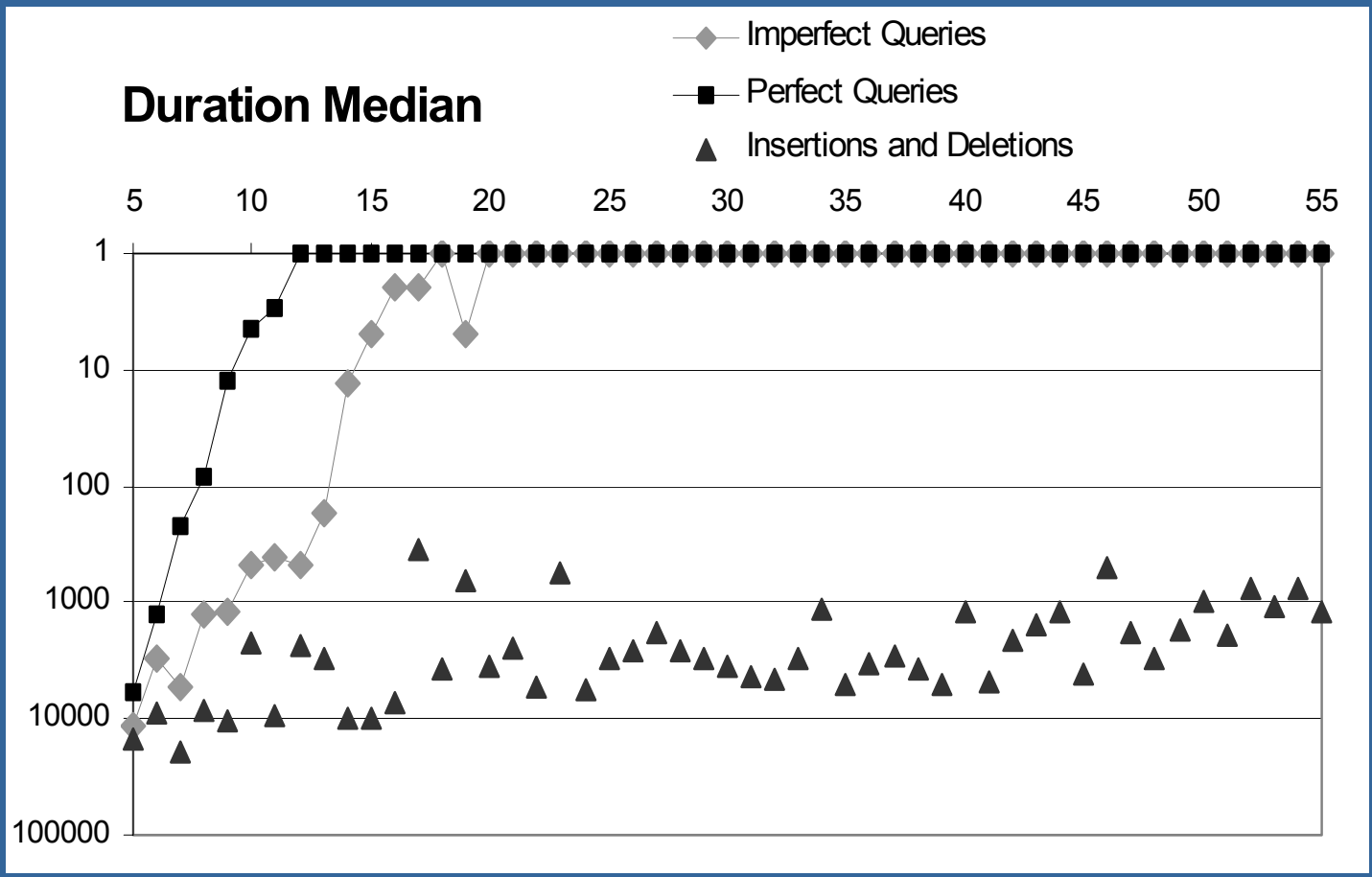
Results:
Pitch Only Matcher

Similar results to Pitch and Duration
Requires more notes to distinguish queries



Results: Duration Matcher

Significant Drop-off
Terrible performance by Insertions and Deletions



Conclusions and Future Work

- Query length is significant
  - Saturation point
  - Query length used as confidence factor
- Topological Improvements
  - Insertion and Deletion states
- Suggests a change to the duration model

Conclusions

- Are individual themes distinguishable at all?
  - Yes. Perfect Queries performed quite well
- Can queries modeled for our system be retrieved?
  - Yes. Results were very encouraging.
- Will “real-world” modeled queries be identified by our system?
  - Not reliably. Needs topological modifications.

Questions