Was Parsons Right?

An experiment in usability of
music representations for melody-based music retrieval

A. L. Uitdenbogerd
and Yaw Wah Yap

Ancient History of Music IR

Offline Music IR:

- Barlow and Morganstern "A Dictionary of Musical Themes" 1948 - represented by a sequence of note names, with the theme transposed into the key of C, for example: EEFGGFED

- Parsons "The directory of tunes and musical themes" 1975

  - designed to be usable for non-musicians
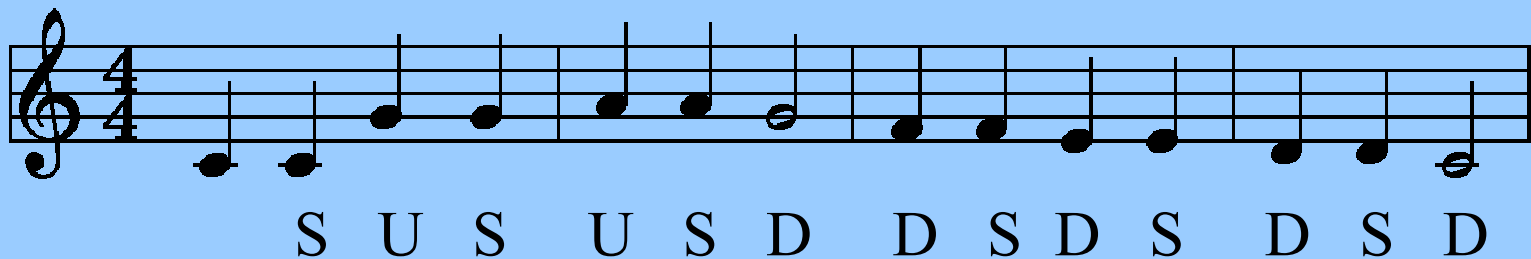 - contour strings (*RUURDDD)

Questions

- Who are we designing content-based music retrieval systems for?

- If it is for non-musicians, what skills can they use for preparing queries?

- Can they use a text-based query system?

Experiment

- Users with a wide range of musical skill – self-assessed into one of four categories
- The user's task: construct a simple text-based melody query to retrieve results from a prototype music search engine for a given piece of music (Twinkle Twinkle Little Star).
- Three text-based representations.
- Definition: successful query = answer retrieved in the ranked list

Contour

- Up represented by U
- Down represented by D
- Same pitch represented by S
- First phrase of Twinkle: SUSUSDDSDSDSD


S U S U S D D S D S D S D

Extended Contour

- An extended version of contour, where larger jumps in pitch are distinguished from smaller ones.

- Step = 1 or 2 semitones – lower case u or d

- Leap > 2 semitones – upper case U or D

- First phrase of Twinkle: SUSuSddSdSdSd

S U S   u S d   d S d S   d S d

Simple Numeric Scale

- Notes of the major scale numbered from 1 to 8.

- Representation of notes outside the range or key ignored.

- Careful choice of query to fit within the limitation.

- First phrases of Twinkle: 1 1 5 5 6 6 5 4 4 3 3 2 2 1

1   1   5   5   6   6   5   4   4   3   3   2   2   1
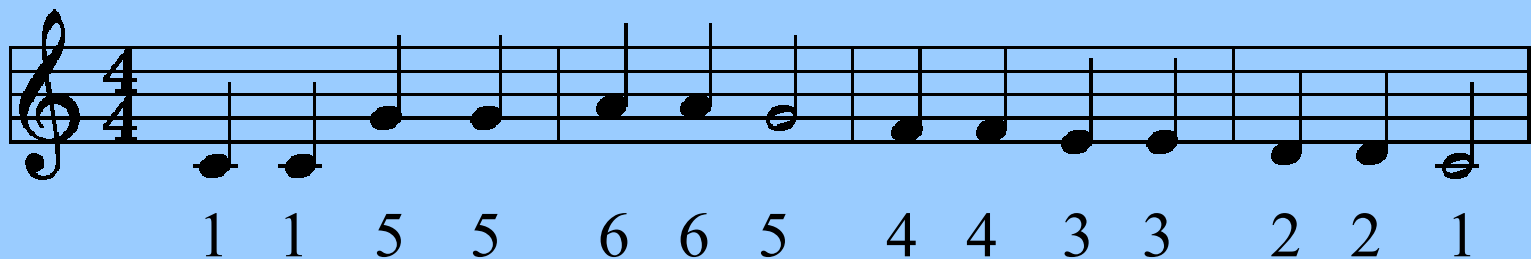
The Collection and the System

- The collection:10,466 MIDI files + 4 simple well-known melodies.

- directed modulo-12 (or contour) melody representation for matching (intervals reduced to octave maximum, plus direction of interval)

- 5-gram coordinate matching

Experiment Participants

Several cohorts:

– Randomly selected from business faculty building

– CS PhD students

– Choristers

– Hand-picked CS staff with musical skill

Music Background	Number of Users	Percentage (%)
Zero	13	36
Novice	10	28
Intermediate	8	22
Professional	5	14
Total	36	100

User Friendliness

no. of users (percentage of responses within user category)

Music Background	Excellent	Good	Moderate	Hard to Use
Zero	1 (07.69%)	7 (53.85%)	4 (30.77%)	1 (07.69%)
Novice	0 (00.00%)	7 (70.00%)	3 (30.00%)	0 (00.00%)
Intermediate	0 (00.00%)	6 (75.00%)	1 (12.50%)	1 (12.50%)
Professional	2 (40.00%)	3 (60.00%)	0 (00.00%)	0 (00.00%)
Total	3 (08.33%)	23 (63.89%)	8 (22.22%)	2 (05.56%)

Learnability

Music Background	Easy	Moderate	Hard	Very Hard
Zero	4 (30.77%)	4 (30.77%)	4 (30.77%)	1 (7.69%)
Novice	2 (20.00%)	5 (50.00%)	3 (30.00%)	0 (0.00%)
Intermediate	1 (12.50%)	6 (75.00%)	1 (12.50%)	0 (0.00%)
Professional	2 (40.00%)	3 (60.00%)	0 (00.00%)	0 (0.00%)
Total	9 (25.00%)	18 (50.00%)	8 (22.22%)	1 (2.78%)

Screen Navigation

Music Background	Excellent	Good	Confusing	Complicated
Zero	2 (15.38%)	11 (84.62%)	0 (00.00%)	0 (0.00%)
Novice	3 (30.00%)	7 (70.00%)	0 (00.00%)	0 (0.00%)
Intermediate	0 (00.00%)	7 (87.50%)	1 (12.50%)	0 (0.00%)
Professional	5 (100.00%)	0 (00.00%)	0 (00.00%)	0 (0.00%)
Total	10 (27.78%)	25 (69.44%)	1 (02.78%)	0 (0.00%)

Overall Satisfaction

Music Background	Very Satisfied	Satisfied	Moderate	Poor	Very Poor
Zero	1 (07.69%)	4 (30.77%)	6 (46.15%)	2 (15.38%)	0 (0.00%)
Novice	1 (10.00%)	5 (50.00%)	4 (40.00%)	0 (00.00%)	0 (0.00%)
Intermediate	1 (12.50%)	5 (62.50%)	1 (12.50%)	1 (12.50%)	0 (0.00%)
Professional	2 (40.00%)	3 (60.00%)	0 (00.00%)	0 (00.00%)	0 (0.00%)
Total	5 (13.89%)	17 (47.22%)	11 (30.56%)	3 (08.33%)	0 (0.00%)

Preferred Representation

- contour followed by scale: 92% of 13 users preferred scale.

- scale followed by contour: 75% of 12 users preferred scale.

- Only novice users preferred contour, but the results are too close to be significant.

Proportion of successes by type of users.

no. of successful queries/total number of trials (proportion as a percentage)

Music Background	Contour		Extended Contour		Numeric Scale	
Zero	0/19	(0.00%)	0/3	(0.00%)	0/14	(0.00%)
Novice	1/16	(6.25%)	1/6	(16.67%)	3/12	(25.00%)
Intermediate	1/11	(9.09%)	1/3	(33.33%)	5/12	(41.67%)
Professional	4/07	(57.14%)	4/6	(66.67%)	4/06	(66.67%)

Typical Errors

- Additional symbol at the start of a contour or extended contour query (eg. SSUSUSD instead of SUSUSD). This was still counted as a successful query as it retrieved the answer.

- Reduced numeric range: 1 1 2 2 3 3 2 was common.

- Incorrect representation (contour for extended contour)

Experimental Results

- Musically unskilled and beginner users had great difficulty in generating effective queries.

- 60% of musically unskilled users listened to music more than 5 times per week.

- Scale representation was preferred over contour.

- Scale queries were more successful.

- Average query length for the task was about 7.

Conclusions

- Contour-based text representations are not useful for non-musicians.

- Numeric representations are moderately useful for those with some musical skill.

- More direct forms of query representation are required for non-experts: recording, singing, possibly simple musical instrument-like interfaces.

- The main skill we know that a non-musician user has is the ability to recognise a known piece of music, and possibly the ability to sing a melody contour.

Future Work

- Testing the usability of keyboard and piano-roll interfaces

- Comparing skills in using graphical representations

- Singing-based queries

- Rhythm queries