CrossSong puzzle:
Generating and unscrambling music mashups with real-time interactivity

Jordan B. L. Smith, Graham Percival, Jun Kato, Masataka Goto, and Satoru Fukayama
Media Interaction Group
National Institute of Advanced Industrial Science and Technology (AIST), Japan

30 July 2015

Oral Session II:
Content Processing of Music Audio Signals

1. Music, games and puzzles

- Music, games and puzzles are all incredibly popular pastimes
- Many activities combine two of these pursuits:
    - *Guitar Hero* and *Rock Band* and others combine music and games
    - Countless games include puzzle elements
    - Figured bass exercises are music puzzles
- But nothing combines all three: music, games and puzzles!

Music + Games + Puzzles

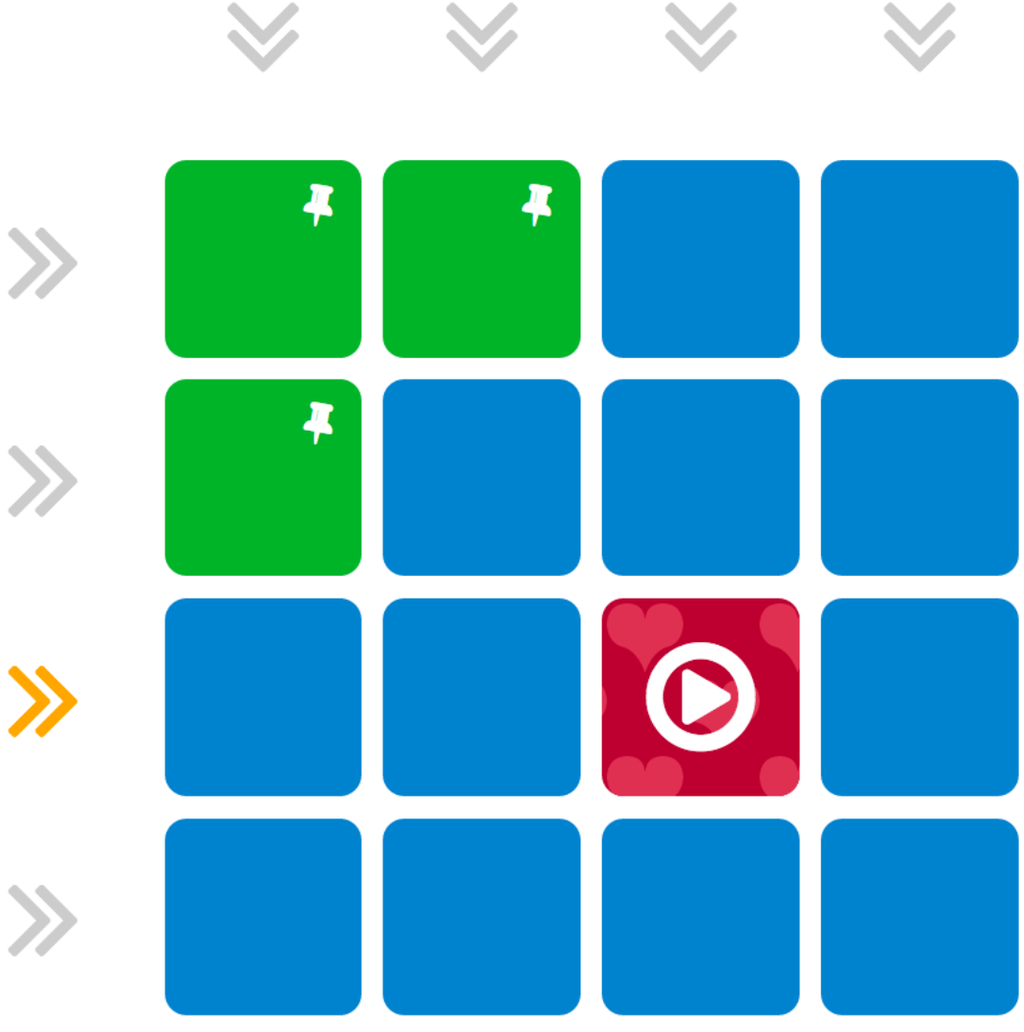
Goal: make something that combines these!

1. Music, games and puzzles: motivation

- …but **why**? (other than because it hasn't been done)
    - Puzzle pleasure:
      pattern identification and fulfilment
    - Music pleasure:
      pattern identification and fulfilment (or denial)
- Interesting to play with both at once

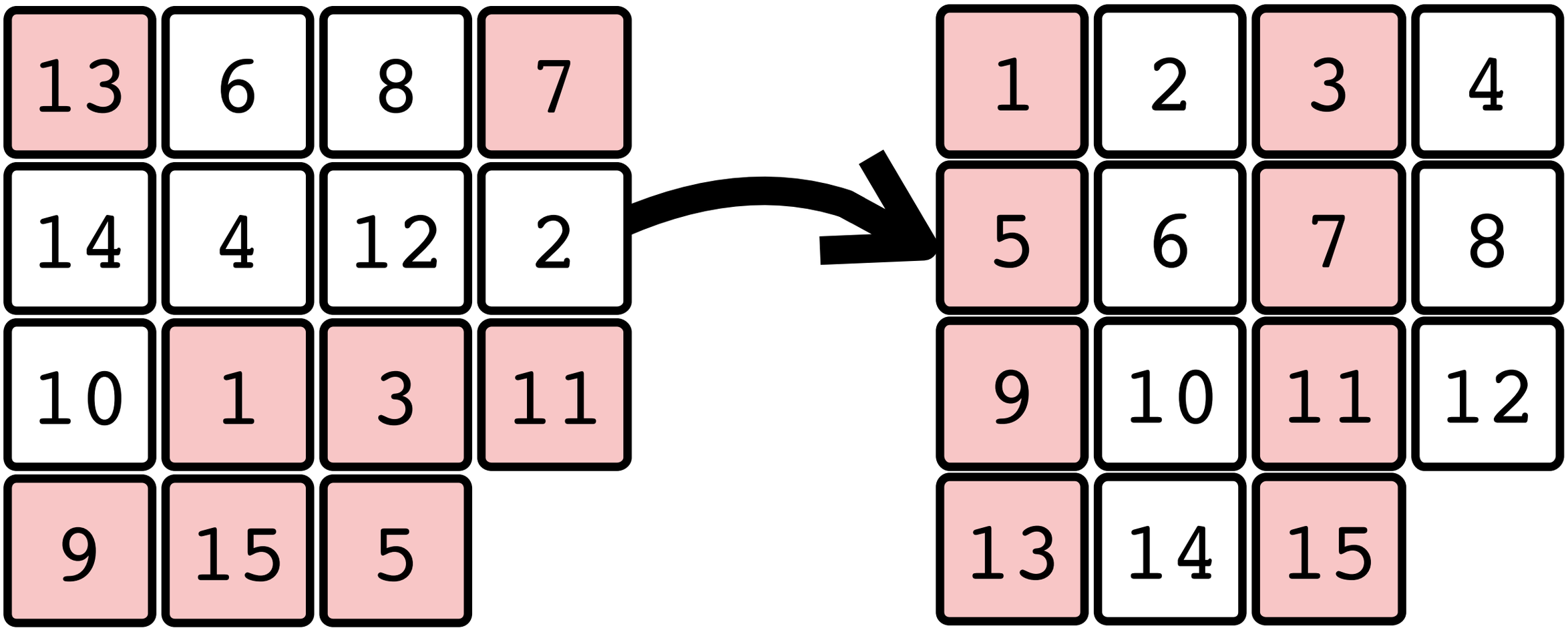
2. CrossSong puzzle

- Design goals:
  - Puzzle should **require musical thinking**
    - Most “music” puzzles are just puzzles that have music
  - Puzzle should **be fun to listen to**
    - An engaging musical experience
    - Music should be continuous / have a beat



2. CrossSong puzzle

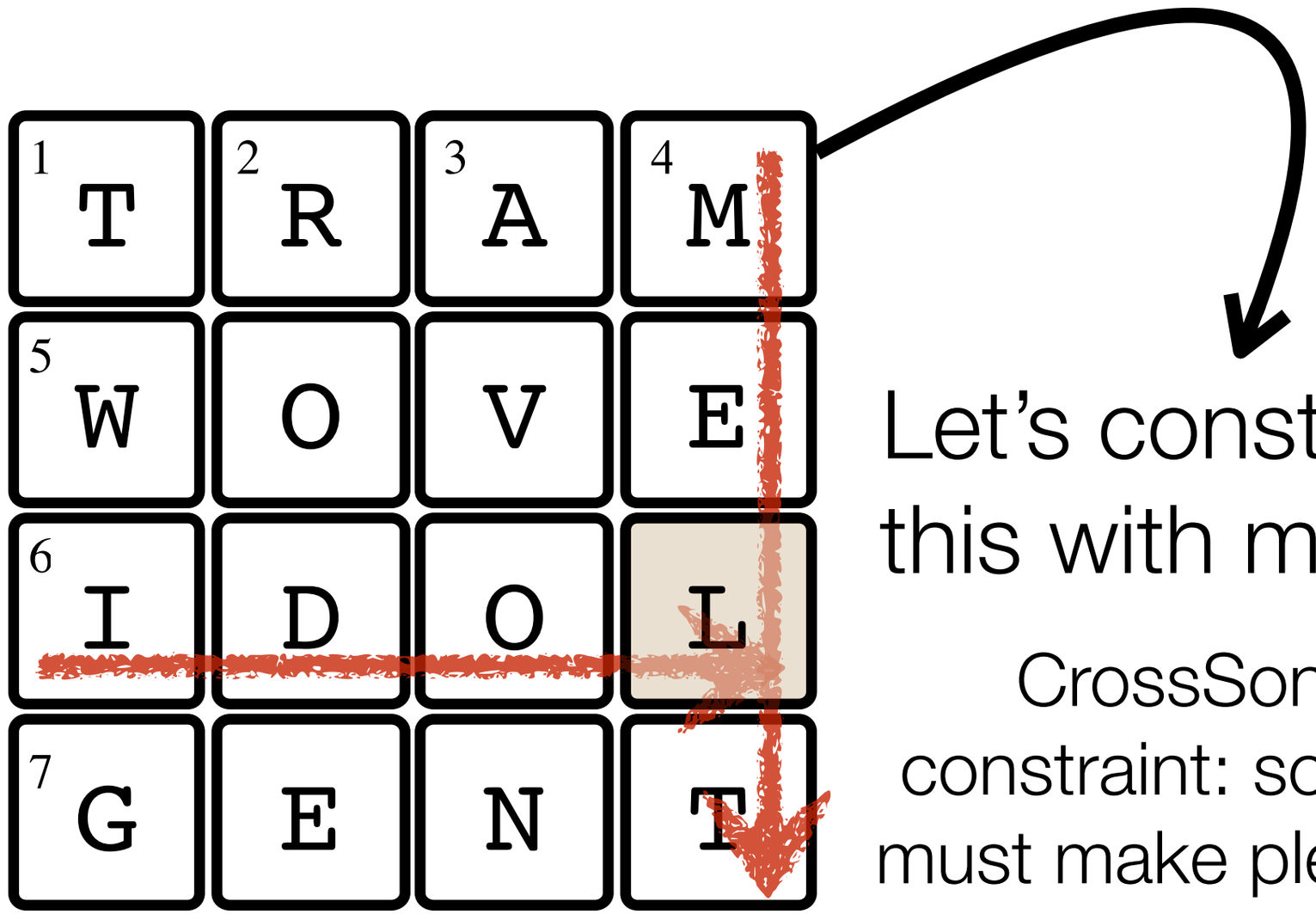
Goal:
Unscramble the tiles



2. CrossSong puzzle

Construction:
Parts of puzzle fit in interlocking rows and columns

Constraint at intersection for crosswords: letter in each word must be the same



Let's construct this with music

CrossSong constraint: sounds must make pleasing mashup!

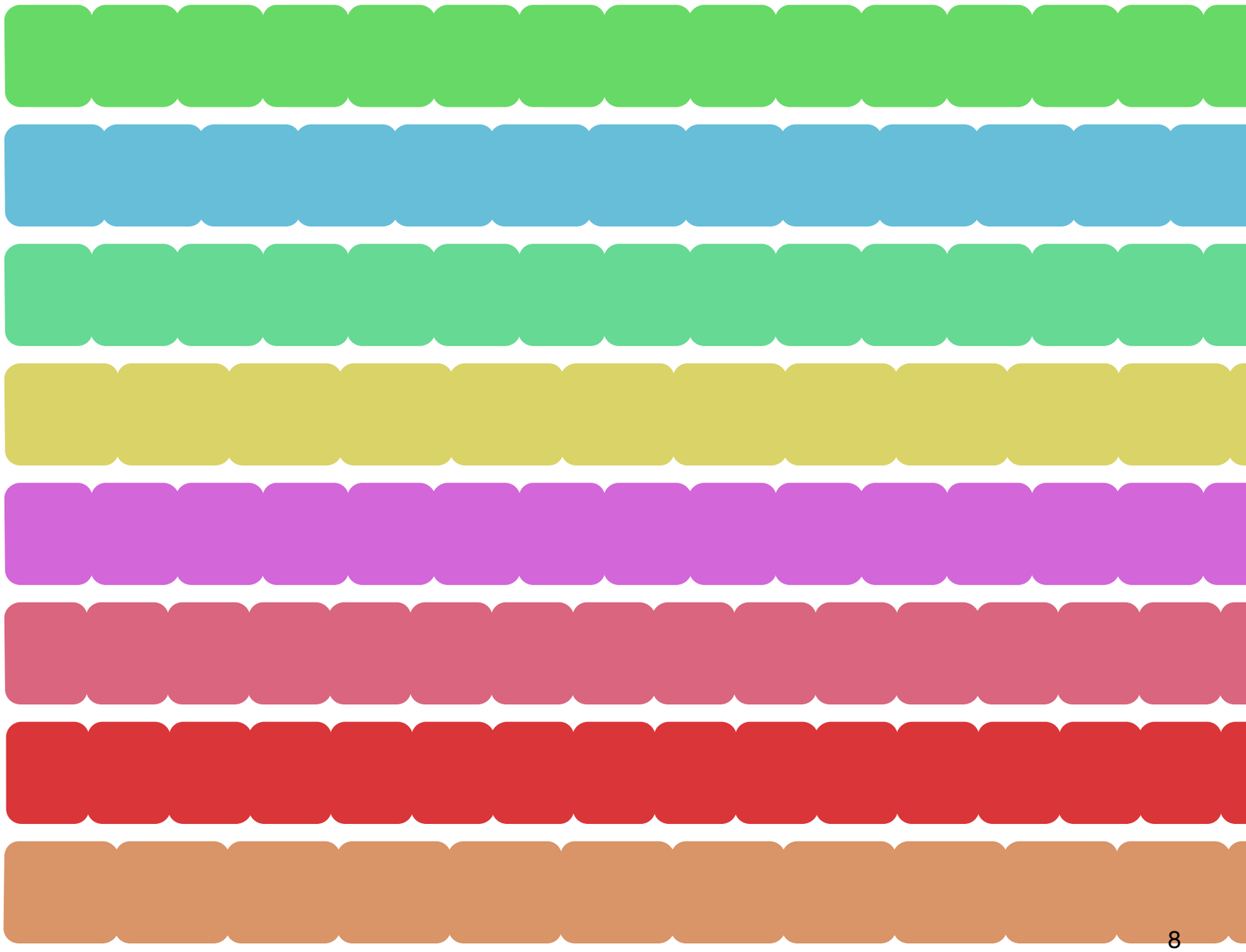
2. CrossSong puzzle

User:

Input: 8 songs

System:

Step 1: extract 4-measure clips

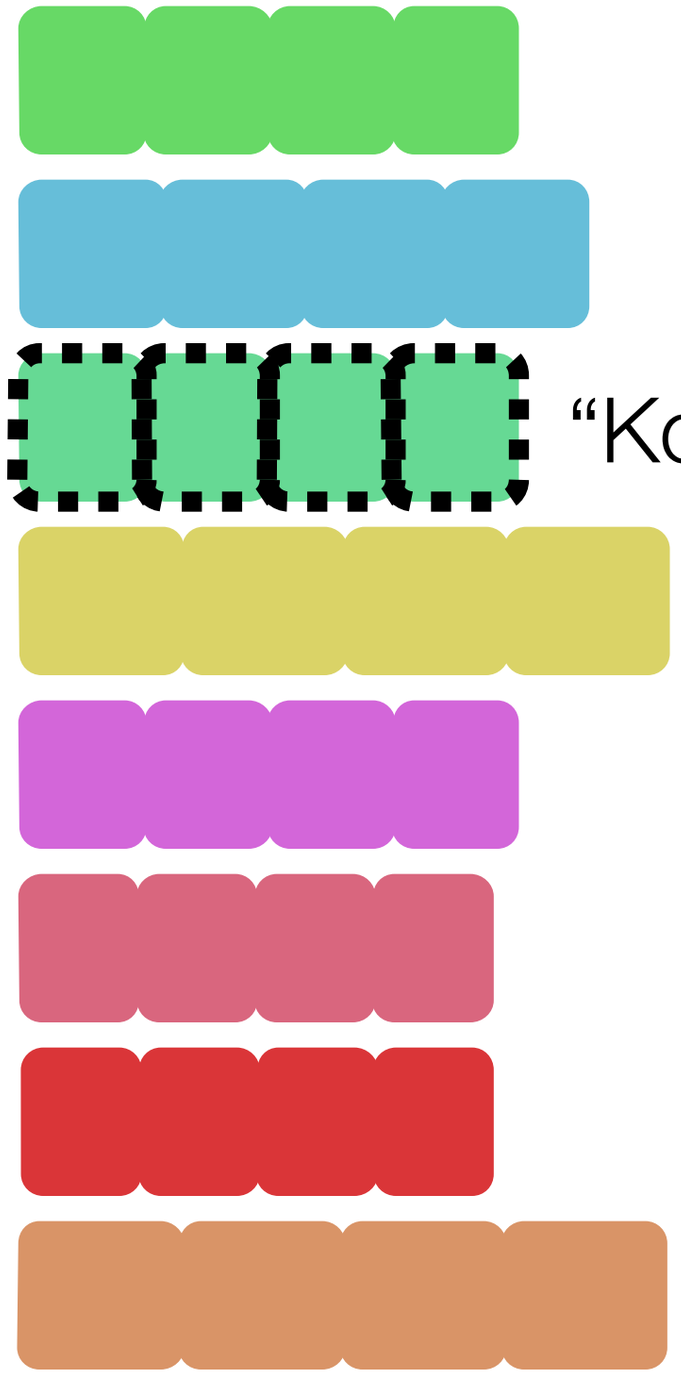


(All music examples from RWC Popular Database. See M. Goto, H. Hashiguchi, T. Nishimura and R. Oka. 2002. RWC Music Database: Popular, classical, and jazz music databases. In Proceedings of ISMIR. 287–8.)

2. CrossSong puzzle

User:

Input: 8 songs



System:

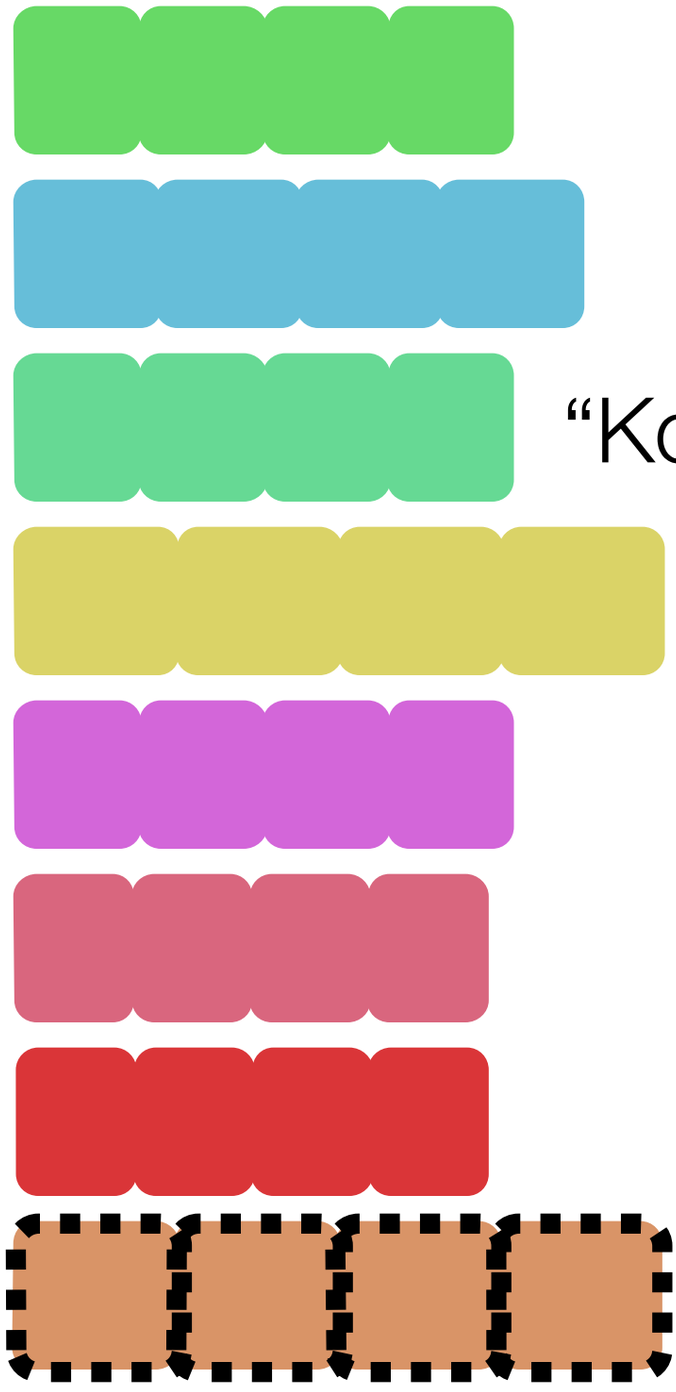
Step 1: extract
4-measure clips

“Koino Ver.2.4” by Eves RWC Pop #5

2. CrossSong puzzle

User:

Input: 8 songs



System:

Step 1: extract
4-measure clips

“Koino Ver.2.4” by Eves

“Life Is What You Make It To Be” by Donna Burke RWC Pop #89

2. CrossSong puzzle

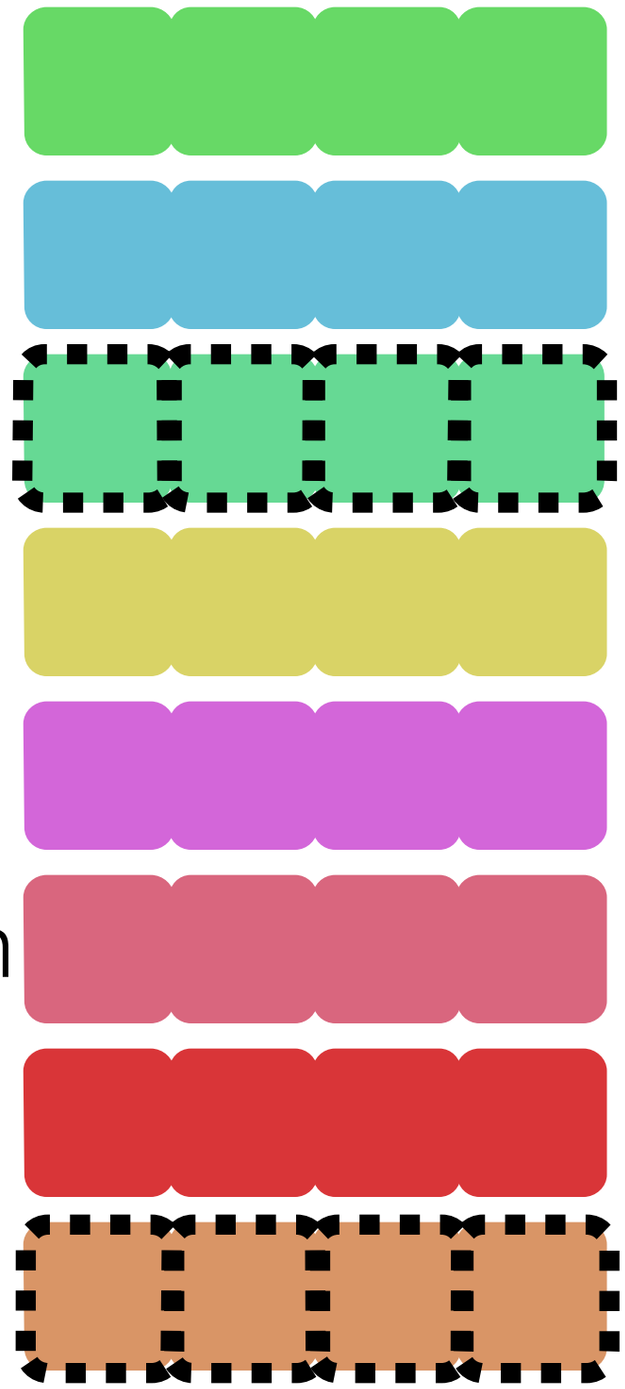
User:

Input: 8 songs

System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch



“Koino Ver.2.4” by Eves

mashed-up with

“Life Is What You Make It To Be” by Donna Burke

2. CrossSong puzzle

User:

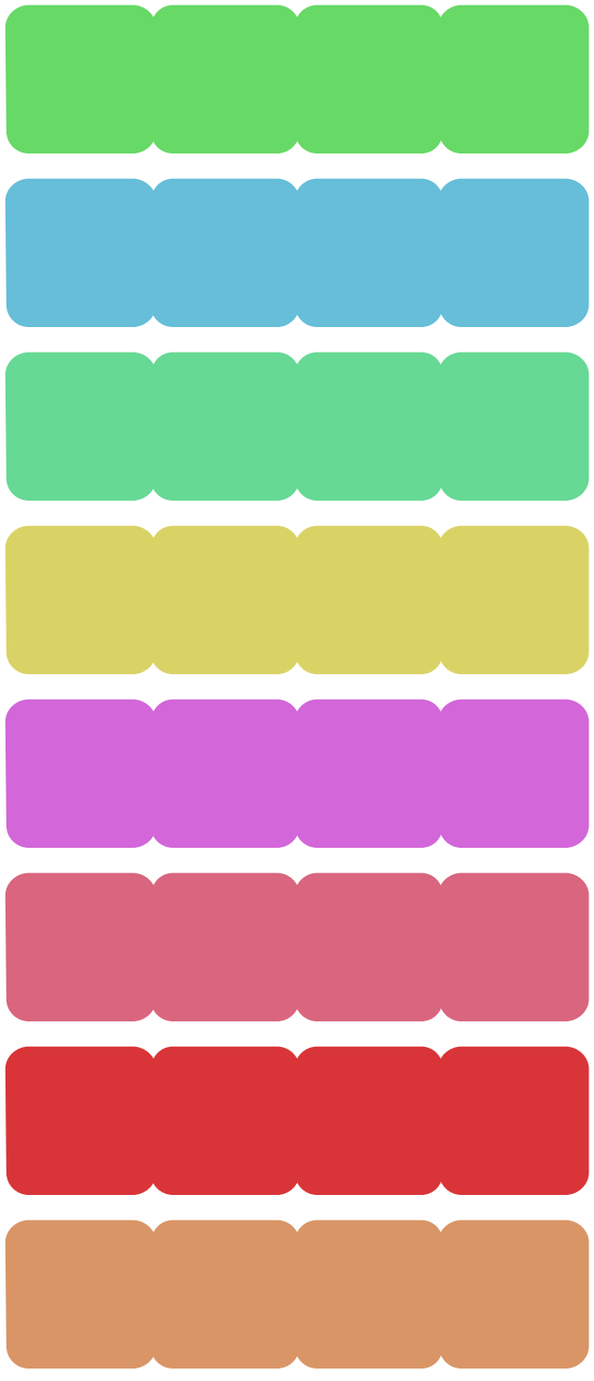
Input: 8 songs

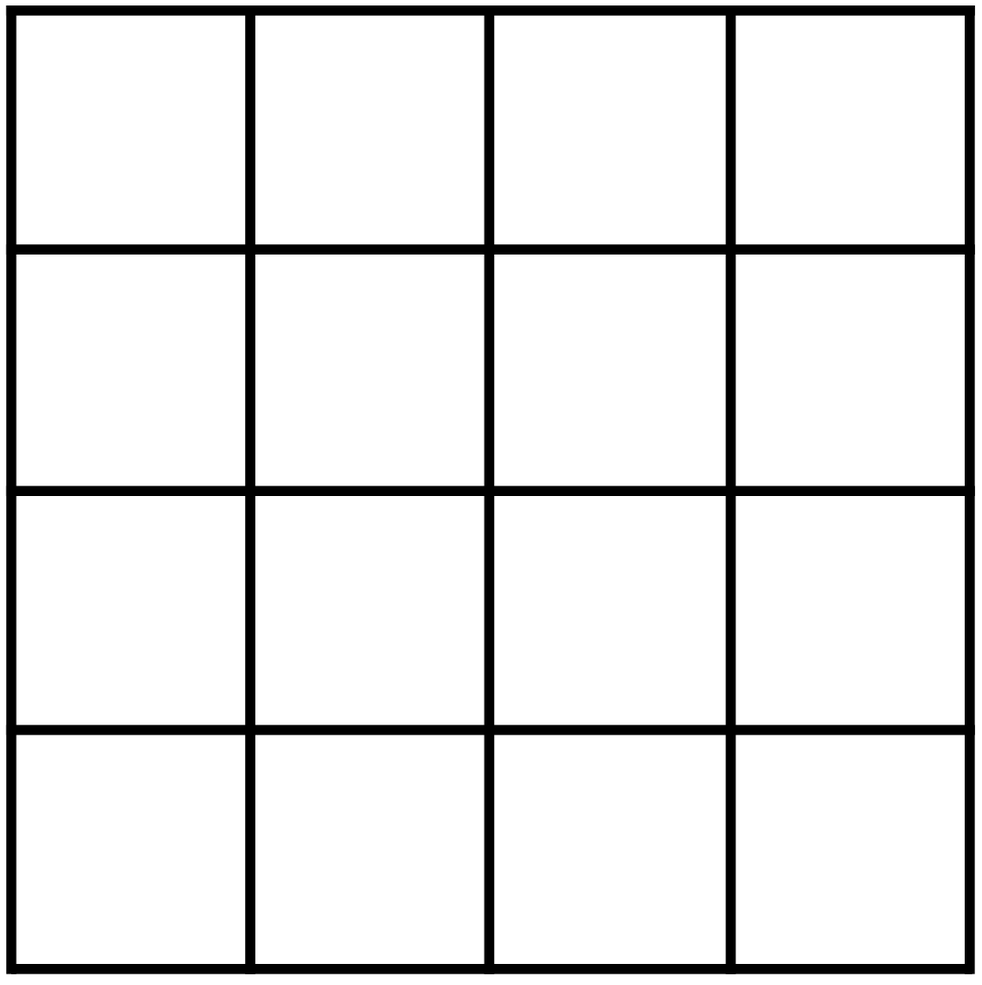
System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch

Step 3: arrange songs in grid





2. CrossSong puzzle

User:

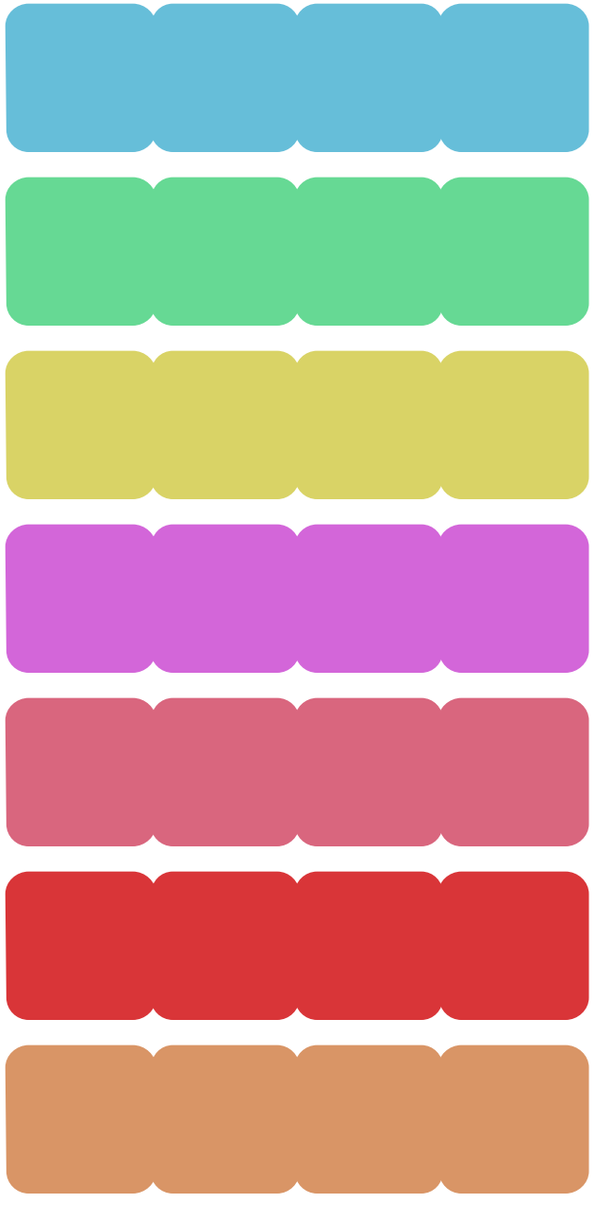
Input: 8 songs

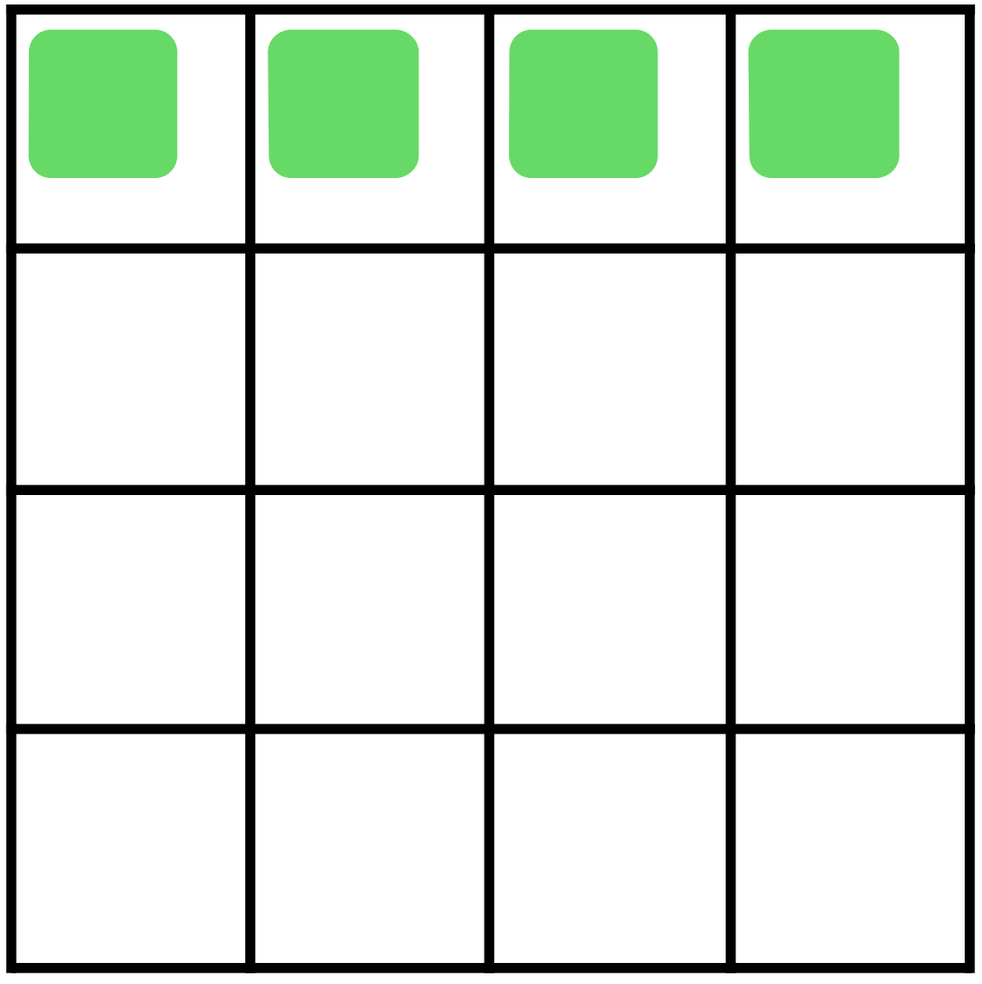
System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch

Step 3: arrange songs in grid





2. CrossSong puzzle

User:

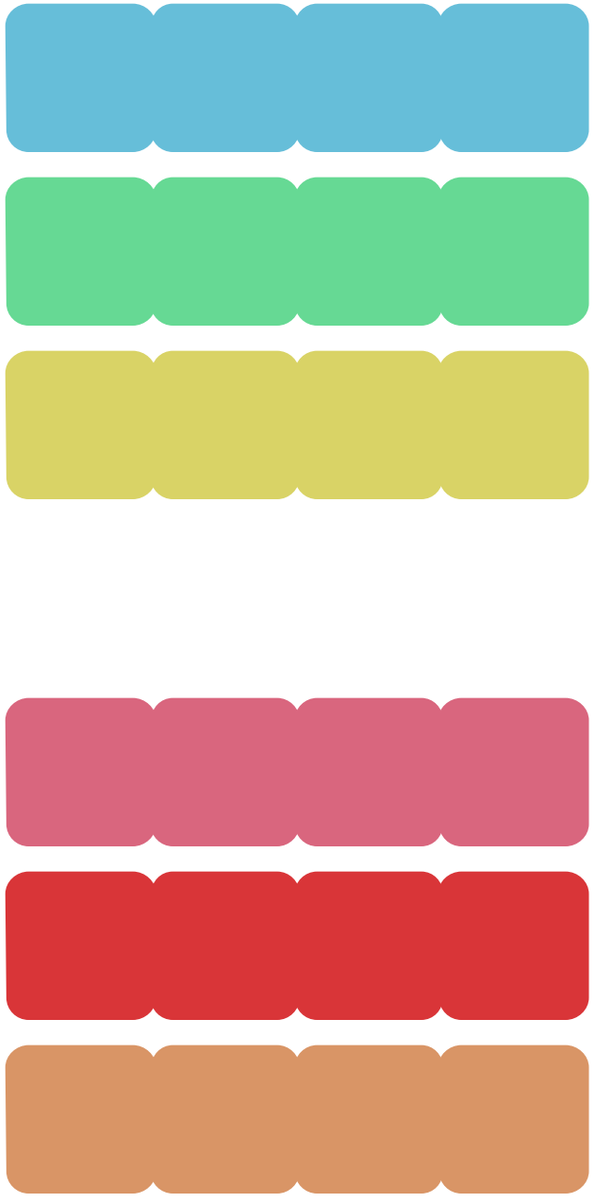
Input: 8 songs

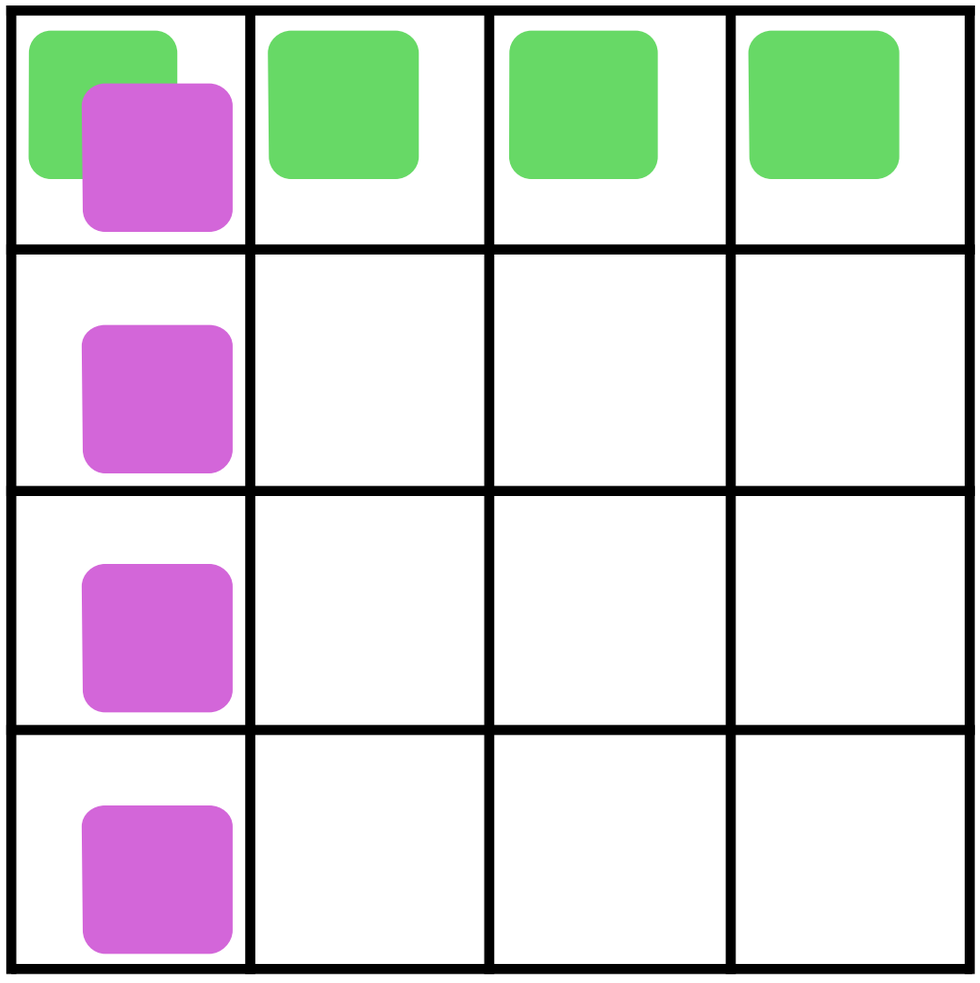
System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch

Step 3: arrange songs in grid





2. CrossSong puzzle

User:

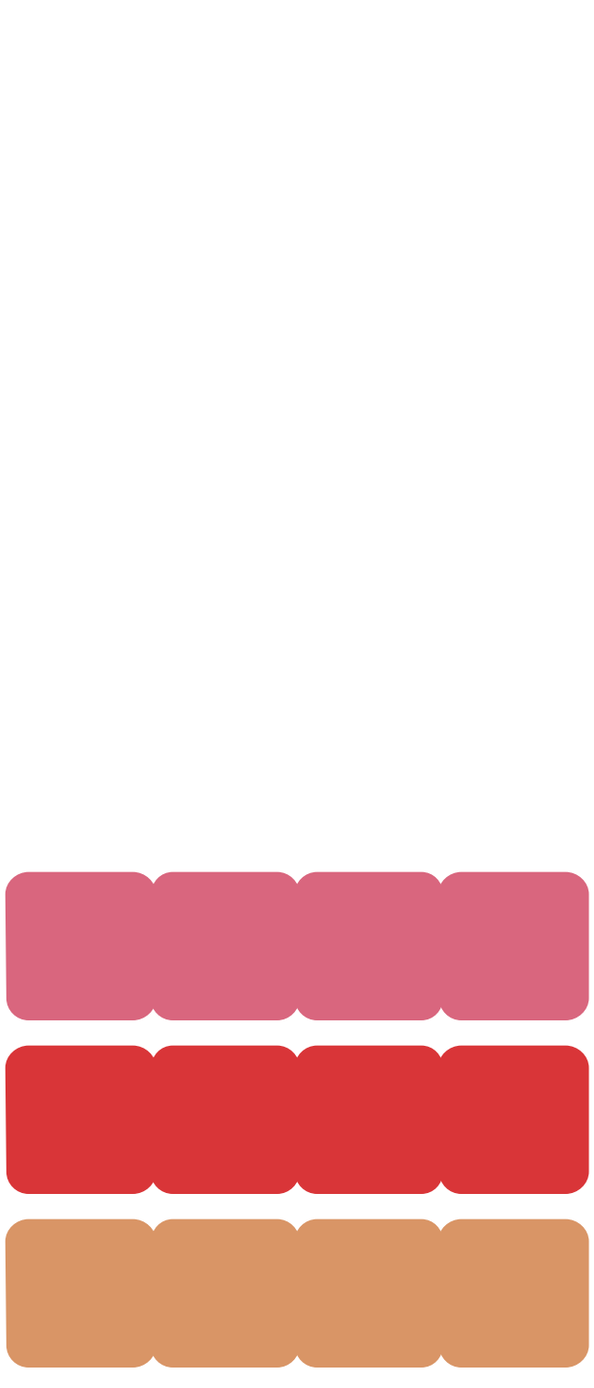
Input: 8 songs

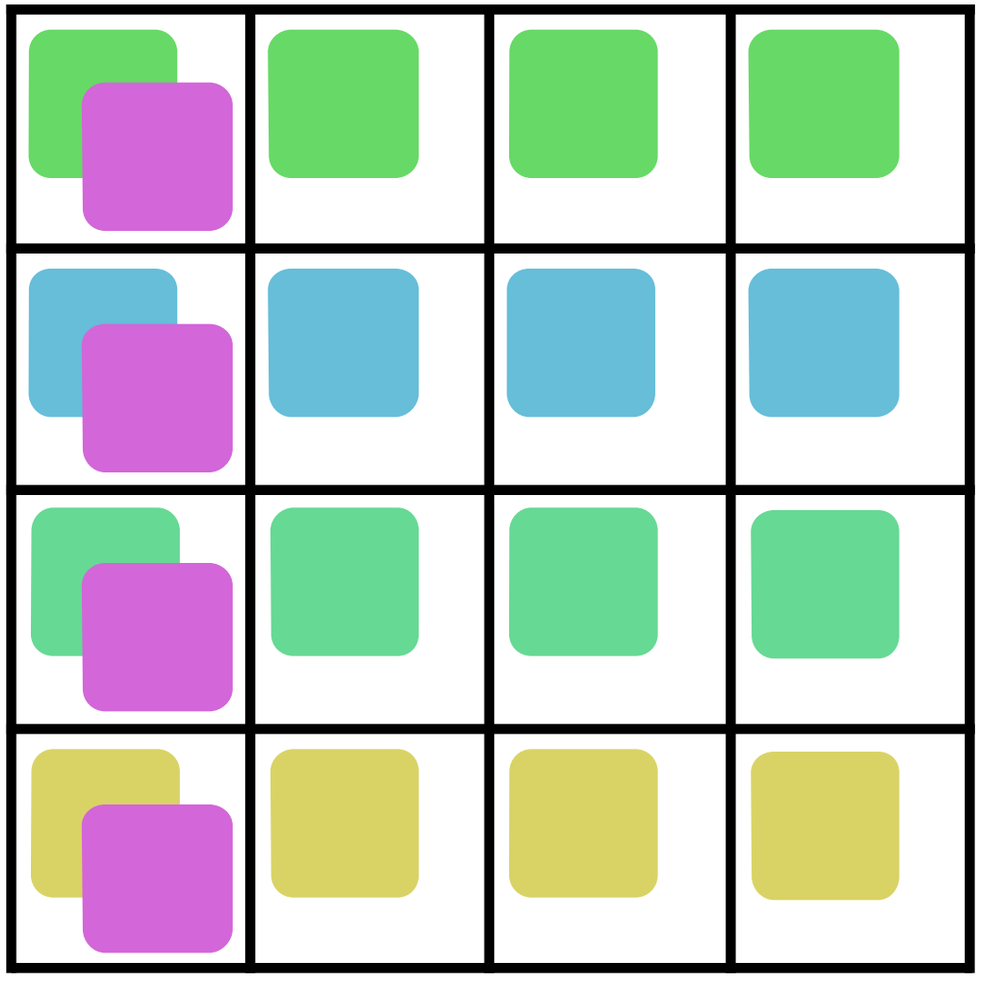
System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch

Step 3: arrange songs in grid





2. CrossSong puzzle

User:

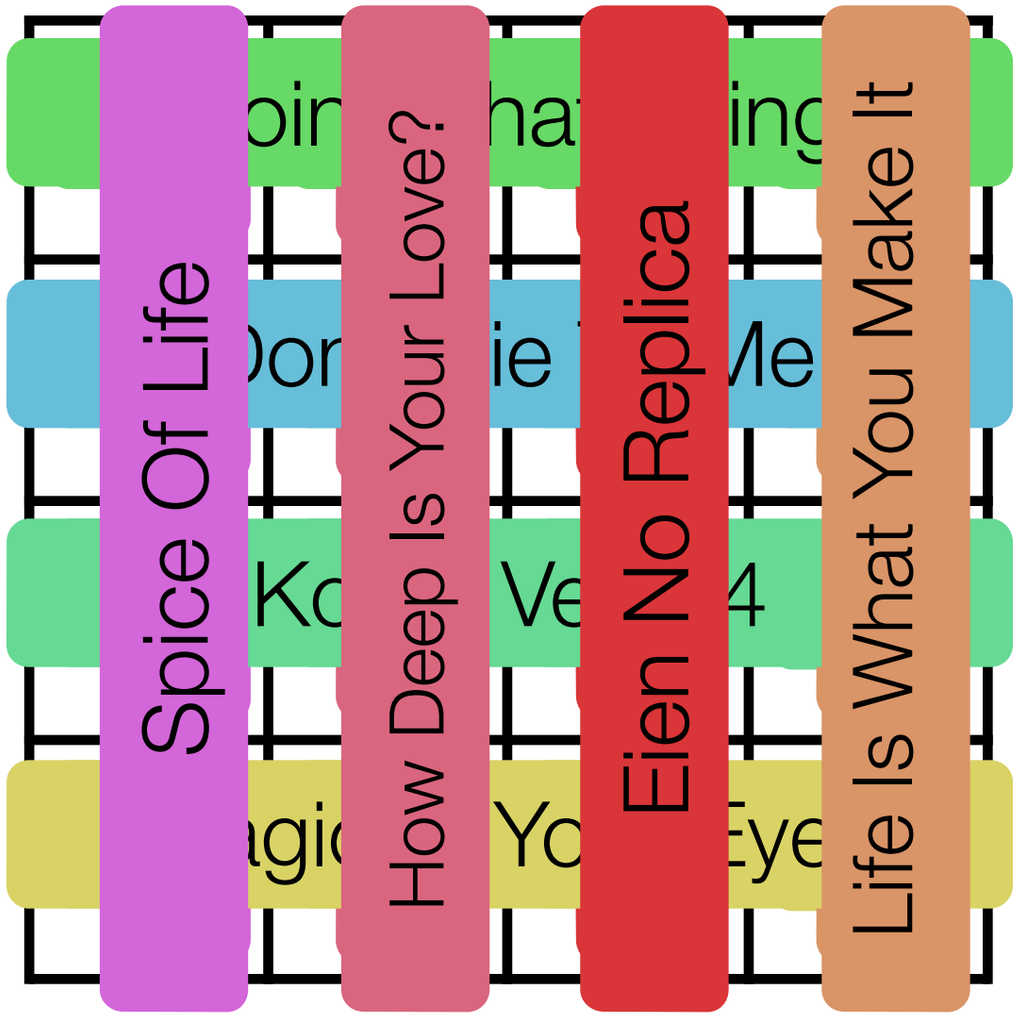
Input: 8 songs

System:

Step 1: extract 4-measure clips

Step 2: match tempo and pitch

Step 3: arrange songs in grid



2. CrossSong puzzle

Doing That Thing

Don’t Lie To Me

Koino Ver.2.4

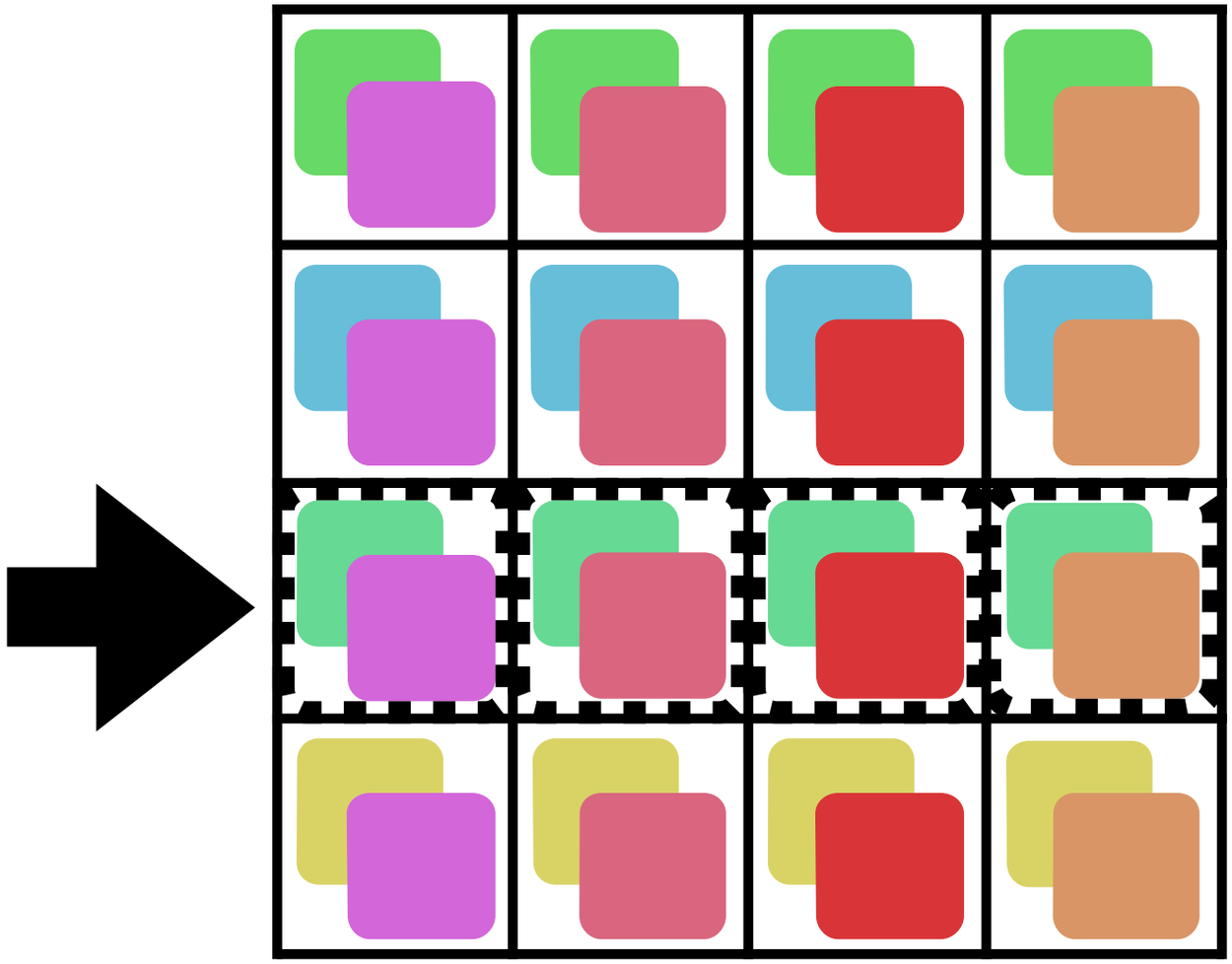
Magic In Your Eyes

Spice Of Life

How Deep Is Your Love?

Eien No Replica

Life Is What You Make It To Be

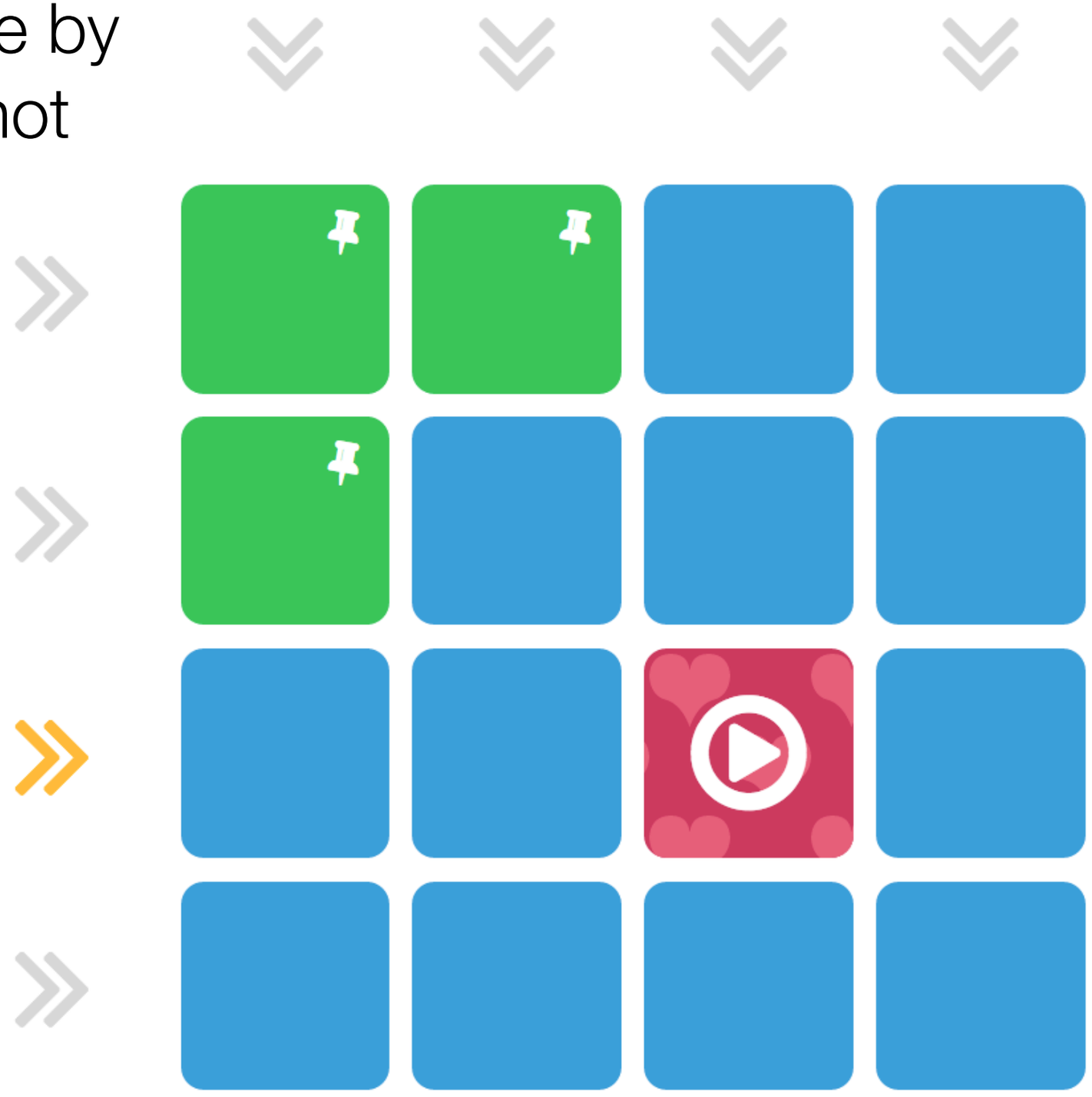


2. CrossSong puzzle

Solve puzzle by **listening**, not **looking**

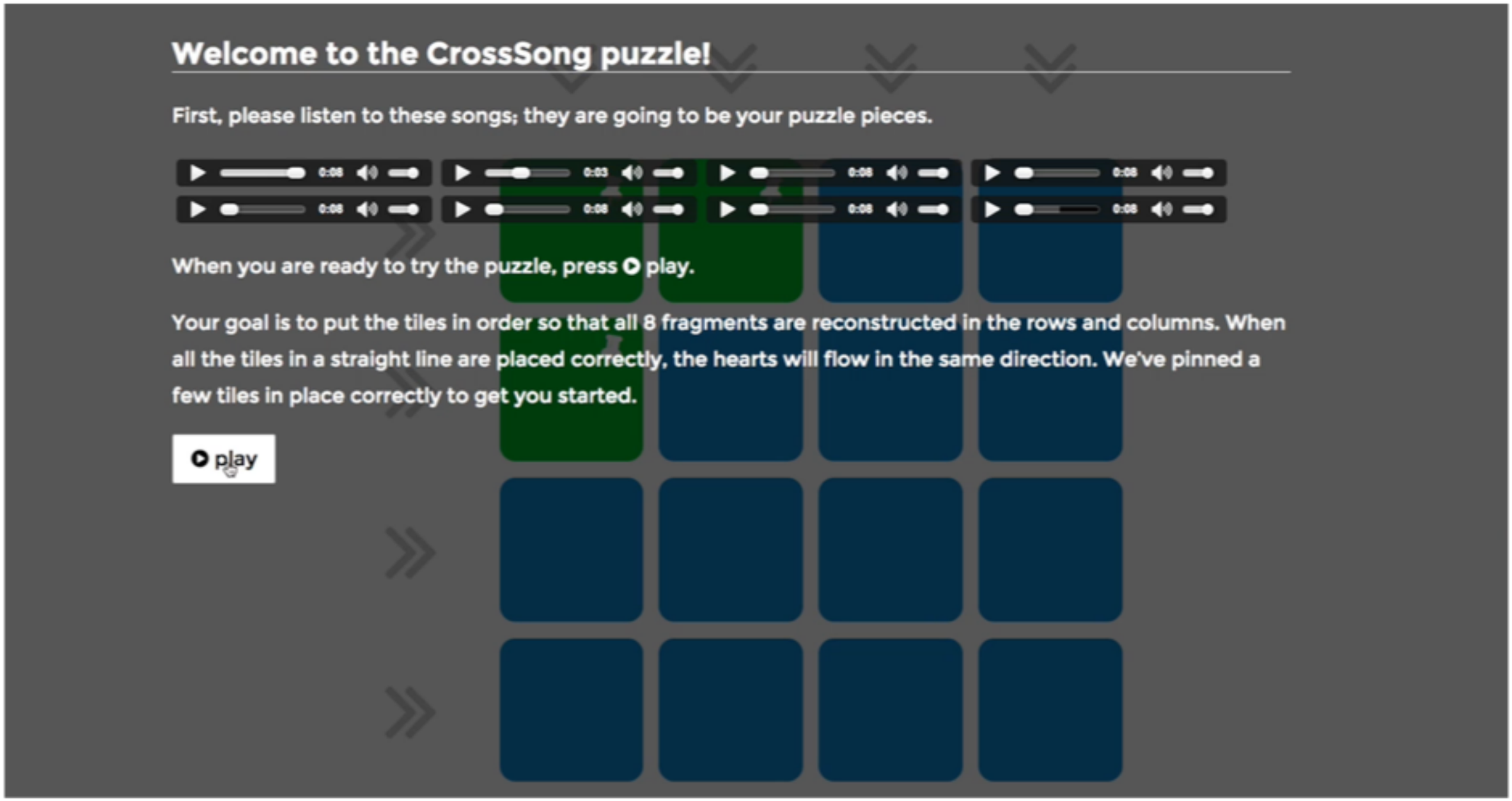
2. CrossSong puzzle

Solve puzzle by **listening**, not **looking**



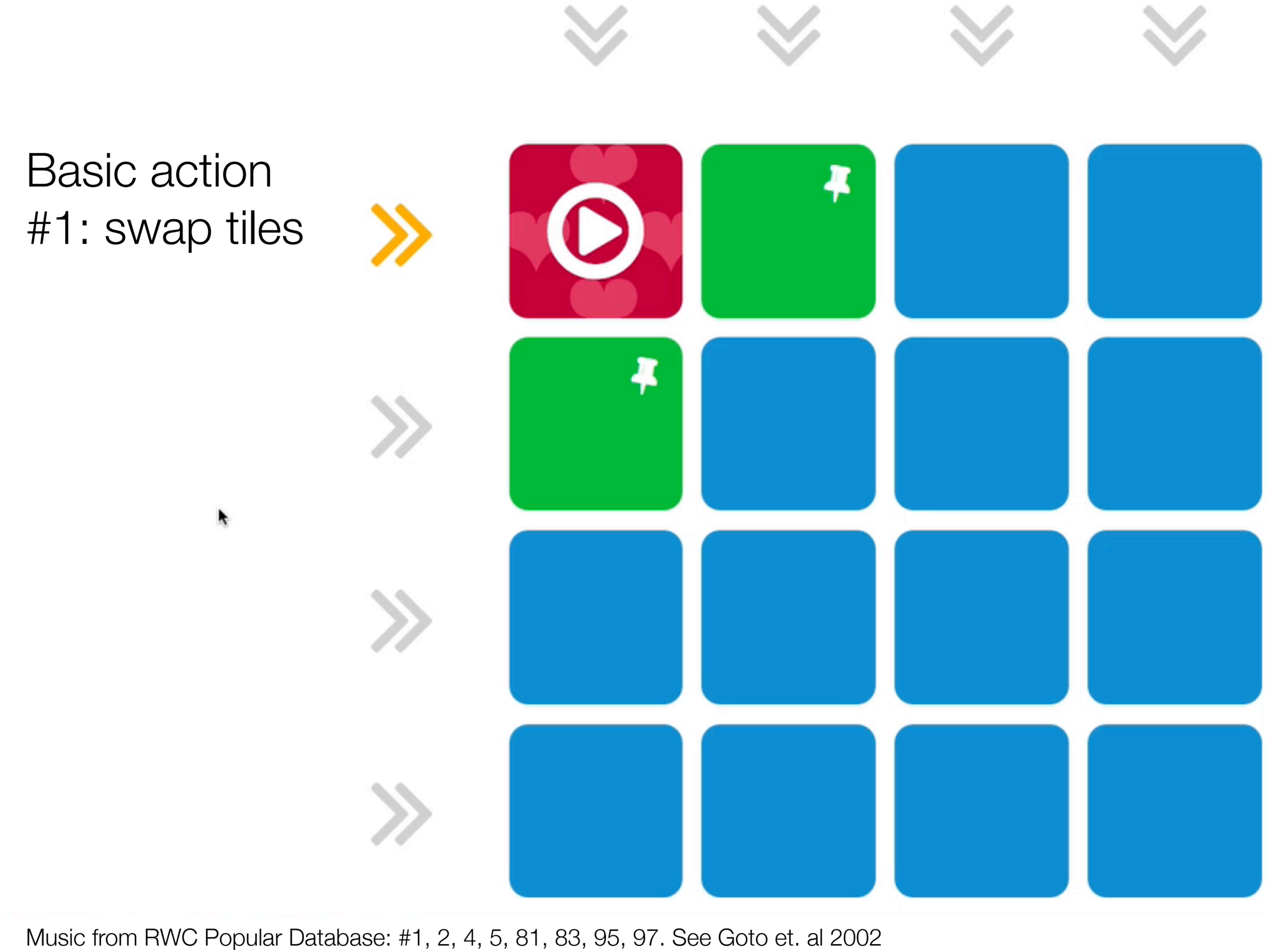
2. CrossSong puzzle: Demo

Solver can listen to isolated tracks before starting



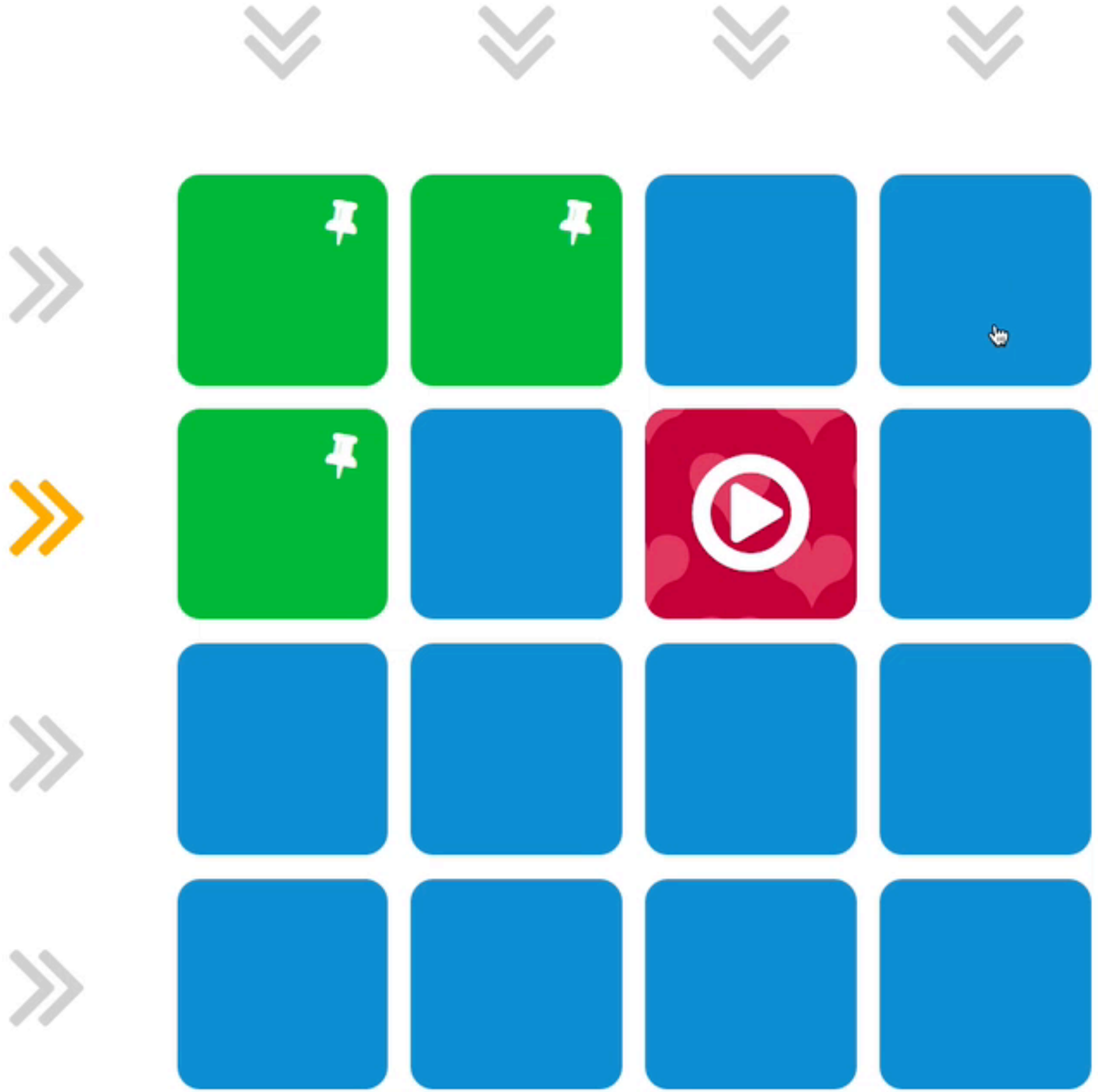
Music from RWC Popular Database: #1, 2, 4, 5, 81, 83, 95, 97. See Goto et. al 2002

Basic action
#1: swap tiles



Music from RWC Popular Database: #1, 2, 4, 5, 81, 83, 95, 97. See Goto et. al 2002

Basic action #2: choose row or column to play next



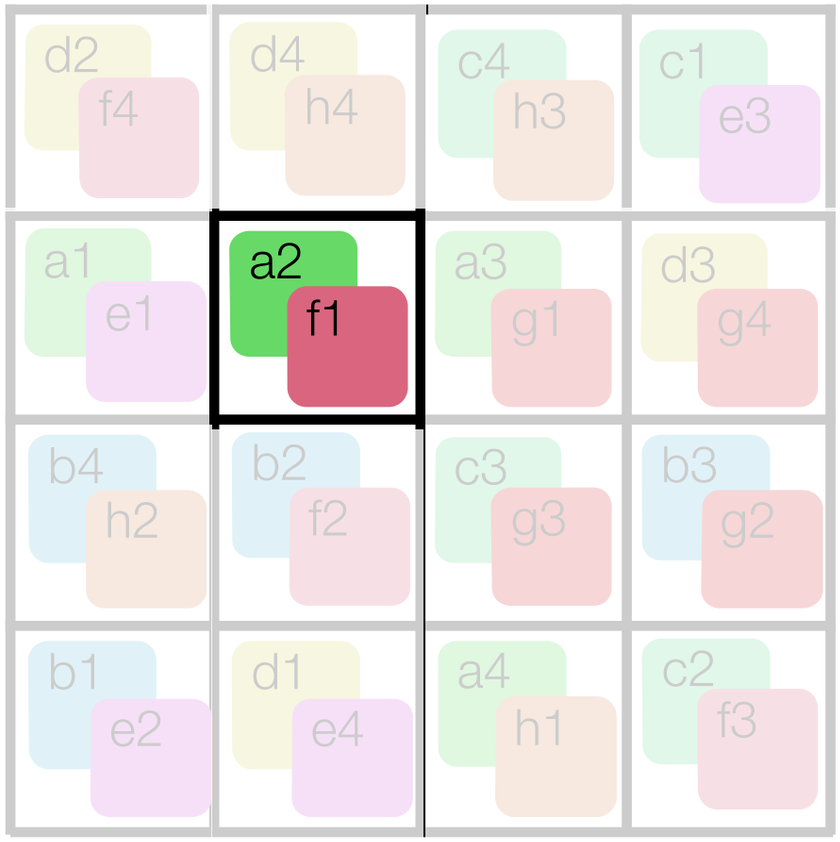
Music from RWC Popular Database: #1, 2, 4, 5, 81, 83, 95, 97. See Goto et. al 2002

2. CrossSong puzzle: confirming progress

- Visual feedback:
    - Row locks in place **after** full hearing
        - Visual feedback confirms progress, but solver **must listen** to music to **make progress**
- When entire puzzle correct, instant win
    - Not likely to happen by chance (6 billion possible tile arrangements)

2. CrossSong puzzle: confirming progress

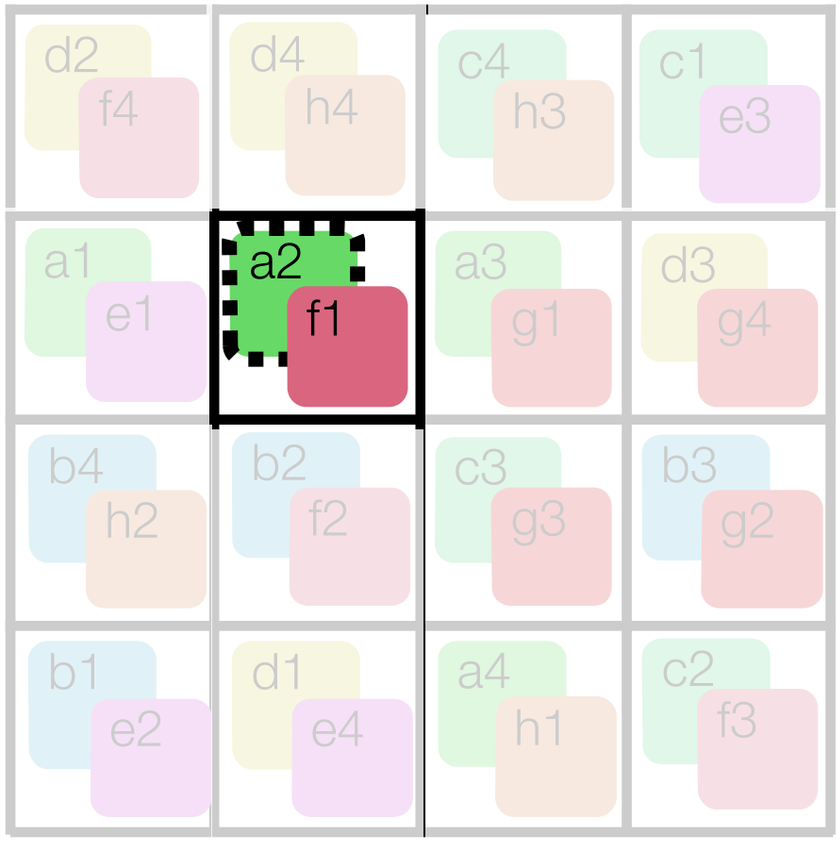
- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours



2. CrossSong puzzle: confirming progress

- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours

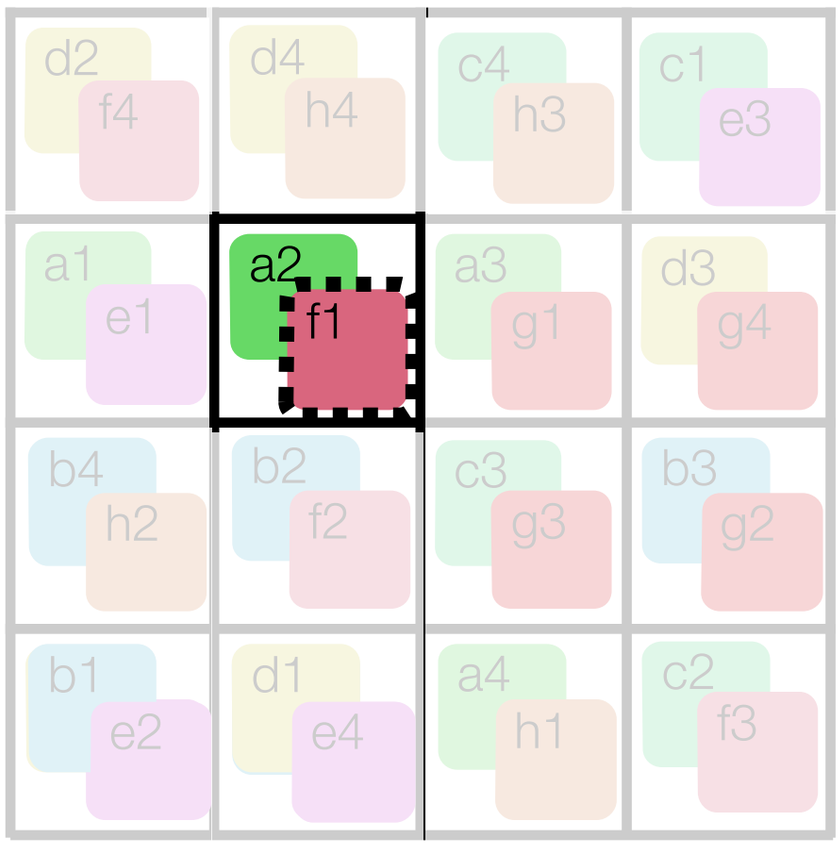
Across song part: *a2*



2. CrossSong puzzle: confirming progress

- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours

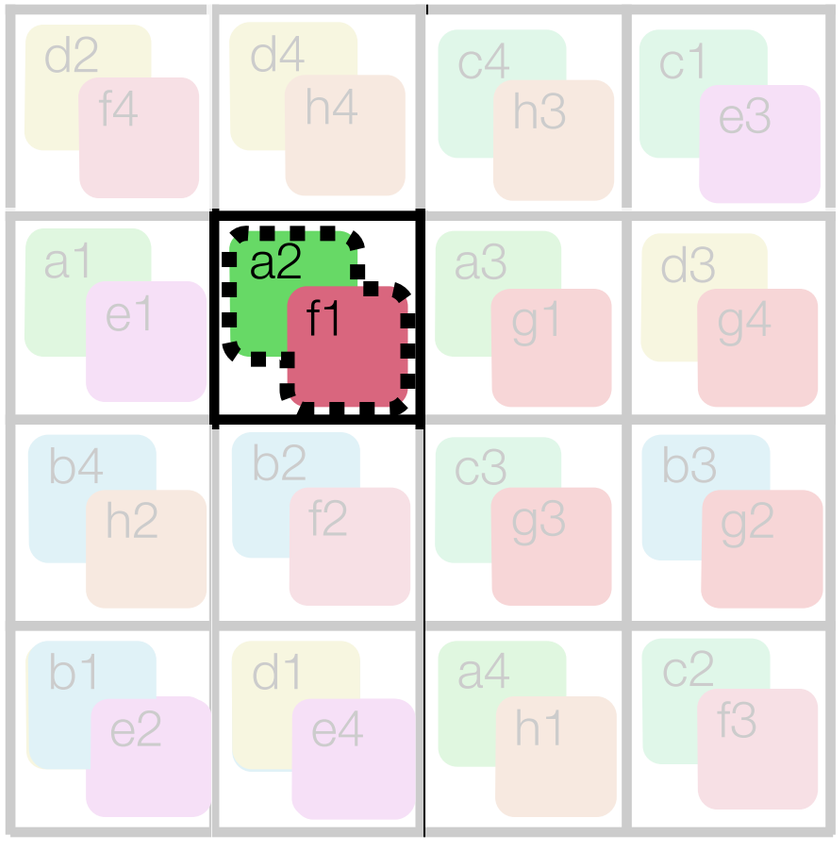
Down song part: *f1*



2. CrossSong puzzle: confirming progress

- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours

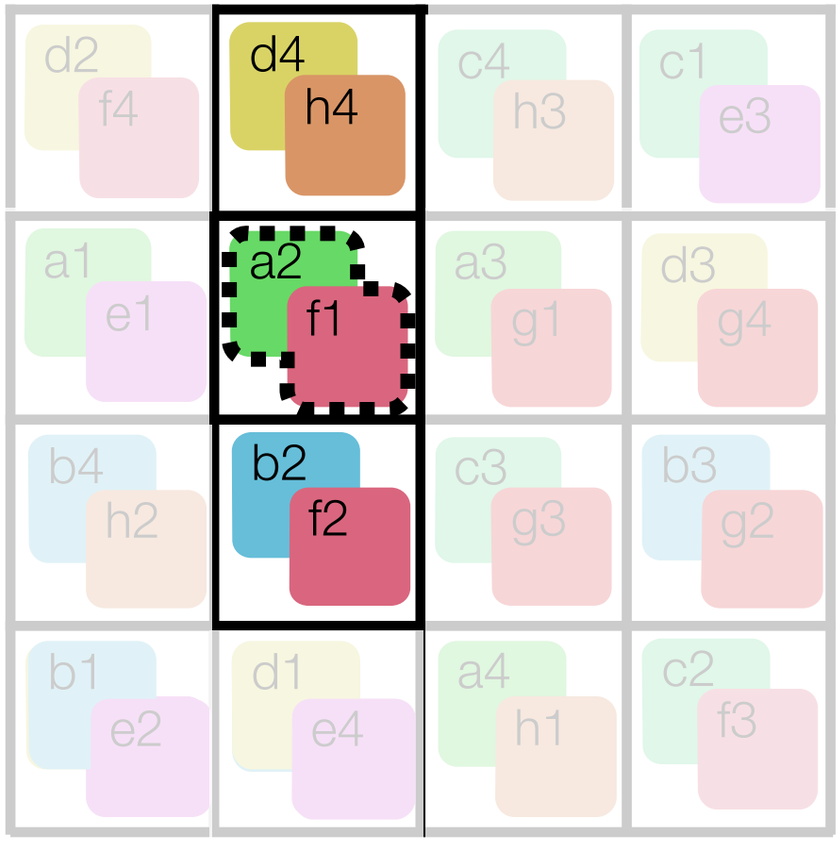
Normal balance: 50/50 *a2/f1*



2. CrossSong puzzle: confirming progress

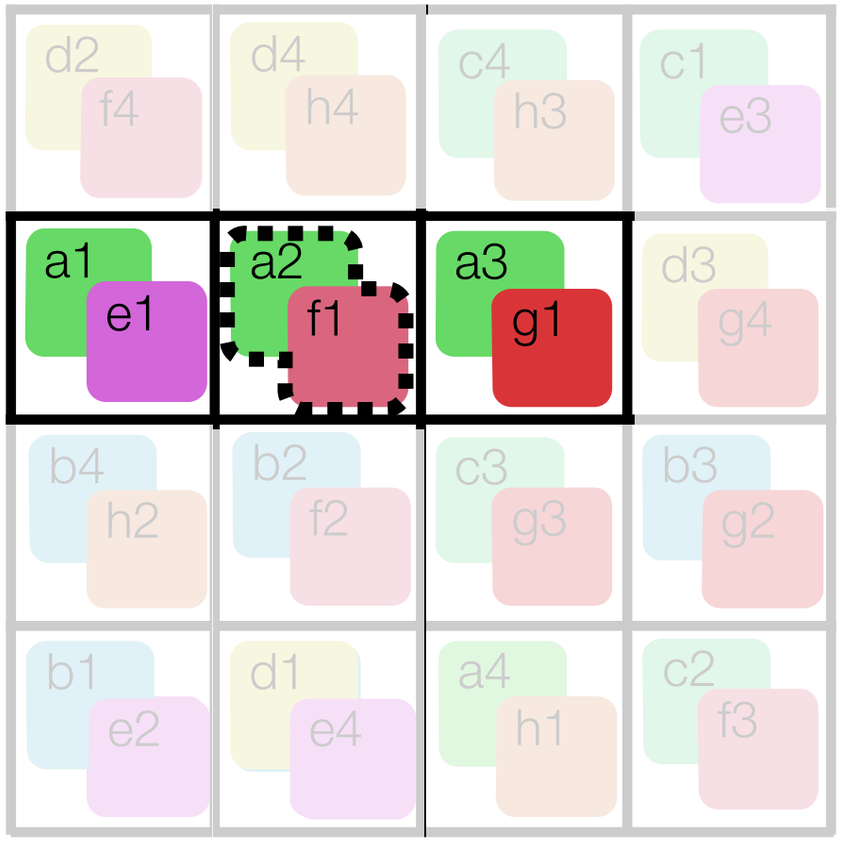
- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours

One neighbour in playing direction correct → audio balance favours correct direction (*f1* louder in playback of column)



2. CrossSong puzzle: confirming progress

- Audio feedback:
    - Balance of mash-up in a single tile depends on tile's correct neighbours



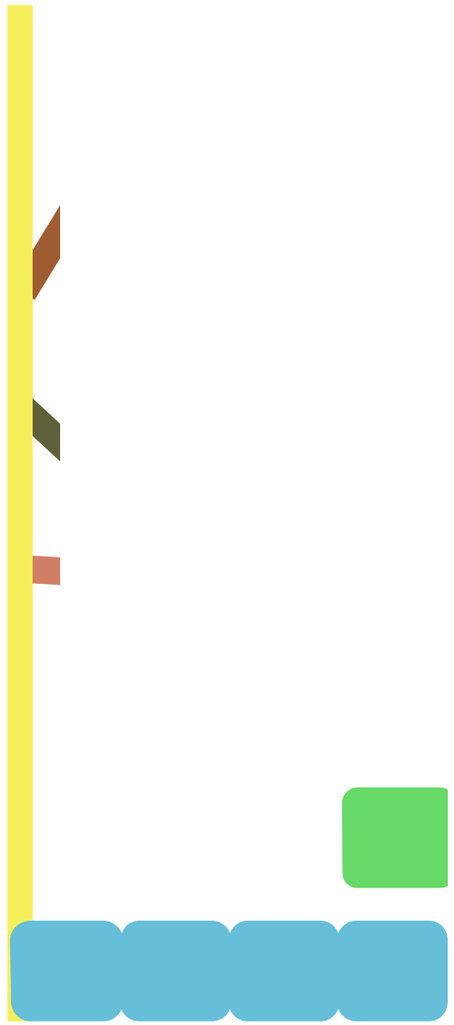
Both neighbours in playing direction correct → audio isolated (all *a2* in playback of row)

Problem: how do we create puzzles?

- Taking random excerpts, the tiles will sound terrible!
- Designing a puzzle by hand is too difficult

3. Puzzle creation algorithm

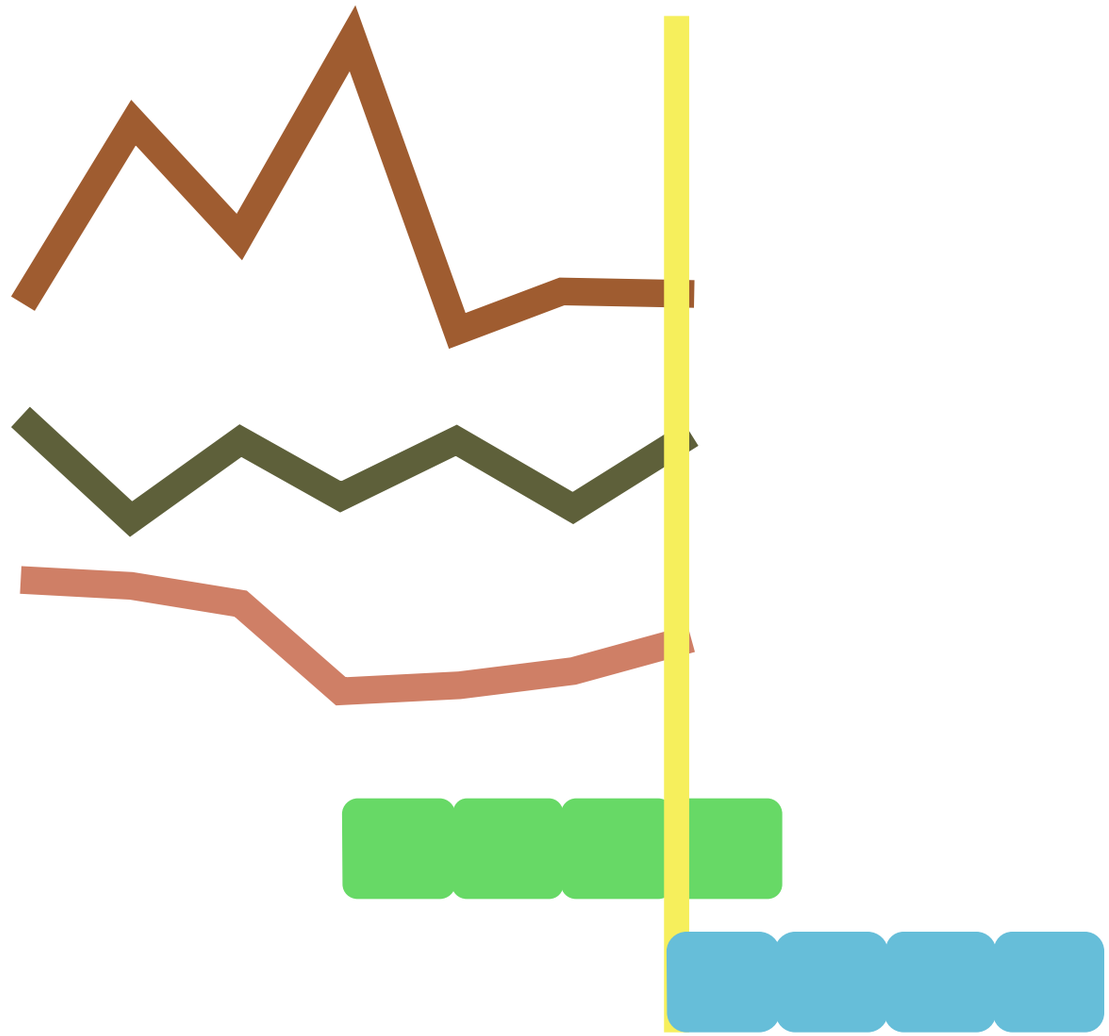
- “Mashability” estimation algorithm (Davies et al. 2014)
  - Estimates the quality of the match between two aligned music sequences
    - Harmonic match:
    - Rhythmic match:
    - Spectral balance:



- M. E. P. Davies, P. Hamel, K. Yoshii and M. Goto. 2014. AutoMashUpper: Automatic creation of multi-song music mashups. *IEEE Transactions on Audio, Speech, and Language Processing* 22 (12). 1726–37.

3. Puzzle creation algorithm

- "Mashability" estimation algorithm (Davies et al. 2014)
  - Estimates the quality of the match between two aligned music sequences
    - Harmonic match:
    - Rhythmic match:
    - Spectral balance:



- M. E. P. Davies, P. Hamel, K. Yoshii and M. Goto. 2014. AutoMashUpper: Automatic creation of multi-song music mashups. *IEEE Transactions on Audio, Speech, and Language Processing* 22 (12). 1726–37.

3. Puzzle creation algorithm

1. Detect section breaks

   - We will only consider 4-measure excerpts at the beginning of a section

2. Compute the mashability of all excerpts with each other

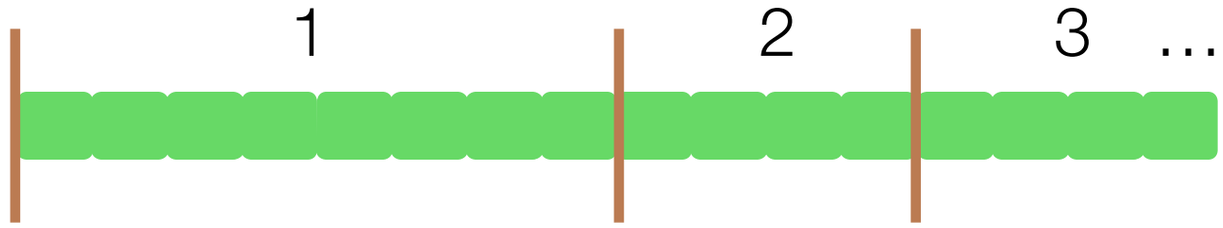
   - Automatic mashability estimation (Davies et al. 2014)

3. Search space of arrangements to find the one with the greatest estimated mashability

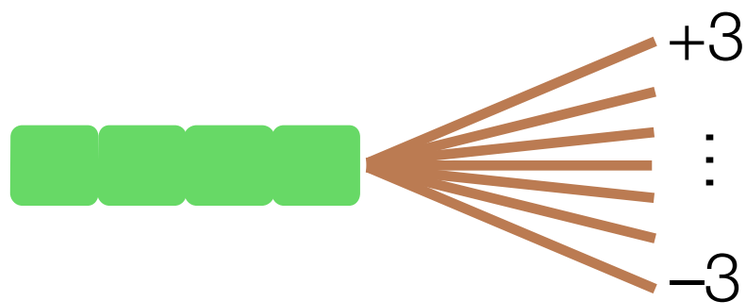
4. Apply pitch shifting / time stretching / loudness matching

3. Puzzle creation algorithm

- Each song: ~ 15 segments to choose from



- Each segment: can be transposed up to 3 semitones



- Each set of segments: can be arranged in 8!/2 ways

Total number of solutions to check:

$$15^8 \times 7^8 \times 8! / 2 \sim 10^{20}$$

3. Puzzle creation algorithm

- Search shortcuts and search space reductions:
  - Only consider **optimal transpositions**
    - reduce by factor of 10^6
  - Use **simulated annealing** to converge faster (new)
    - reduce by factor of ~4
  - **Ignore repetitions** to reduce space of segments (future)
    - reduce by factor of ~10^5

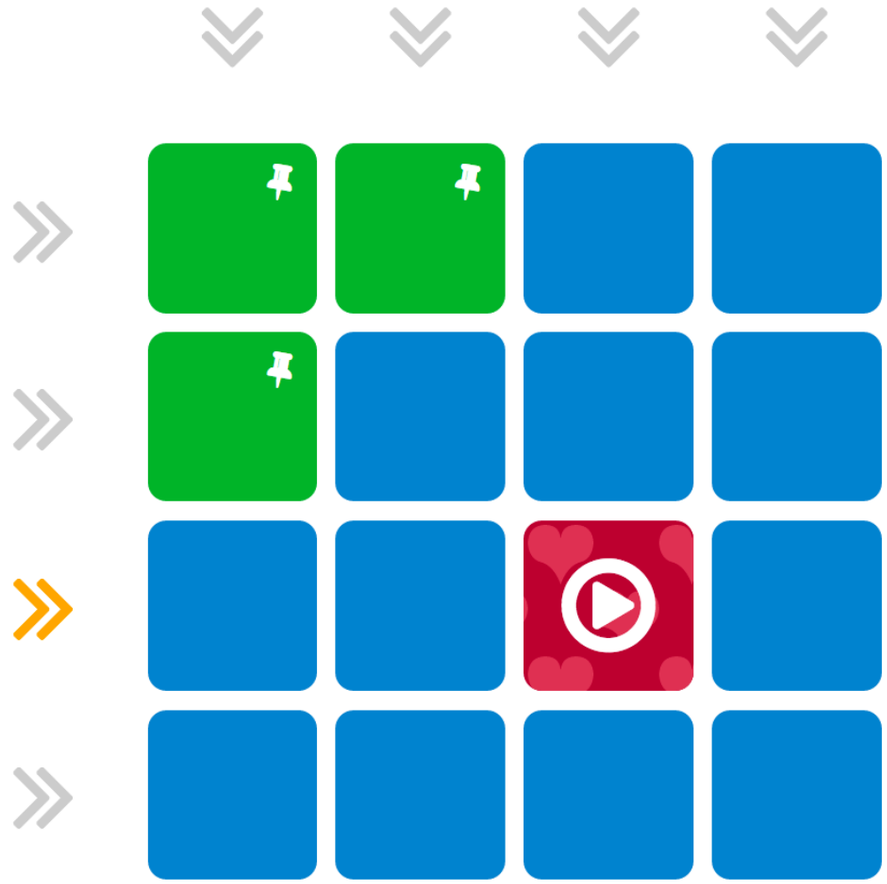
4. Conclusion

- CrossSong is a **music**-based **puzzle game** with real-time gameplay

- **Constructing puzzles** is hard, but mashability estimation and search algorithm make it **easy**

- Works with many kinds of music

- **It's fun!**
  …but please judge for yourself:
  https://staff.aist.go.jp/jun.kato/CrossSong/



Thank you!

CrossSong playable at:

https://staff.aist.go.jp/jun.kato/CrossSong/

Works on any device with a Chrome browser

If you send us your music, we'll turn it into a puzzle!

→ jordan.smith@aist.go.jp

AIST