

Computational Model for Automatic Chord Voicing based on Bayesian Network

***Tetsuro Kitahara, Makiko Katsura,
Haruhiro Katayose, and Noriko Nagata***

***CrestMuse Project, CREST, JST, Japan/
Kwansei Gakuin University, Japan***

Our Goal

To develop a computational model for automatic *chord voicing*

Chord voicing:

Simultaneous vertical placement of notes in order to obtain the effective sound of a chord

The image displays two musical staves illustrating chord voicing. The top staff shows a sequence of chords: G, CM7, Gm7, C7, and FM7. A green arrow points down to a second staff that shows the same sequence of chords with their respective voicings in the bass clef. The bottom staff shows a bass line with notes corresponding to the chords.

What's the difficulty in chord voicing

3 issues:

Extended notes

e.g. 9ths, 11ths, ...
Essential especially
for jazz music


Omitted notes

Only adding extended
notes may cause
dissonant sound.
Physical limitation

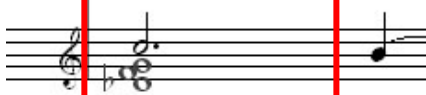
Inversions

Important to make
temporally smooth
voicings

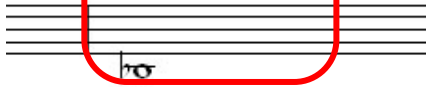
DbM7



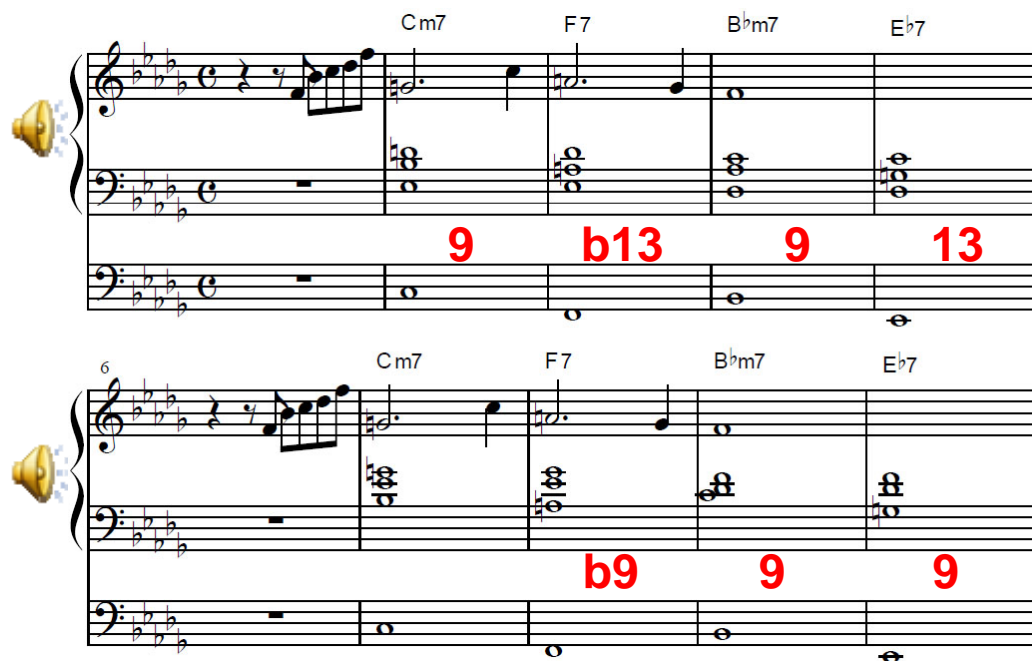
DbM7(#11)



DbM7(9,13)



A red box highlights the notes of the DbM7 chord in the first staff.



Chord progression: Cm7, F7, Bbm7, Eb7

Extensions: 9, b13, 9, 13

Chord progression: Cm7, F7, Bbm7, Eb7

Extensions: b9, 9, 9

Two musical staves showing chord progressions. The first staff shows Cm7, F7, Bbm7, Eb7 with extensions 9, b13, 9, 13. The second staff shows Cm7, F7, Bbm7, Eb7 with extensions b9, 9, 9. A yellow bell icon is present on the left of each staff.

What's the difficulty in chord voicing

3 issues:

Extended notes

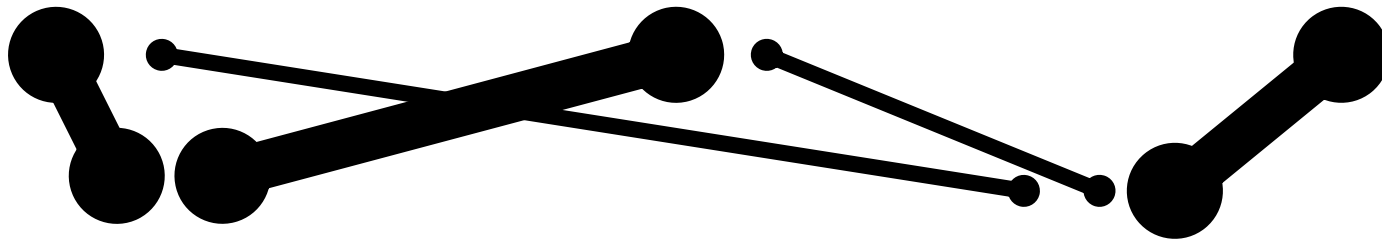
e.g. 9ths, 11ths, ...
Essential especially
for jazz music

Omitted notes

Only adding extend-
ed notes may cause
dissonant sound.
Physical limitation

Inversions

Important to make
temporally smooth
voicings



Many-to-many
relationship

2 aspects:

Musical simultaneity

Matching with the simulta-
neous notes of other parts
(e.g. melody)

Musical sequentiality

Smooth connection of
previous and following
chords

3 possible solutions

- ***Rule-based***

Design a set of rules manually

--- [Emura, 2008]

- ***Case-based***

Use existing examples (corpus) of chord voicings

- ***Instance-based***

e.g. Find part of the target chord progression in a corpus and copy-and-paste their voicings

--- [Hirata, 2001] etc.

- ***Probability-based***

e.g. Train a probabilistic model with a corpus

--- Few trials in previous studies

Merits (M) and Demerits (D)

- ***Rule-based***

[M] If an expert carefully design rules, high quality can be expected.

[D] Difficult to design a good rule set.

⇒ Rules satisfying both simultaneity and sequentiality

[D] Difficult for users to control the behavior

- ***Case-based***

[M] No need to design a thorough rule set.

[M] Users can control the behavior by switching the corpus.

[D] If a probabilistic model, a large corpus is needed.

In this study...

To develop a *probabilistic* model for automatic chord voicing

Aim: To show that...

- We can generate voicings satisfying simultaneity and sequentiality with a model where they are described as probabilistic dependencies.
- If a particular genre corpus is used, the generated voicing also has the characteristics of that genre.

Problem Statement

- **Target instrument:** Electronic organ (Electone)
 - Upper (right-hand) keyboard: Melody line
 - Lower (left-hand) keyboard: Chord
 - Pedal keyboard: Bass line
- **Input:** Data written on a lead sheet
 - A melody line
 - a sequence of chord names
- **Output:** A left-hand voicing and a bass note for each chord
 - No passing chords or passing bass notes are added.
- **Target genre:** Jazz

The image displays a musical score transformation. The top staff shows a melody line in treble clef with four measures. Above the notes are chord names: G, CM7, Gm7, and C7. A large green arrow points downwards to a second staff, which is a full score. This staff includes the same melody line, but also adds a left-hand keyboard part in bass clef with chord voicings and a separate bass line in bass clef with single notes. The chord voicings in the left hand are G, CM7, Gm7, and C7, corresponding to the chords in the lead sheet. The bass line consists of single notes: G, C, Bb, and C.

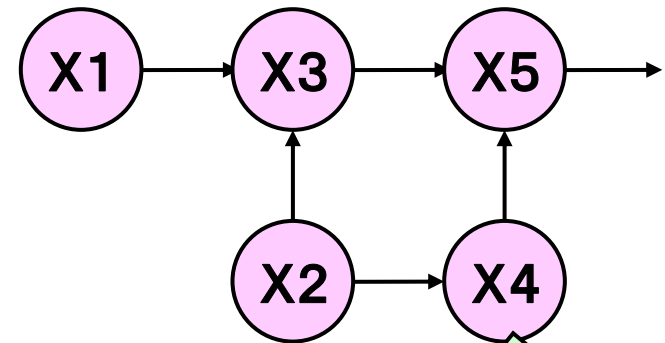
Probabilistic model we use

We use a *Bayesian network*

Bayesian network:

a probabilistic graphical model that represents variables and their probabilistic independencies

- Nodes: Random variables (typically discrete)
- Links: Conditional dependencies between variables
- Each node has a conditional probability table (CPT).
- After observable variables are set, the probabilities of the others are estimated.



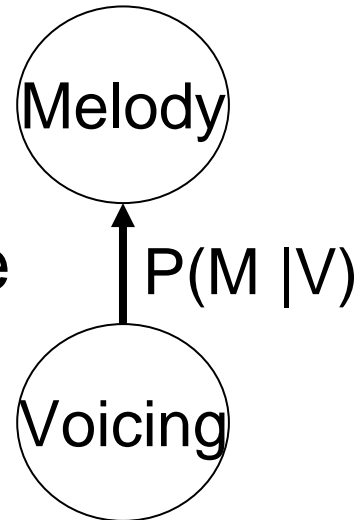
CPT (Conditional Probability Table)

	X2		
X4		0	1
P(X4 X2)	0	0.8	0.4
	1	0.2	0.6

Basic policies

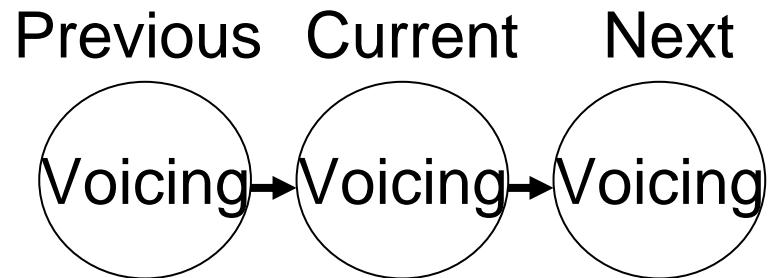
- **Simultaneity**

- Melody and voicing nodes are linked.
- We consider an observed melody to be generated from the harmony including the voicings.



- **Sequentiality**

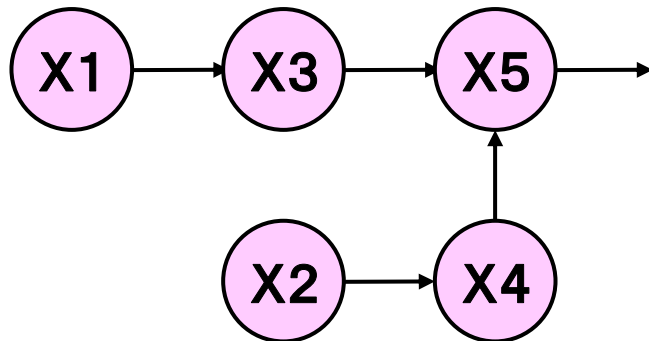
- The voicing nodes for the previous, current, and next chords are linked.
- The nodes for top, middle, and bottom notes are separated.



Basic policies

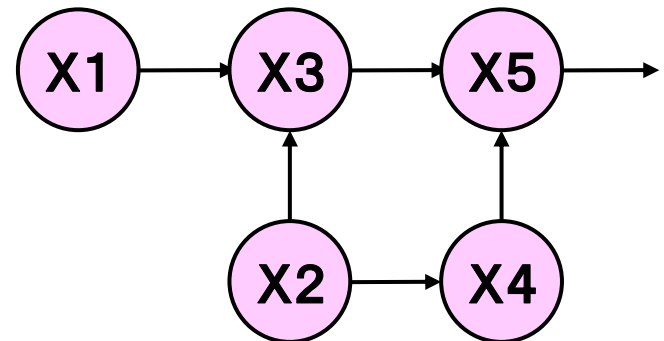
- Design a singly connected network

Singly connected network



A simple $O(n)$ algorithm for probability calculation exists.

Multiply connected network



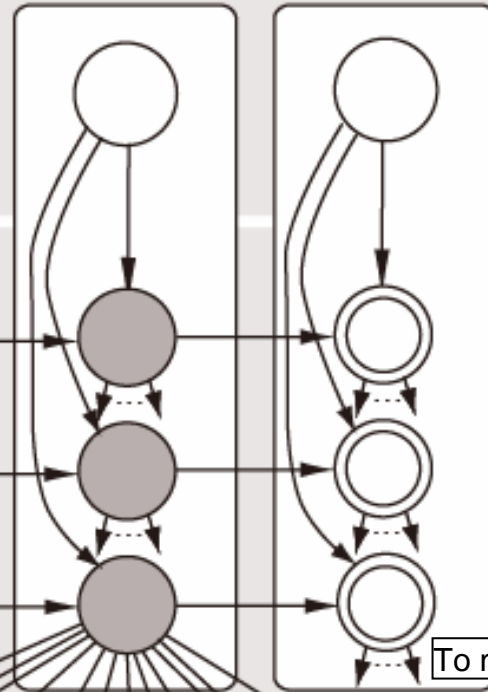
A complicated algorithm is needed.

- Design separate models for left-hand voicing and bass note determination
 - Models should be simple for training with limited data

Left-hand voicing model

Previous Current Next

Chord names



$\{C, C\#, \dots, B\}$
 $\times \{\text{maj.}, \text{min.}, \dots\}$

Voicing

bottom note

middle notes

top note

$\{C, C\#, \dots, B\}$

$\{C, C\#, \dots, B\}^+$

$\{C, C\#, \dots, B\}$

To melody nodes

Melody

From voicing nodes

$\{0.0, 0.2, \dots, 1.0\}$
Relative length of appearance of each pitch class

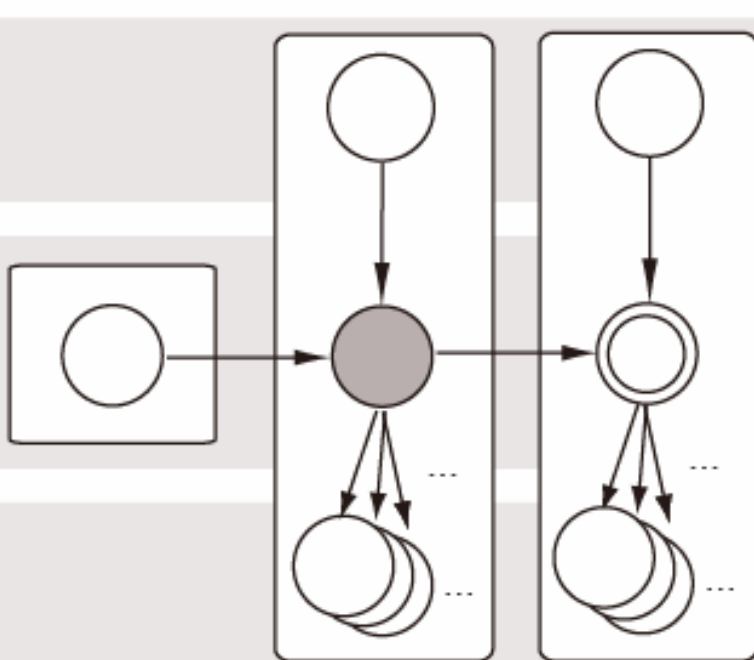
Bass model

Previous Current Next

Chord names

Bass

Melody



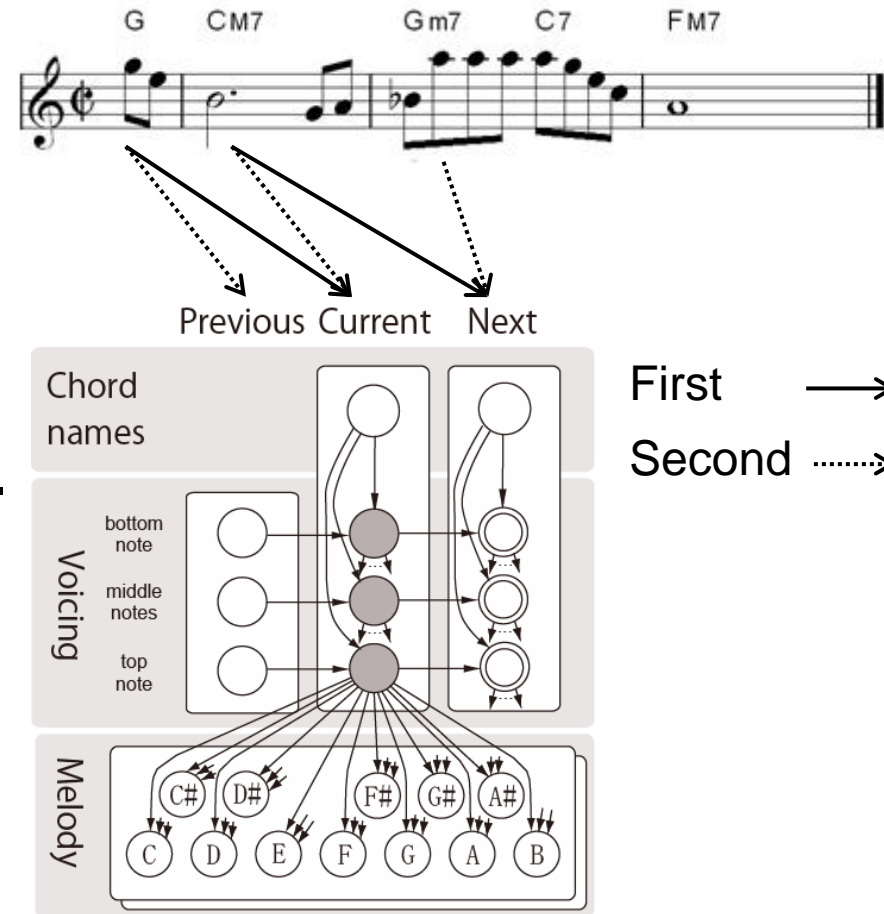
{C, C#, ..., B}
× {maj., min., ...}

{C, C#, ..., B}

{0.0, 0.2, ..., 1.0}
Relative length of
appearance of each
pitch class

System implementation

- Our models are applied to each chord from the beginning to the end of a given chord progression by shifting the current chord.
- After the probabilities of the nodes are calculated, the note with the highest probability is chosen from each current voicing node.
- Pitch range
Left-hand: C3-A#4
Bass: C2-G3
- Training: 30 jazz pieces



Example of Outputs

“Misty” 🎷
(by Erroll Garner)

Left-hand voicing

- Extended notes were added.
- Natural inversions were chosen.

Bass

- Mostly root notes were chosen.

The image displays a musical score for the jazz standard "Misty" by Erroll Garner, featuring three systems of music. Each system consists of a treble clef staff with a melody line and two bass clef staves for the left hand. The score is annotated with various boxes and callouts to highlight specific voicing and bass line choices.

- System 1:** Chords G, CM7, Gm7, C7, FM7, Fm, Bb. A red box highlights a chord voicing in the second bass staff, with a callout pointing to it.
- System 2:** Chords CM7, Am, G7, E7, A7, D7, G7. A red box labeled "b13th" points to a chord voicing in the second bass staff. Another red box labeled "9th" points to a chord voicing in the third bass staff.
- System 3:** Chords CM7, Gm7, C7, FM7, Fm, Bb. A yellow box highlights a chord voicing in the second bass staff. A red box labeled "9th" points to a chord voicing in the third bass staff. A yellow box labeled "Top tones are smoothly connected" points to the melody line in the treble staff. A blue box labeled "Non-root note" points to a chord voicing in the third bass staff.
- System 4:** Chords CM7, D, G7, C, Am7, Dm7, G7, Bdim, C6. A red box highlights a chord voicing in the second bass staff. A blue box labeled "Non-root note" points to a chord voicing in the third bass staff.

Qualitative evaluation of simultaneity and sequentiality

1. No previous- and next-chord nodes

- Extended notes match the melody.
- ✗ Temporal connection is not smooth.

Chord progression: G, CM7, Gm7, C7, FM7

The score shows a melody in the treble clef and a bass line in the bass clef. Red arrows point from the bass line notes to the corresponding notes in the melody, indicating a lack of smooth temporal connection.

2. No melody nodes

- Temporal connection is smooth.
- ✗ Some notes do not match the melody

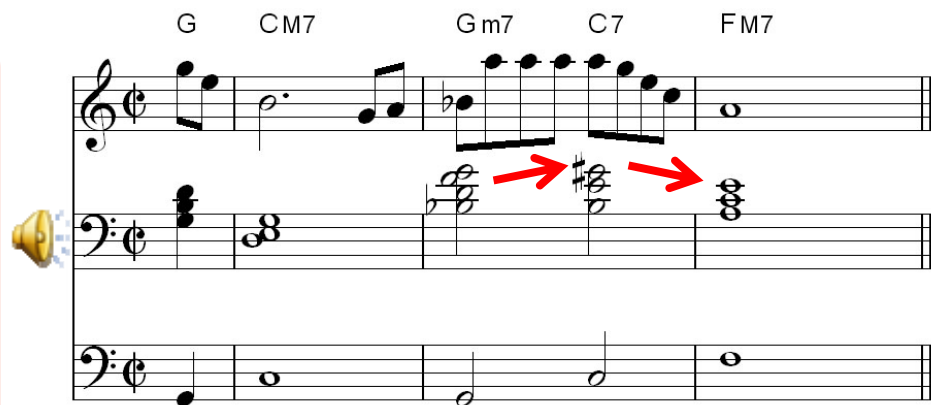
Chord progression: G, CM7, Gm7, C7, FM7

The score shows a melody in the treble clef and a bass line in the bass clef. Red question marks are placed above the bass line notes for the Gm7 and C7 chords, indicating that some notes do not match the melody.

Qualitative evaluation of simultaneity and sequentiality

3. Our model

- The matching with the melody
- Smooth temporal connection



A musical score for guitar and bass. The guitar part is in the treble clef, and the bass part is in the bass clef. The score is in 4/4 time. The guitar part has a melody line with notes and rests. The bass part has a bass line with notes and rests. Above the guitar staff, the chords G, CM7, Gm7, C7, and FM7 are indicated. A speaker icon is located to the left of the bass staff. Two red arrows point from the Gm7 chord to the C7 chord and from the C7 chord to the FM7 chord, indicating a smooth temporal connection between the chords.

Subjective evaluation of simultaneity by music experts

Subjects rate the voicing for each chord

◎ : Beautiful

○ : Good

△ : Acceptable

× : Unacceptable

3 subjects

e.g. Electone instructor
Demonstrator

5 pieces (608 chords)

Results

	# of chords	Percentage
Beautiful	82	9.0%
Good	719	78.8%
Acceptable	63	6.9%
Unacceptable	48	5.3%

94.7%

Subjective evaluation of sequentiality by music experts

Subjects rate continuity of each chord change

○ : Good

△ : Acceptable

× : Unacceptable

2 subjects

5 pieces
(299 chord changes)

Results

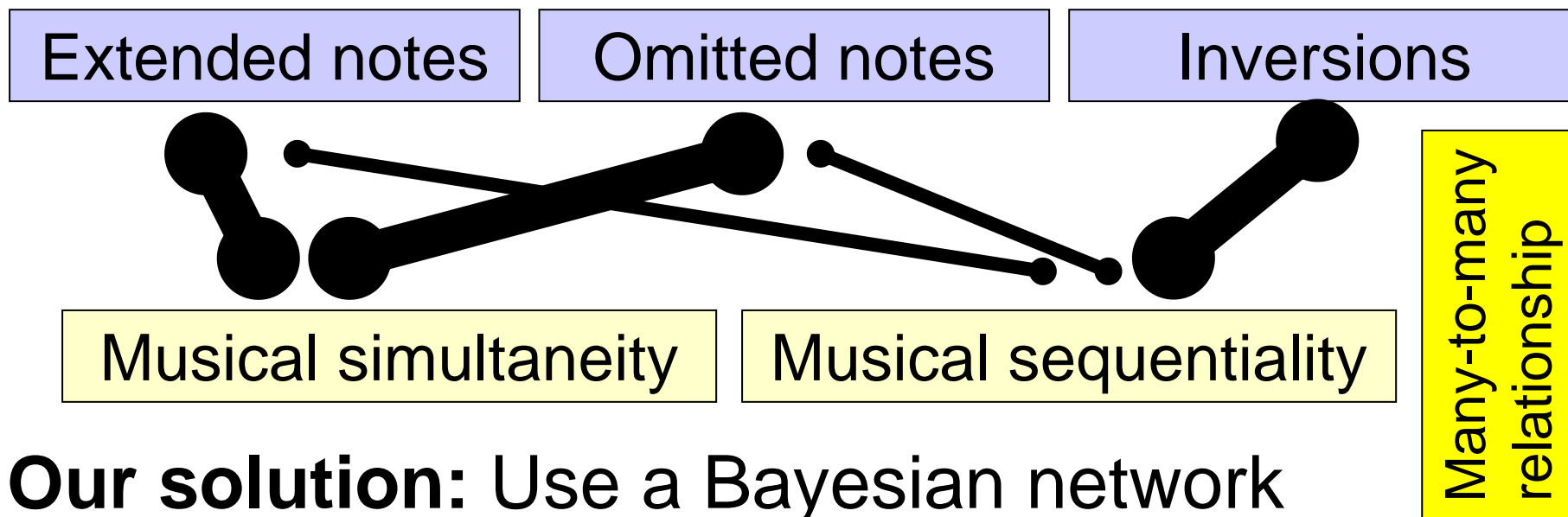
	# of chord changes	Percentage
Good	180	25.8%
Acceptable	407	58.3%
Unacceptable	111	15.9%

84.1%

Conclusions

Our goal: To develop a computational model for automatic chord voicing

Issue:



Our solution: Use a Bayesian network

Results: Generated voicings satisfy both simultaneity and sequentiality

