Chord-conditioned Melody and Bass Generation

Alexandra C. Salem[1], Mohammad Shokri[1], Johanna Devaney[1,2]
[1]CUNY Graduate Center, [2]Brooklyn College

LUMaA

INTRODUCTION

- Our goal was to **generate melody and bass scores conditioned on a chord input**
- We evaluate five **Transformer**-based [3] strategies for this task
- We use music-theory motivated metrics capturing **pitch content**, **pitch interval size**, and **chord tone usage** in generated files

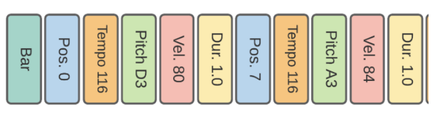
DATA

We use the **TAVERN** (Theme and Variation Encodings with Roman Numerals) corpus [1]:

- 27 sets of themes and variations composed by **Mozart** and **Beethoven**
- Piano, High classical style
- Training:
  - 25 sets of theme and variations
  - 1,702 phrases
- Testing:
  - 2 sets (1 Beethoven, 1 Mozart)
  - 164 phrases

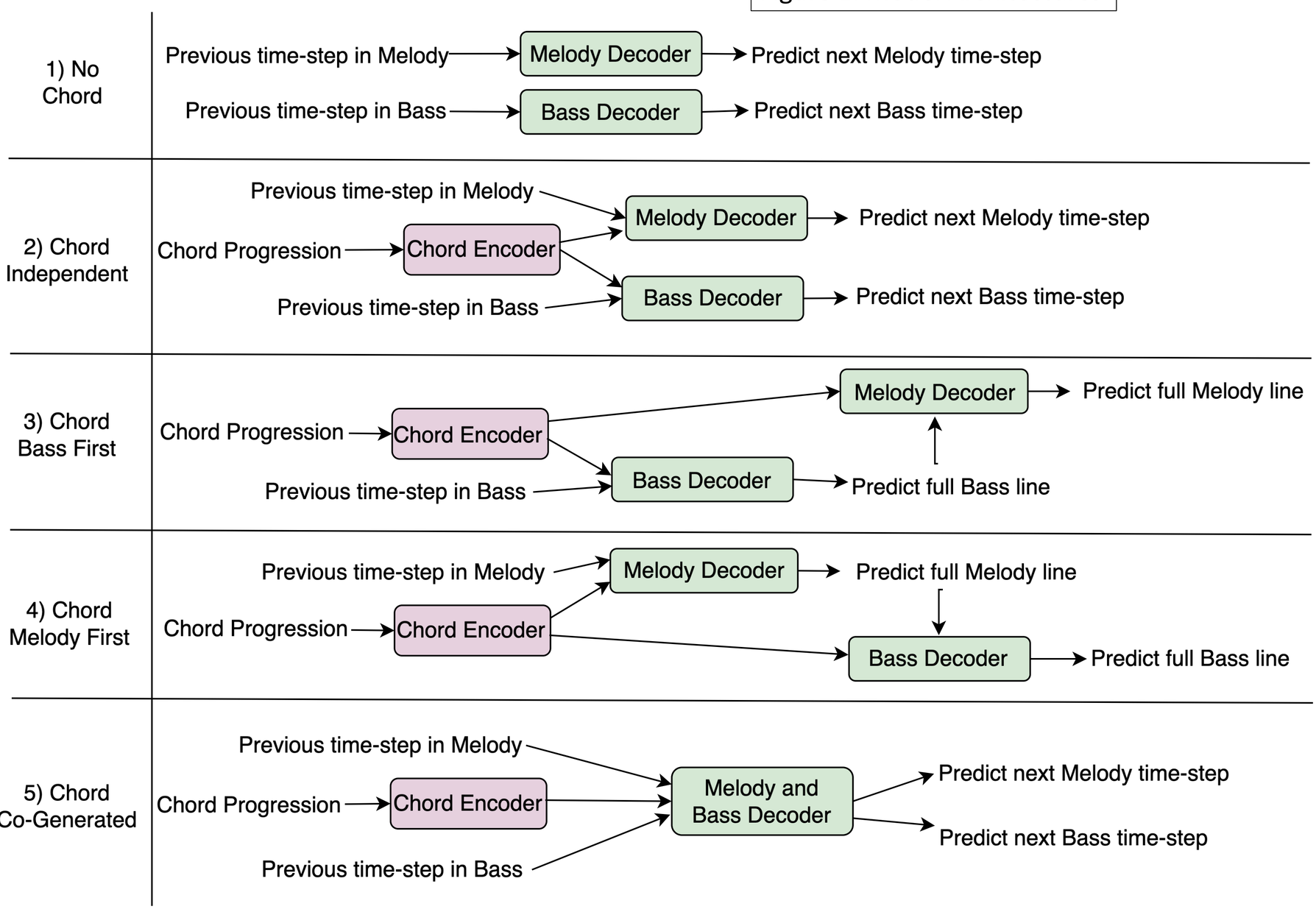
Pre-processing:

- Create **score skeletons** that capture overall bass or melody
- Transpose to C major or C minor
- Chord input :
  - **kern file [2] → Chord sequence
  - Ex: C:min G:maj G:maj C:min
- Melody and Bass input:
  - **kern file → MIDI → REMI
  - Ex:



TRANSFORMER-BASED MODELS

Figure 1. Model Architectures



EVALUATION & RESULTS

Pitch Content

Pitch class histogram entropy

pitch classes used / # notes used

of unique pitches

Range in frequency of pitches

Unique pitch class ratio, melody vs bass
Pitch consonance score, melody vs bass

Pitch Interval Size

Average pitch interval size

Chord Tone Usage

Chord-tone to non-chord tone ratio

Table 1. Metrics Across Model Architectures

Metric	Voice	No Chord	Chord *Ind.*	Chord *Mel-1st*	Chord *Bass-1st*	Chord *Co-Gen*	GT
PC Entropy	Mel.	**1.456**	1.627	1.574	1.665	1.773	1.506
	Bass	1.210	**1.540**	**1.523**	**1.423**	**1.547**	1.474
PCs Used	Mel.	0.519	**0.501**	0.521	**0.493**	0.528	0.482
	Bass	0.433	**0.491**	**0.501**	**0.464**	**0.476**	0.471
Unique Pitches	Mel.	5.805	**7.494**	**6.683**	9.421	8.841	7.152
	Bass	5.280	7.695	**7.207**	**6.811**	7.927	7.335
Pitch Range	Mel.	12.634	**15.323**	**13.939**	19.945	16.884	17.390
	Bass	14.177	**17.976**	**17.646**	**18.049**	**17.573**	15.211
Unique PC Ratio		**0.853**	**0.836**	**0.839**	0.840	0.864	0.820
PCS		0.248	**0.360**	0.261	0.333	0.319	0.402
Pitch Interval	Mel.	**3.356**	**3.296**	3.298	**4.216**	**3.274**	4.467
	Bass	4.001	**4.104**	**4.423**	**4.128**	3.654	3.731
CT Ratio	Mel.	NA	0.728	0.714	**0.716**	0.744	0.662
	Bass	NA	**0.676**	0.717	**0.683**	0.721	0.611

CONCLUSIONS

- All models benefited from the information in the chords
- **Best-performing model** was **Chord Bass-1st**
  - This aligns with how students are typically instructed to do this task
- Somewhat surprisingly, Chord Co-Gen did not perform as well
  - Bass-melody generation ordering is more beneficial than jointly estimating
- In future work, we will test additional model variations
- This work presents a first step in a **pedagogy-inspired musical score generation** approach that could benefit downstream tasks such as chord-controlled training data generation

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (NSF) award 2228910.

KEY REFERENCES

[1] Johanna Devaney et. al. Theme and variation encodings with roman numerals (tavern): A new data set for symbolic music analysis. In Proceedings of International Society for Music Information Retrieval Conference (ISMIR), pages 728–34, 2015.

[2] David Huron. Music information processing using the humdrum toolkit: Concepts, examples, and lessons. Computer Music Journal, 26(2):11–26, 2002.

[3] Ashish Vaswani et al. Attention is all you need. In Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc., 2017.