

Score Transformer

Generating Musical Score from Note-level Representation

Masahiro Suzuki - R&D Division, Yamaha Corporation



Background

- ✓ MIDI tokens x Transformers = fruitful results (reported)
 - Score tokens x Transformers = ??? } unexplored
 - ↳ How to tokenize scores
- ↳ We explore **score token** representation to **generate scores** with **transformers**

Method

1. **Design** score token representation
2. **Tokenize** MIDI and score data
3. **Train** the Transformer model in a seq-to-seq manner

down-convert

Musical Score (Original)



measure differences

MIDI (Down-converted)

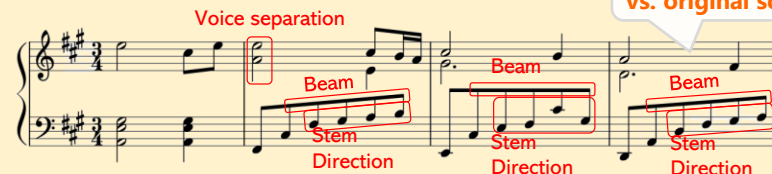


tokenize

```
bar note_76 len_48 note_56 len_48 note_52 len_48 note_45 len_48 beat beat note_73 len_12 note_56 len_24 note_52 len_24 note_45 len_24 pos_12 note_76 len_12 bar ...
```

MIDI Token Representation (REMI-based)

Musical Score (Generated)



differences vs. original score

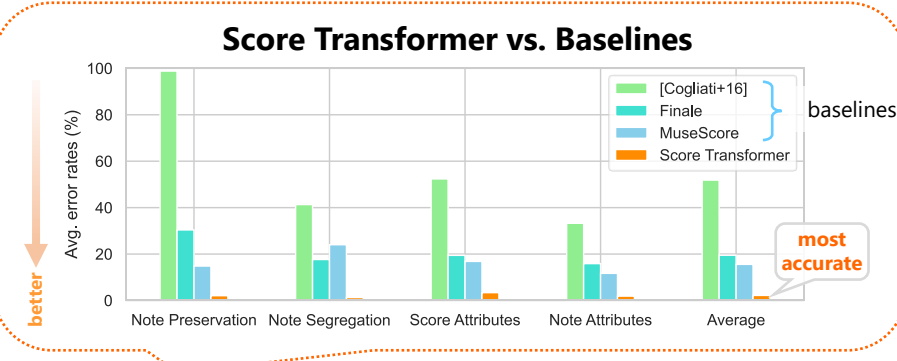
detokenize

```
R bar clef_treble key_sharp_3 time_3/4 note_E5 len_2 stem_down note_C#5 len_1/2 stem_down beam_start note_E5 len_1/2 stem_down beam_stop bar ... L bar clef_bass key_sharp_3 time_3/4 note_G#3 note_E3 note_A2 len_2 stem_down note_G#3 note_E3 note_A2 len_1 stem_down bar ...
```

Score Token Representation

train / infer with Transformer

Score token performs the best.



Get the tokenization tools!

<https://github.com/suzuqn/ScoreTransformer>

- ### Token design:
1. One token per symbol or attribute
 2. music21-friendly attributes
 3. Concatenated R and L sequences
- Can be **shorter** sequences w/o performance degradation

Experiment

- [Dataset] 7,161 pop / 354 classical piano scores
- [Model] A small vanilla Transformer (param: ~4M)
- [Metric] Error rates for 12 musical aspects / 4 categories

Results

- ✓ Score Transformer works **significantly better** than baselines.
- ✓ Score token **outperforms** existing formats.
- ✓ Score Transformer also works with **unquantized** data.

