

Impact of time and note duration tokenizations on symbolic music modeling

Nathan Fradet, Nicolas Gutowski, Fabien Chhel, Jean-Pierre Briot

Background

- Music modeling is commonly handled with tokens
- Music can be tokenized by different ways
- Time, note duration, pitch, instrument ...
- How does this tokenizations impact model perf?



Focus on time and note duration



What explicit information?

- TimeShift —> time intervals
- Bar + Position —> Onset / Offset positions
- Duration —> Duration
- NoteOff —> Note offsets



Tokenizat	tion Top-20 compose	ers \uparrow Top-100 composer	s \uparrow Emotion \uparrow
TS + Dur	· 0.973	0.941	0.983
TS + NOf	ff 0.962	0.930	0.962
Pos + Du	r 0.969	0.927	0.963
Pos + NO	<i>Off</i> 0.963	0.925	0.956

TS+Dur TS+NOff Pos+Dur Pos+NOff -0.2

TS+Dur TS+NOff Pos+Dur Pos+NOf -0.2

Model considerations

• Transformers struggle at « reasoning » but are better at capturing explicit information

• Transformers can attend information differently

Causal attention: only past context



Full attention: past and future context

Classification

Sequence embedding



Mean cosine similarity between seq. embedding and augmented seq. embedding

Tokenization	lPCA \uparrow	$\mathbf{MOM} \uparrow$	TwoNN \uparrow	FisherS ↑
TS + Dur	213	42.6	34.3	17.5
TS + NOff	161	43.7	32.7	17.5
Pos + Dur	146	39.1	33.1	17.1
Pos + NOff	177	45.2	35.6	17.8

Intrinsic dimension estimations of the learned sequence embeddings



Prediction error ratios: token type, time, duplicated note, NoteOff errors







Generation

$\mathbf{TSE_{type}} \downarrow$	$\mathbf{TSE_{time}} \downarrow$	$\mathbf{TSE_{dupn}} \downarrow$	$\mathbf{TSE_{nnon}}\downarrow$	$\mathbf{TSE_{nnof}} \downarrow$
$< 10^{-3}$	_	0.014	-	-
$< 10^{-3}$	-	0.001	0.109	0.040
0.002	0.113	0.032	-	-
0.002	0.127	0.005	0.095	0.066

Continuation generated from the same prompt with the four strategies, from a checkpoint during training

Note characteristics distributions

Future work

• This is only a step in symbolic music modeling • Experiment on other tasks / tokenizations • Design musical reasoning tasks to assess models