Playing Along in Time and Key

Autoregressive transformer-based generator follows beat and chord progression using causal predictors

Kevin Bretz Derya Soydaner Aske Plaat LIACS, Leiden University ISMIR 2026 Late-Breaking Demo

kevin-bretz.github.io | linkedin.com/in/kevin-bretz | k.o.bretz@umail.leidenuniv.nl

Scan for the demo, the website and me

DEMO VIDEO | WEBSITE | LINKEDIN

kevin-bretz.github.io

full paper in the works, more coming soon, follow for updates

1 The musicians play, the AI listens

input

heard so far | now | not yet heard

causal audio codec

neural network

plain computation

input comes from outside, the musicians or you

2 It tracks and predicts tempo

on a CPU at a few percent of real time

beat predictor

tempo and phase tracker

a dynamic Bayesian network (DBN), a bar-pointer model, turns the predictor's noisy guesses into one steady tempo and phase, so the clock never jumps

125 BPM

3/4

the same clock, over time

solid is the predicted value, dashed is the true value

3 ...and harmony

chord predictor

C | G

input texture reference

a short clip of the sound you want

clock

chords

texture

tokens

accompaniment generator

SMG, Wu et al.

SiMe, Bretz et al.

causal scorer

trained on the generator's own tries, labelled afterwards with the beat and realism scores they got, it predicts those scores for all eight and keeps the best

340 ms | 500 ms

on the GPU, an RTX 2080 Ti

4 Eight different rollouts, scored and ranked by a learned causal scorer

5 Chunk committed and queued for next boundary

0.5 s ahead

generated ahead

the committed chunk becomes the next context