

VampNet: Music Generation via Masked Acoustic Token Modeling

Hugo Flores García^b, Prem Seetharaman[#], Rithesh Kumar[#], Bryan Pardo^b

^bNorthwestern University, USA [#]Descript Inc.

hugofg@u.northwestern.edu

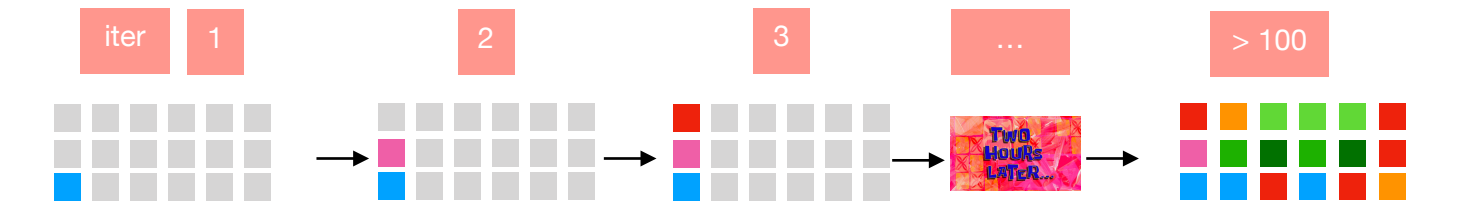
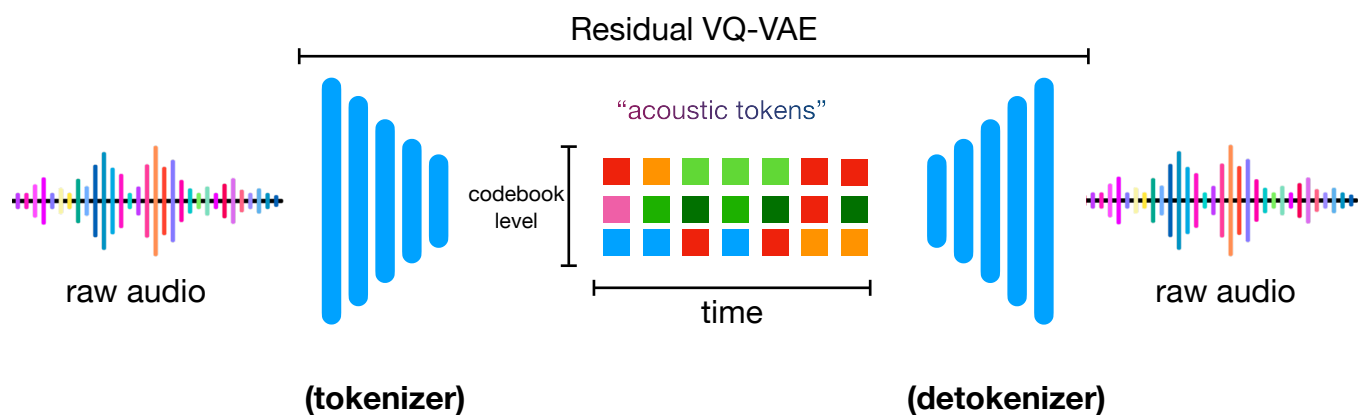


Audio Generation via Acoustic Token Modeling

Prev. Work: Autoregressive Acoustic Token Modeling

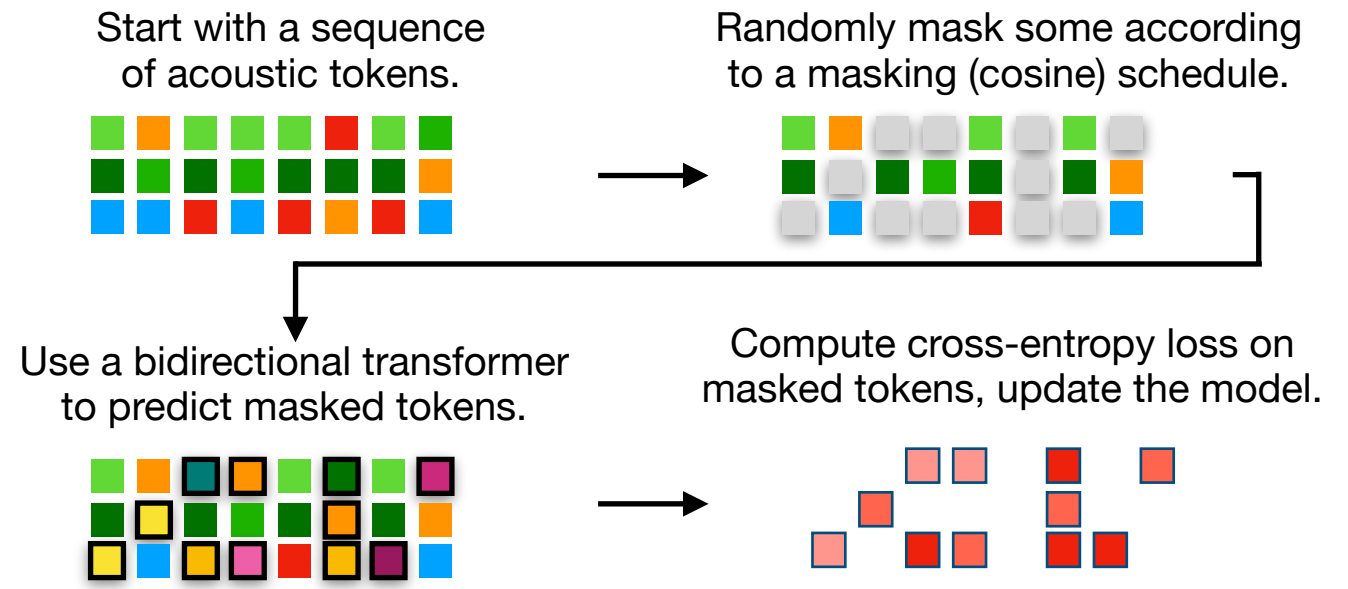
Stage 1. Tokenizing / Detokenizing

An RVQ-VAE encodes continuous audio signals into a hierarchical sequence of discrete tokens



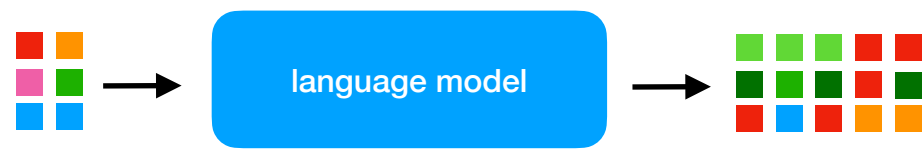
Autoregressive methods for acoustic token modeling are known to generate high quality sequences, but **require many (>100) inference steps** and **can only attend to the past**.

VampNet: Masked Acoustic Token Modeling

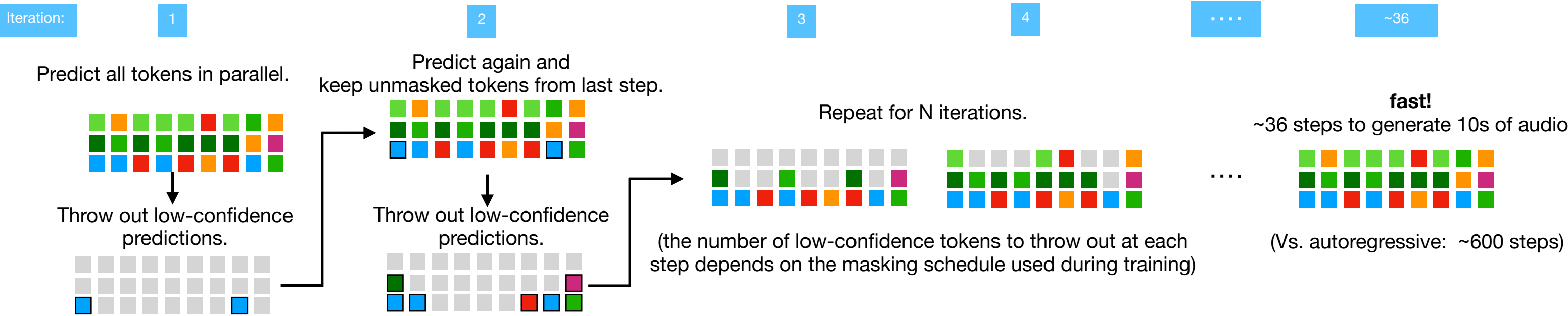


Stage 2. Language Modeling

A language model learns to generate new sequences of acoustic tokens, to later be detokenized into an audio signal.

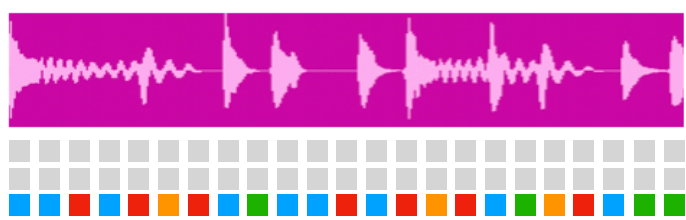


Sampling from VampNet: Iterative Parallel Decoding



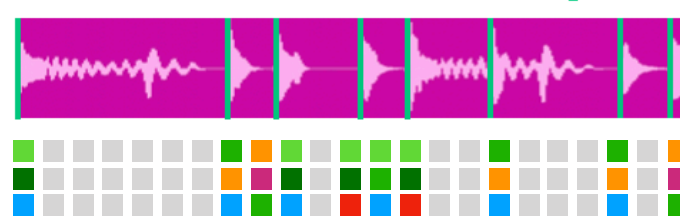
The (many) ways to prompt VampNet

Compression



- Remove (mask) all of the codebook levels except the lowest
- Use VampNet to generate all of the upper codebook levels
- Can compress music down to **~600bps** (ref: Soundstream: 3kbps)

Beat-Driven



Use a beat tracker to find all the beats in the music, and only **unmask at the beat, generate the rest** generates variations while preserving the beat!

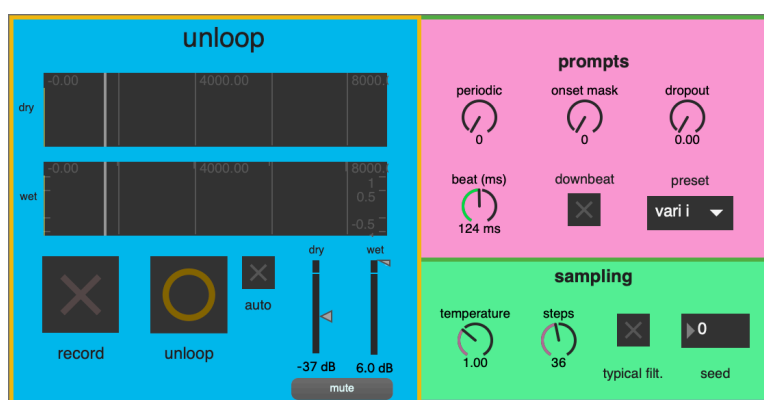
Periodic



Unmask only every **P**th timestep, where **P** is the *period* of the prompt.

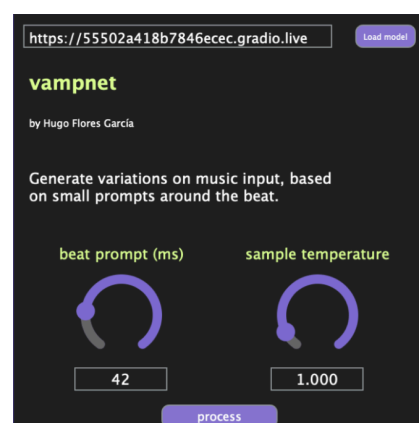
VampNet as a Co-Creation Tool

unloop: Vamp in Max/MSP



It's a looper that doesn't repeat itself!
(<https://github.com/hugofloresgarcia/unloop>)

HARP: Vamp in a DAW

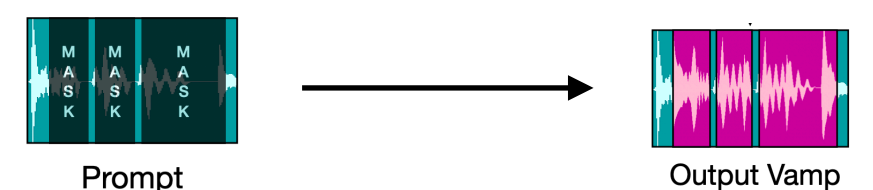


available as a VST3 plugin via the HARP plugin
(<https://github.com/audacitorch/pyharp>)

Takeaways!

VampNet is a music generation approach that leverages masked acoustic token modeling techniques.

VampNet enables applications like **musical variations** and **generative compression**, making it an excellent music production tool for interactive Human-AI music co-creation.



VampNet is open source, with code, weights, and a generative looper interface available online.

Listen to VampNet audio examples, co-created loops, fine-tuned models, and more!



Try VampNet on HuggingFace Spaces:
(<https://huggingface.co/spaces/descript/vampnet>)

code/weights:
(<https://github.com/hugofloresgarcia/vampnet>)