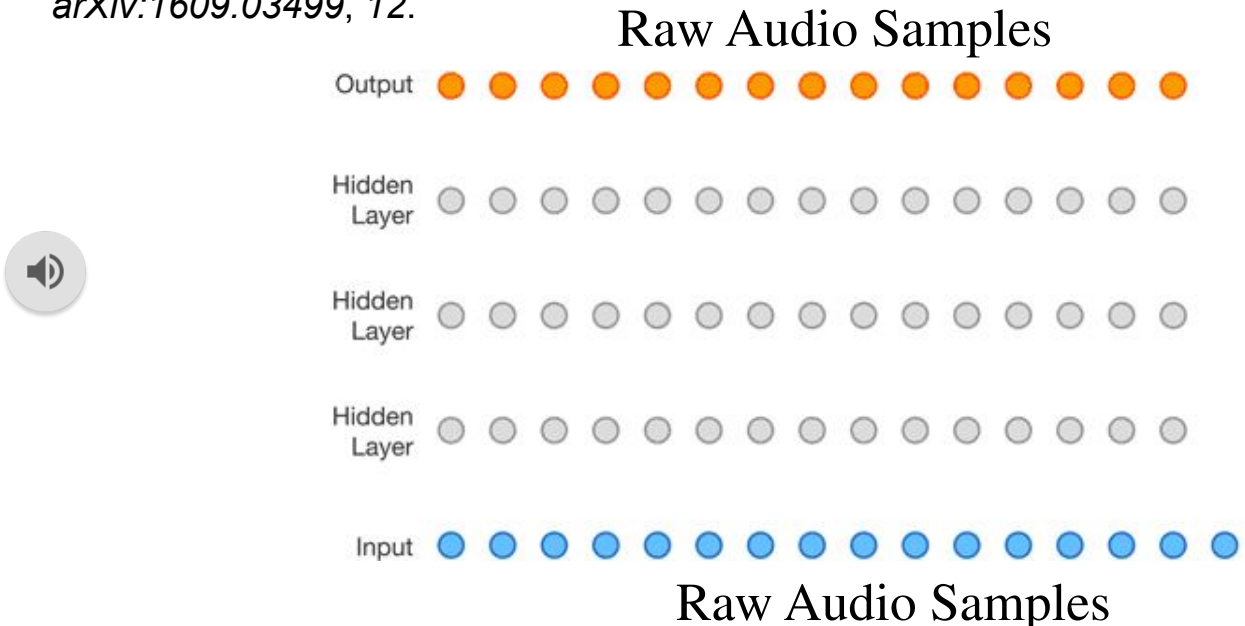
WHAT IS COMING IN
PART II
GENERATIVE AUDIO

Lonce Wyse

Shifting gears from Symbolic Music to General Audio

- Wavenet
  - Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (**2016**). **Wavenet**: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, *12*.

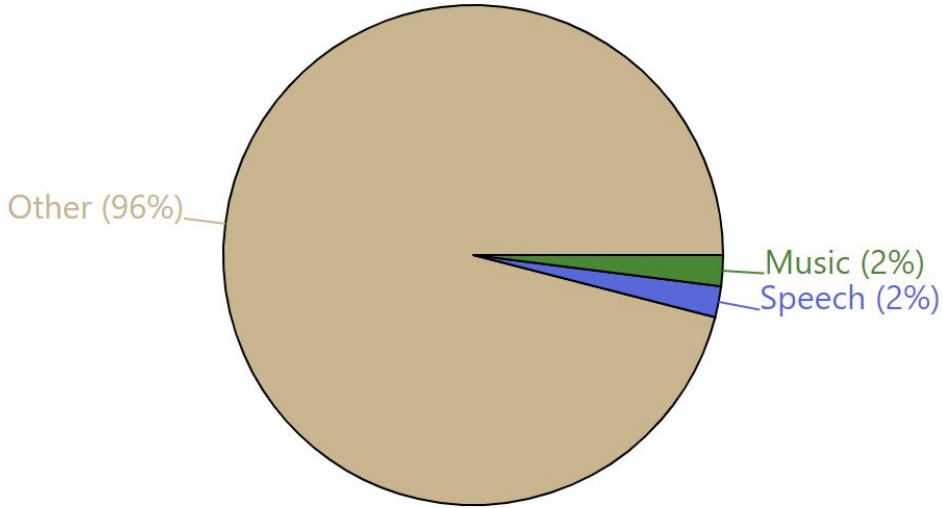


Shifting gears from Symbolic Music to General Audio

- Some of the issues/topics that arise:
  - Other “subsymbolic” audio representations
    - frame-based
    - codecs
  - Conditioning
  - Managing long time dependencies
  - Latent spaces for creative interaction
  - Managing computational overhead
  - Playability
    - Offline vs online / Composition vs performance / Text vs RT Interaction

Why raw audio?

1. What we hear
2. Space of Sound



For: Sound effects, game sound, “experimental” music

Architectures

- Convolutional (e.g. Wavenet)
- GAN (e.g. Sound Model Factory)
- VAE (e.g. Rave)
- DDSP (e.g. DDSP)
- RNNs (not dead yet!)
- Transformers (Vampnet)
  - (Building on Part 1)

Logistics

- Course style
  - Part II: Feb 2rd – March 11th (5 weeks)
  - Research paper orientation (home work)
  - Unflipped classroom
    - Mondays: in-class lecture & discussion,
    - Wednesdays: Hands-on

Feb 2 & 4 (week 5)
Feb 9 & 11 (week 6)
Feb 16 & 17 (Part I, Symbolic)
Feb 23 & 25 (week 8)
Mar 2 & 4 (week 9)
Mar 9 & 11 (week 10)

- Final Project
  - Generative audio + write up
  - Due Friday, March 20th midnight.
    - (no extensions – already as late as we can push it!)

What does AI bring to CMC?

- Productivity
- Anyone can do it without training
- New tools for artists

Yeah, yeah ... is that all?

Github resource

- Course website
- contact
  - DM on our Slack channel (#smc24-musicgen, @lonce wyse), or email: lonce.wyse@upf.edu
  - **Office: Tangier 55-318**