

# Generating Multitrack Music using Deep Learning

**Hao-Wen Dong**

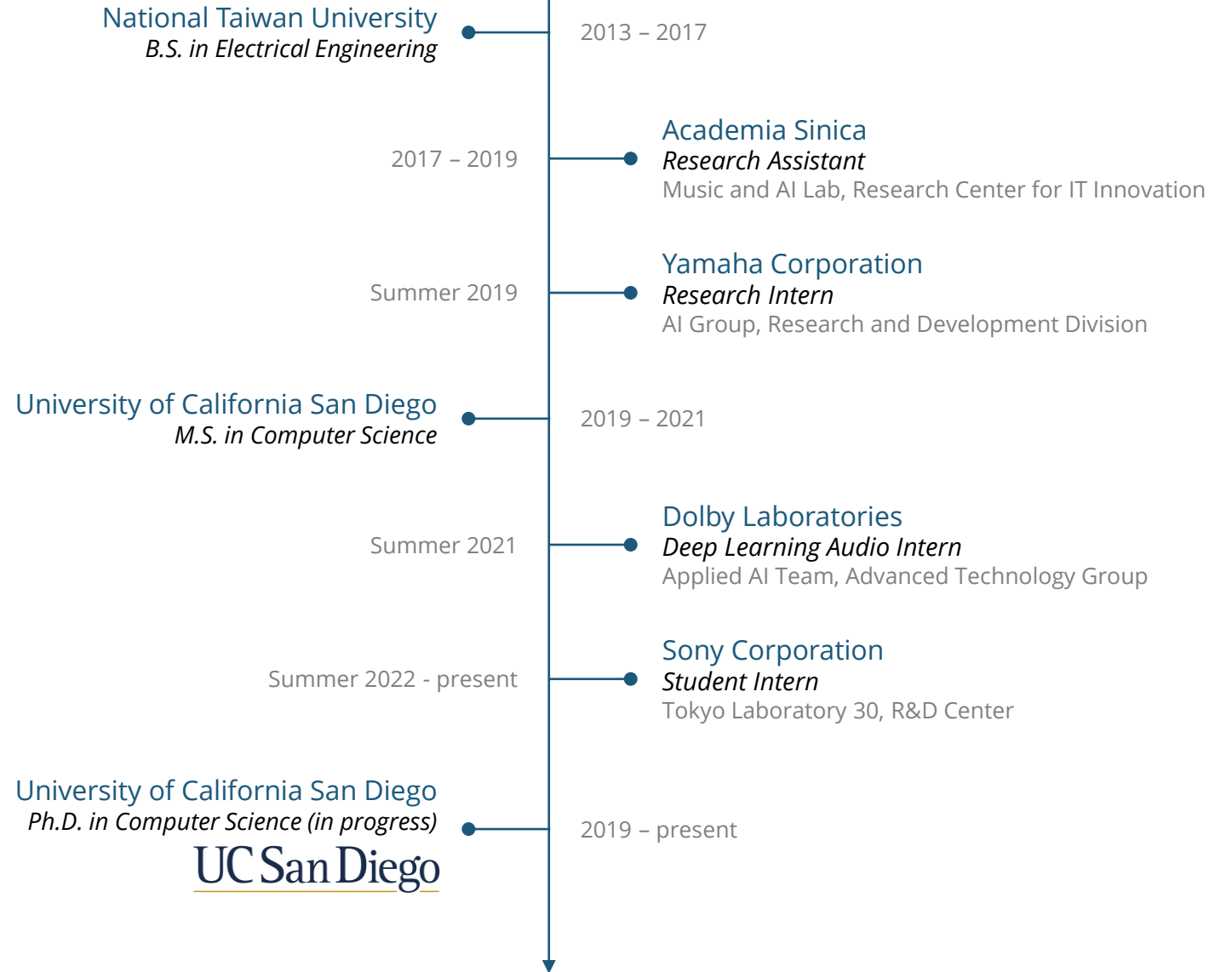
University of California San Diego

UC San Diego

# About me



Hi, I'm Herman.  
I do **Music x AI** research.  
I love music and movies!



# Outlines

- **MuseGAN** for multitrack music generation (AAAI 2018)
- **Arranger** for automatic instrumentation (ISMIR 2021)
- **Multitrack Music Transformer** for multitrack music generation (submitted to ISMIR 2022)

# MuseGAN

Generating multitrack music using convolutional GANs  
(AAAI 2018)



Wen-Yi Hsiao



Li-Chia Yang

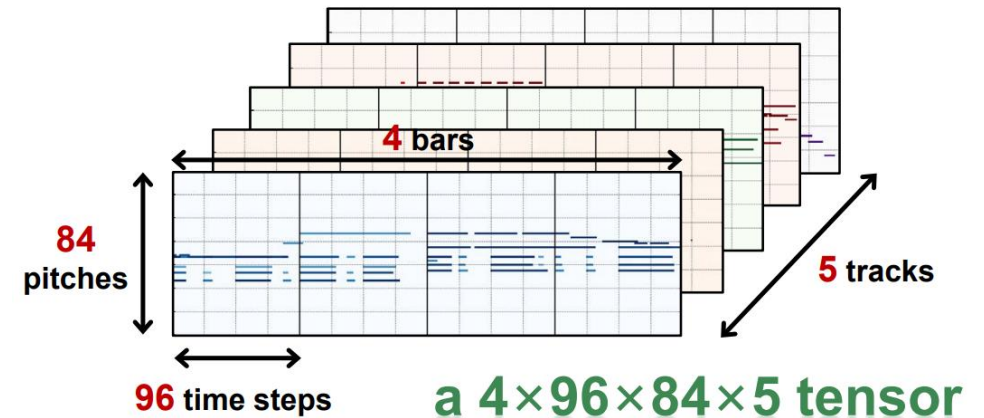


Yi-Hsuan Yang

# Overview

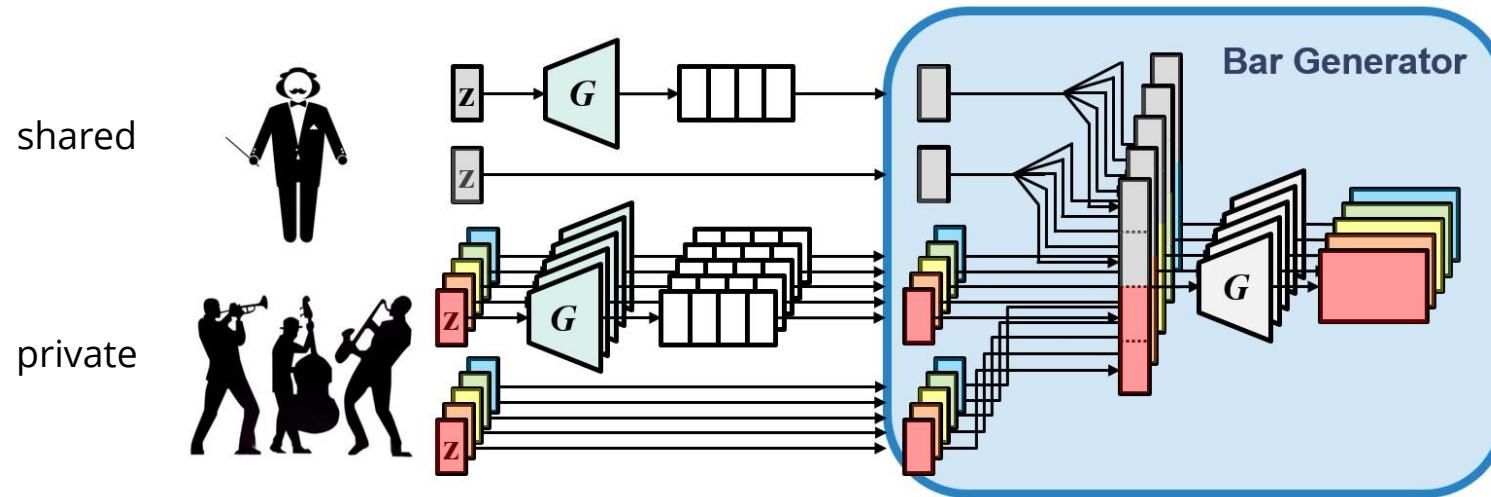
Generate pop music

- of five polyphonic tracks
- in the piano-roll format
- using convolutional GANs (generative adversarial networks)
- on the Lakh MIDI Dataset



# Model

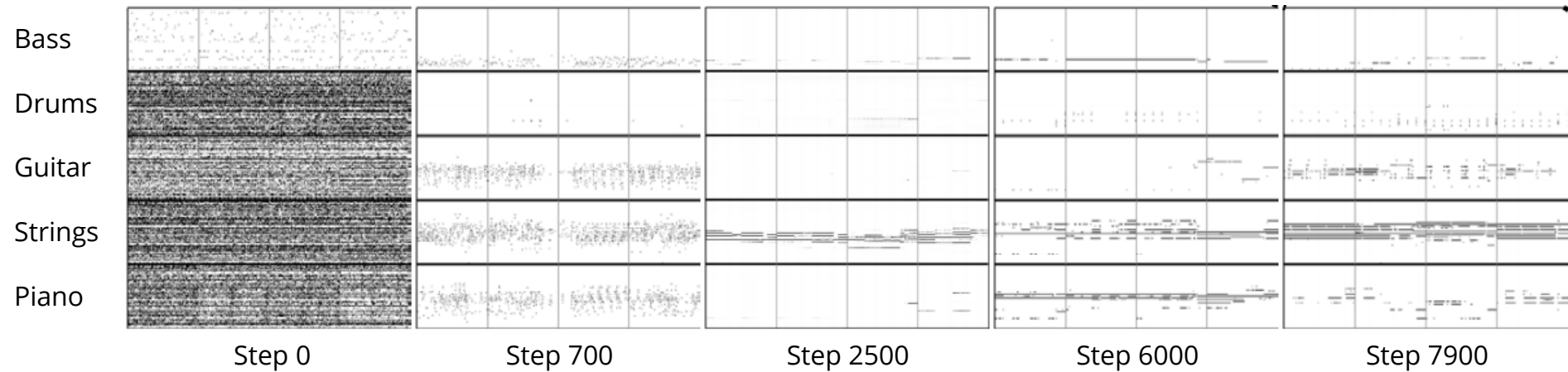
- Each track takes a **shared** and a **private** random vectors as inputs
- Offer better **controllability** than one single random vector input



# Demo



Unconditional generation samples



Training progress

# Summary

- Proposed the **first deep learning model** for generating music consisting of multiple polyphonic tracks
- Proposed the shared and private latent variables to **enhance the controllability**
- Showed that the proposed model can **learn basic musical concepts**



# Arranger

Approaching automatic instrumentation by learning to separate parts  
(ISMIR 2021)



Chris Donahue



Taylor Berg-Kirkpatrick

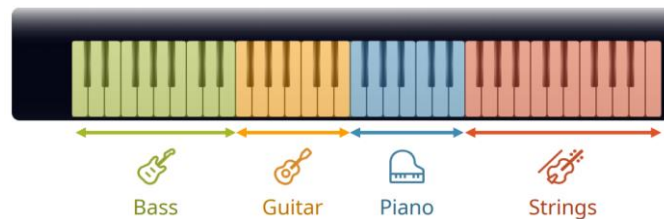


Julian McAuley

# Overview

Dynamically assign instruments to notes in solo music

- by learning to separate parts from a mixture
- using LSTMs and transformers
- on four diverse datasets (Bach chorales, string quartets, game music, pop music)



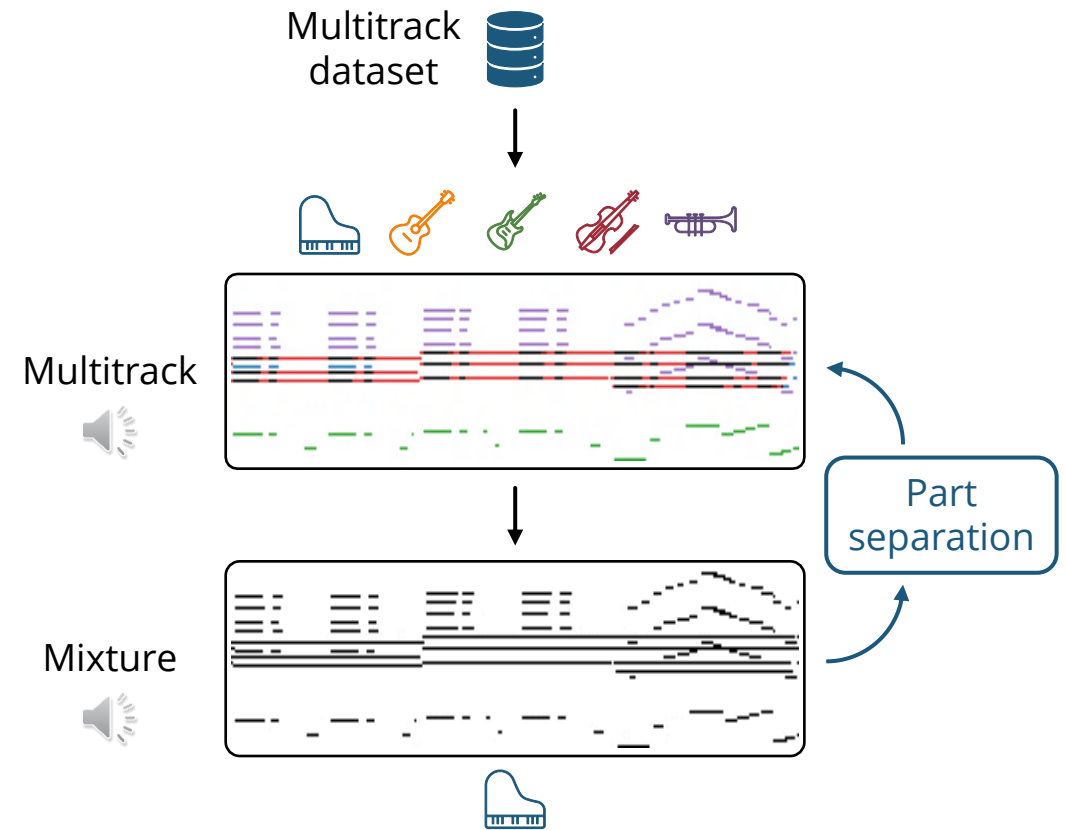
Intelligent keyboard



Assistive composing tools

# Pipeline

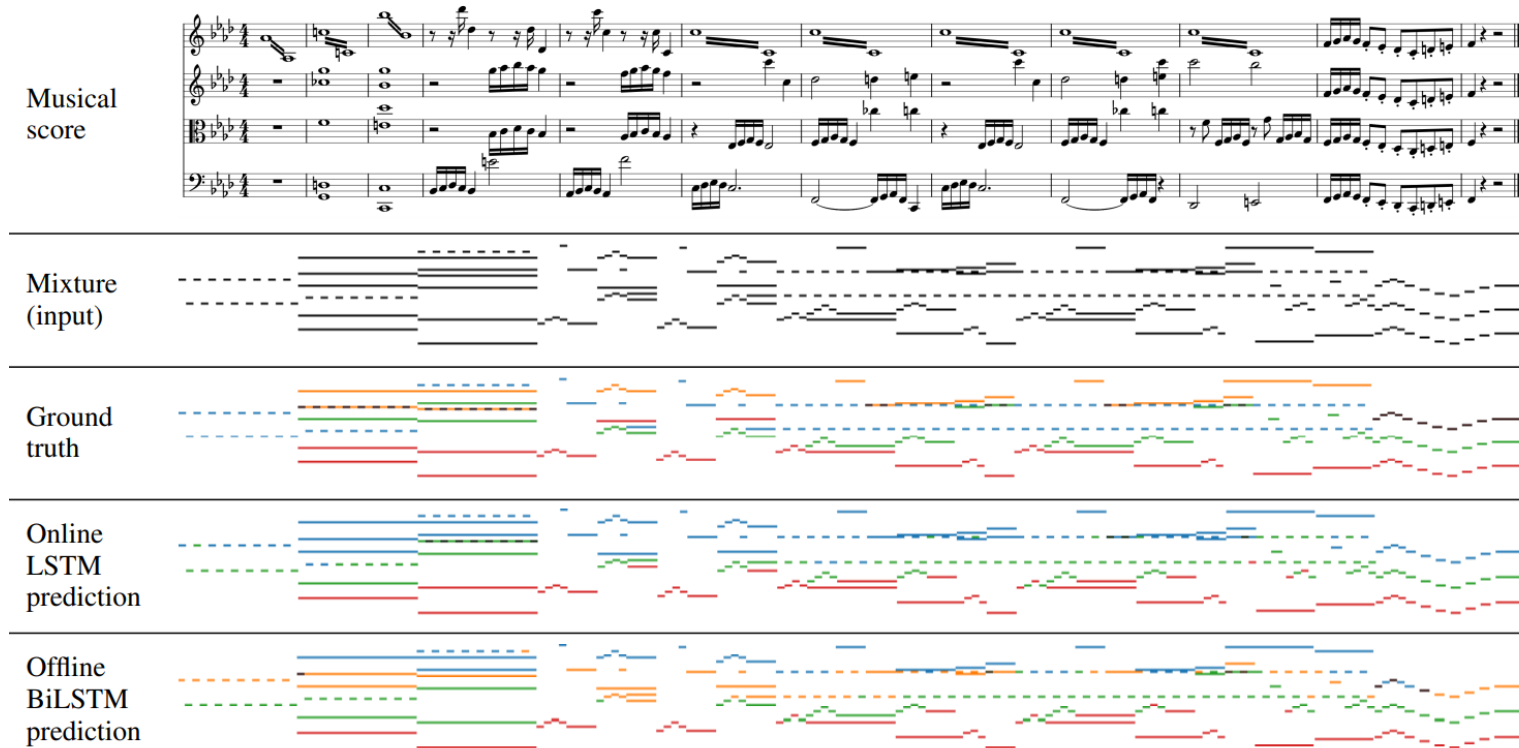
- Downmix multitracks into single-track mixtures
- Train the model to predict the part label for each note in a mixture
- Treat input from a keyboard player as a downmixed mixture and separate out the relevant parts



# A challenging example

## Beethoven's String Quartet No. 11 in F minor

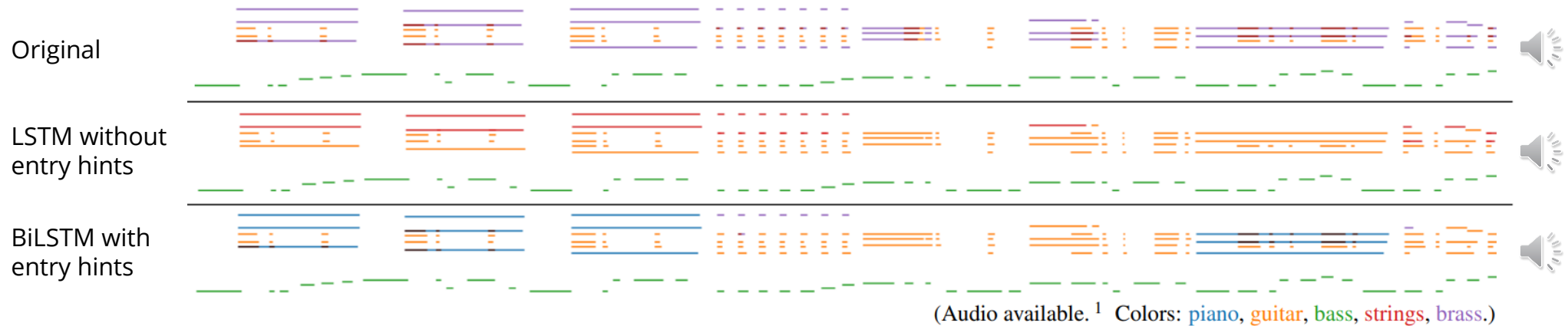
(op. 95, movement 1, measures 72-83)



(Audio available. <sup>1</sup> Colors: first violin, second violin, viola, cello.)

# Demo

- The proposed models can produce alternative convincing instrumentations for an existing arrangement



# Summary

- Approached automatic instrumentation by learning to separate parts
- Showed that our proposed models outperform various baselines

# Multitrack Music Transformer

Generating multitrack music using transformers  
(submitted to ISMIR 2022)



Ke Chen



Shlomo Dubnov



Julian McAuley

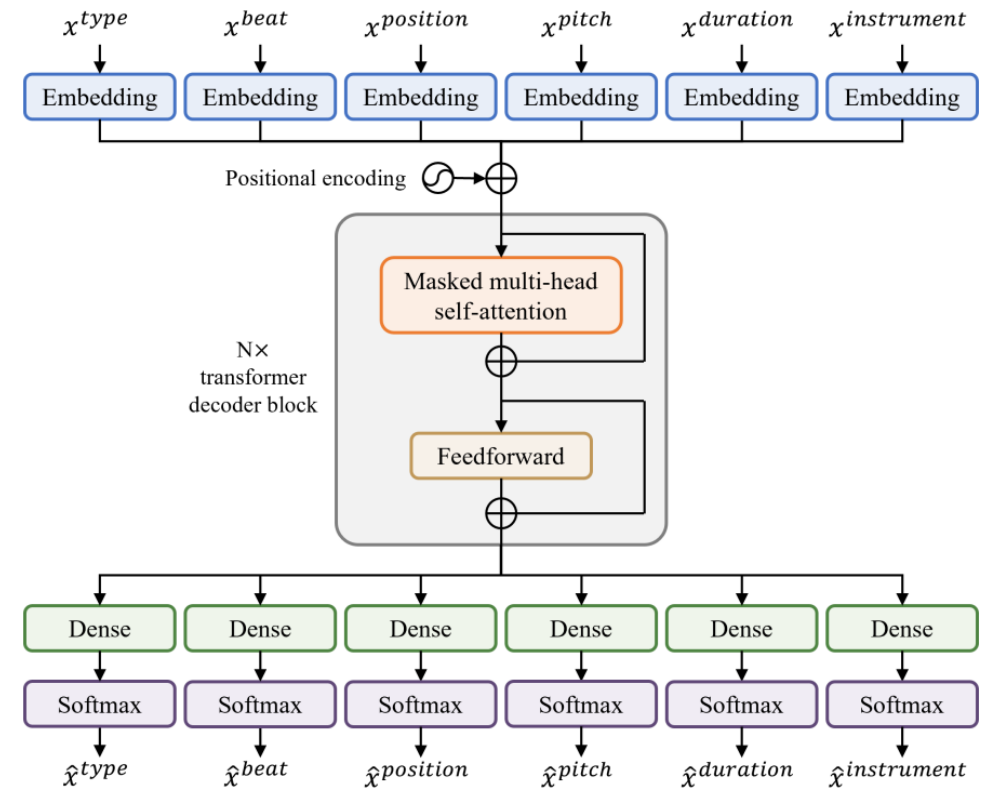


Taylor Berg-Kirkpatrick

# Overview

## Generate music

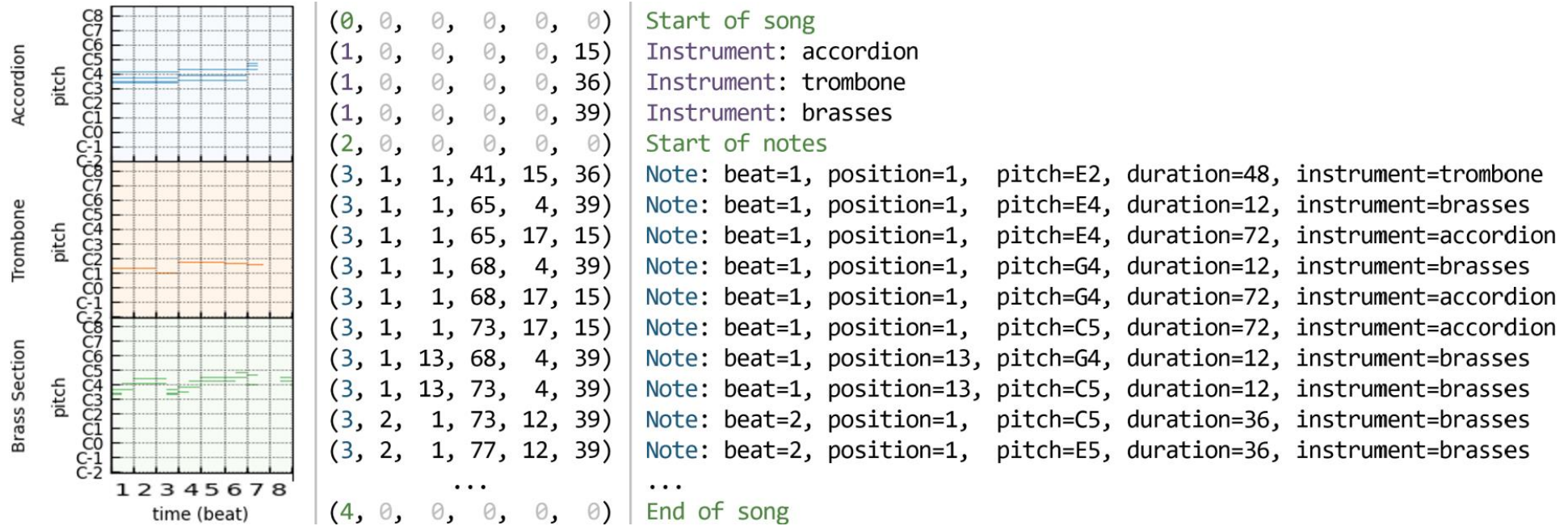
- of diverse instruments
- with a multi-dimensional transformer
- using a new compact representation
- on **pop** and **orchestral** music datasets





# Representation

- Represent 2-4x longer music within the same sequence length (compared to existing representations)



# Example results



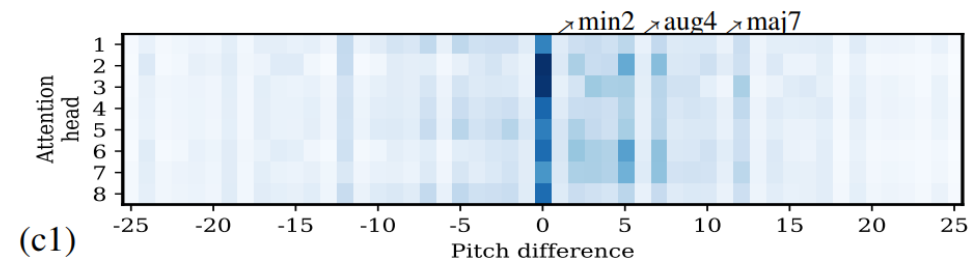
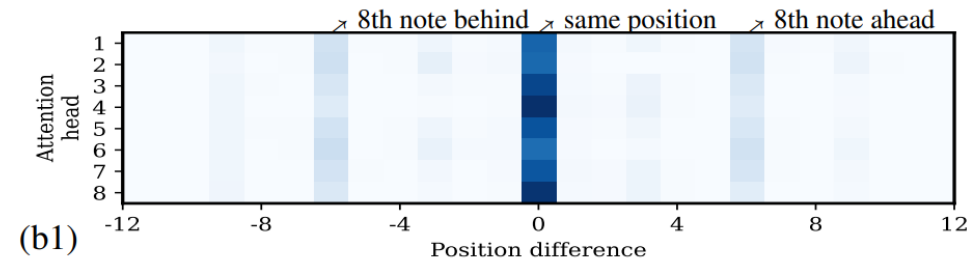
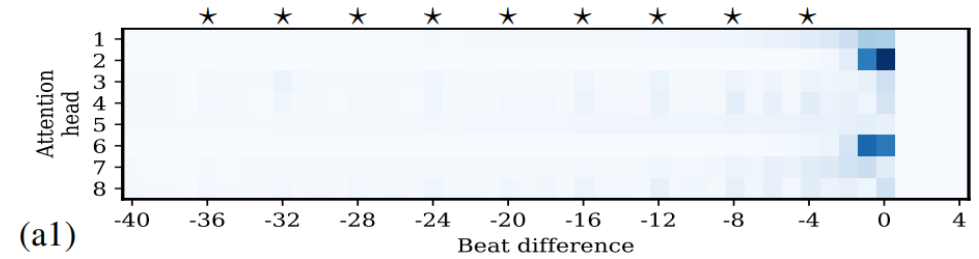
Unconditional generation 1



Unconditional generation 2



4-beat continuation



Attention visualization

# Summary

- Proposed a new representation that can **represent 2-4x longer multitrack music** within the same sequence length (compared to existing representations)
- Showed that the proposed model can **achieve competitive quality** against two baseline models (of similar sizes)
- Showed that the model can **generate 2-3x more notes** in the same inference time (compared to the two baseline models)

# Acknowledgment



Julian McAuley



Taylor Berg-Kirkpatrick



Shlomo Dubnov



Yi-Hsuan Yang



Ke Chen



Chris Donahue



Wen-Yi Hsiao

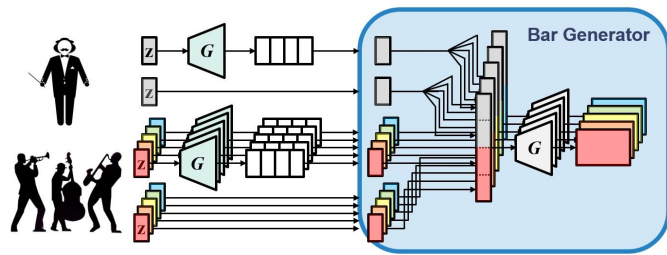


Li-Chia Yang

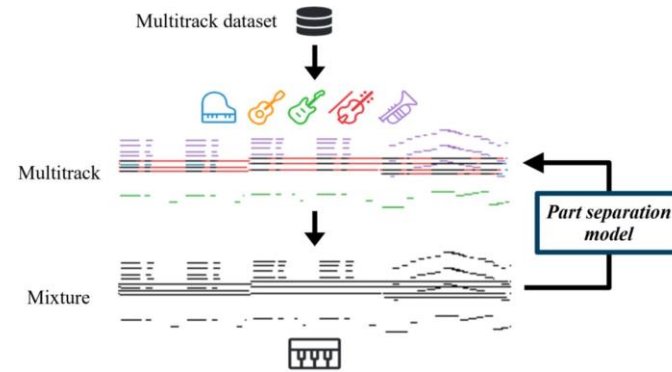
I would like to thank **J. Yang and Family Foundation** and **Taiwan Ministry of Education** for supporting my PhD study.

# Thank you!

## MuseGAN



## Arranger



## Multitrack Music Transformer

