CSE291G Project Proposal

Fanbo Xiang  Yuguang Lin

January 2019

Team Information

The team is composed of 2 students, Fanbo Xiang and Yuguang Lin.

Goals of Project

For this project, we plan to explore the topic "Symbolic Music Generation". Our general goal is to build a generative model for symbolic(sheet) music. Our goal to reach is to come up with new architectures or implement areas of improvement to such generative model to make it more powerful in generating long-term structures as well as details.

Related work

In the last few years, Google AI Magenta team has developed several generations of symbolic music generators. They started with MusicVAE[1], where they used a Variational Autoencoder with LSTM encoders and decoders to construct a latent space for music phrases. They further extended the model to multi-track music and general midi representations in [2]. They have recently used relative self-attention technique to produce long-term repetitive music structures[3]. On the other hand, Dong et al. explored CNN and GAN based music generators [4].

The MusicVAE method is reported to give a meaningful latent space of music phrases allowing much artistic control, but it is not trained to generate future music phrases, so it cannot generate long-term music structures. The self-attention model can generate long music pieces but allows little detailed control. We plan to combine the methods to create models with benefits of the two.

Training data

The training data we plan to use is the Lakh MIDI Dataset [4][5]. This data set contains about 100,000 pieces of midi music with 4/4 time signature and is used extensively in previous researches. This dataset also has a cleaned subset,
LPD.5, which classifies all instruments into 5 categories and gives the piano roll representation of each instrument category.

**Initial thoughts**

Our initial thought to address our goal is to apply the transformer model to the music latent space learned by VAEs. Currently attention was not combined with latent space model in these papers, and we think this might be worth trying.

**References**


