Statement of Purpose

Hao-Wen Dong
hwdong@ucsd.edu

My motivation for pursuing a Ph.D. in Computer Science and Engineering comes from my interests in machine learning and passions in doing research. I have always been fascinated by the idea of building a machine that can learn and think just like a human. While most successfully deployed machine learning models today are based on supervised training, human brains, however, can easily extract a lot more information and figure out the underlying structures and patterns even from unlabelled data. I am thus interested in machine learning models that require less supervisions and mimic human thinking to some extent.

I have also always found it enjoyable to do research on topics of my interest. I began my very first research project early when I was in high school, and have completed four research projects during high school, where two of them are co-advised by my high school teachers and university professors/researchers. In college, I worked with Prof. Hung Yu-Wei to build a game theoretic model for resource allocation and pricing in wireless mobile networks. After college, I worked as a research assistant in the Music and AI Lab directed by Dr. Yi-Hsuan Yang at Academia Sinica for nearly two years, where my research was in music information retrieval with a focus on music generation. Last summer, I did a research internship at Yamaha, where I worked on music synthesizers using deep learning approaches. Throughout these years of research experience, I always found myself most engaged when tackling unsolved challenging research problems.

With my passion and experiences on research, I am now determined to pursue a career in academia, and I believe doing a Ph.D. is the best next step for my research career. My research interests lie at the intersection of artificial intelligence and music. In particular, I am most interested in generative models for music, and I would love to pursue the following questions for my graduate studies.

- *How to learn compact yet efficient representations of music using deep neural networks?*
- *How to learn/generate temporal and instrumental structures of music in an unsupervised way?*
- *More generally, how to build an intelligent system that can compose music just like how human do?*

Specifically, I plan to study music generation along the following research directions.

- **Modeling multitrack interdependencies**—Most prior arts on automatic music composition focused on generating melodies, lead sheets (melodies and chords) and four-part chorales. However, music today usually consists of multiple instruments or tracks. Hence, I am interested in how we can design special network architectures to model multitrack interdependencies. As a first attempt toward this ambitious goal, I developed a novel multi-track sequential generative adversarial network for multitrack music generation. This model, which we call MuseGAN, represents the first model in the literature that can generate polyphonic multitrack music from scratch. With the MuseGAN model and its improved version, I published two first-authored papers, one at AAAI [1] and one at ISMIR [2] .

For future research, I would like to integrate the MuseGAN model with my another work on adopting binary neurons in generative adversarial networks [3]. Specifically, I plan to adopt binary neurons as gate units in network architectures so that the neural networks can dynamically activate/deactivate certain instruments in the middle of the generated music. Moreover, I am also interested in developing more efficient model for handling multitrack interdependencies. One possible approach has been proposed in Prof. Julian McAuley and Prof. Garrison Cottrell's paper [4].

- **Learning efficient representations of music unsupervisedly**—In [1, 2], we are only able to learn the mapping from the latent space to the data space using generative adversarial networks. However, I am also interested in learning latent representations with two-way mappings (between latent space and data space) for music using other deep generative models such as variational autoencoders.

  Moreover, music has long-term temporal dependencies and high-level structures such as paragraphs, repetitions and variations. I would also like to investigate the possibilities of generating music with structures using deep neural networks. One possible approach that utilizes self-similarity matrices has been investigated in Prof. Taylor Berg-Kirkpatrick's paper [5].

A complete music generation system is even more challenging for it requires modeling multitrack interdependencies and temporal structures both at the same time. As a result, success in natural language generation and melody generation might not be readily applicable, and that's why automatic music composition is interesting and worth investigating. Hence, I would love to keep working on generative models for music for my graduate studies, and I hope I can in the end push the research frontiers one step toward the ultimate goal of creating a sophisticated AI composer.

The Ph.D. program in Computer Science and Engineering at UC San Diego is especially attractive to me for the outstanding research projects carried out by its faculty members. In particular, I would like to work with **Prof. Julian McAuley** and **Prof. Taylor Berg-Kirkpatrick** to extend their works [4,5] and study generative models for music. Moreover, your curriculum covers a wide range of research areas that I would like to explore, which I believe will also be helpful for my research. For the above reasons, I am confident that I am a good fit for your program, and that it is the right place for me to pursue my ambitious dream.

[1] **Hao-Wen Dong**, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. MuseGAN: Multi-Track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, 2018.

[2] **Hao-Wen Dong** and Yi-Hsuan Yang. Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation. In *Proceedings of the 19th International Society for Music Information Retrieval Conference*, 2018.

[3] **Hao-Wen Dong** and Yi-Hsuan Yang. Training Generative Adversarial Networks with Binary Neurons by End-to-end Backpropagation. *arXiv preprint arXiv:1810.04714*, 2018.

[4] Chris Donahue, Huanru Henry Mao, Yiting Ethan Li, Garrison W. Cottrell, and Julian McAuley. Lakhnes: Improving multi-instrumental music generation with cross-domain pre-training. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*, 2019.

[5] Harsh Jhamtani and Taylor Berg-Kirkpatrick. Modeling self-repetition in music generation using generative adversarial networks. In *ICML Workshop on Machine Learning for Music Discovery*, 2019.