Initial Plan

Musical Harmonization via Machine Learning

Module: CM2303 - 40 Credits

Author: Brendan Rhys Lamb (C1610560)

Supervisor: Dave Marshall

Moderator: Matthias Treder

Project Description and Motivation

The idea of music written by a machine may have seemed novel in the past, but as of today many individual pieces of music have been written by artificial intelligences. However, oftentimes this music seems to lack many features of music written by humans. My aim is to investigate how we are able to improve machine-written music by creating a neural network that is able to create harmony for a given melody that shows consideration for structure, the greater piece of music as a whole, and potentially even concepts such as motif and theme.

Many tools to help musical composers in harmonization and chord choice already exist today, without the use of machine learning. These tools, such as Plugin Boutique's *Scaler*[1], analyze a given set of notes and suggest a harmonically "correct" chord to accompany the given melody. However, this basic conceptualization of harmony leads to some problems. If we defined a mere musical bar as our set of notes, or input, we would be failing to take into account the information from previous and future bars. As Leonard Bernstein notes in his 1973 series of lectures *The Unanswered Question*: "It is in the nature of music to be ongoing"[1]. A piece of music does not exist as these isolated fragments, but as a greater whole with overarching structure and form.

Incidentally, Bernstein's lectures serve as a primary motivator for this project, and the use of Deep-Learning technology within it. Bernstein puts forward a novel idea, that music could be conceptualized linguistically via syntax and semantics[1][2]. As techniques already exist for language processing using neural networks, these approaches might be applied to musical processing using neural networks. Hopefully, the use of such a model to conceptualize music will allow for a neural network to process the basic elements that make up a larger musical structure.

Other projects exist that seek to investigate improvements into machine-created music. *Magenta* is a project by Google that seeks to investigate how machine learning can be used in the creation of art and music. Importantly, many recent publications investigate the semantics of music via the use of neural networks, providing validation to Berstein's ideas.[8] In particular one publication making use of *Magenta* took a look into generating musical accompaniment with long-term structure[4], such as repeated passages and recurring motifs. This was accomplished by adapting approaches traditionally used for language processing. Google's previous success in this area provides a solid proof of concept, and provides a solid foundation on which to further iterate.[3][4]

The eventual goal of this project to create an affectionately named 'JazzBot' which could create from a given musical melody, or idea, a complex accompaniment using interesting and artistic techniques.

---

[1] https://www.pluginboutique.com/product/3-Studio-Tools/72-Utility/3933-Scaler

Technology, Objectives and Challenges

In terms of technology and approach required for this project, I believe the use of an RNN will be required due to the fact that the data we will be using, MIDI, is sequential information. Long Short-Term Memory networks seem like a promising start, as they have been shown to serve musical application.[6][5] In one publication, LSTM neural networks have been shown to successfully learn the structure of blues music[5]. This contrasts with other architectures of RNN, which are unable to create such structured compositions.

There are three core objectives that I would like this project to accomplish over the duration of the semester. These core objectives have been selected so that they are completed sequentially, providing a foundation to build from at each milestone.

Core Objective 1: Basic Chord Prediction

The first core objective is that of a basic neural network, and for that neural network to be capable of basic chord prediction. To achieve this milestone, the system should be able to take a melody as an input, and output harmonically functional chords to accompany the melody as an output. Via training, the neural network should be able to achieve a level of functionality that will allow it to suggest chords per bar.

Core Objective 2: Looking at Structure

The second core objective will be achieved when the neural network is able to analyse the overarching structure of the music. In order to do this, the neural network will need to be modified to become and LSTM neural network. Via an approach such as word2vec, which has shown promise in analysing musical data[7][8], the neural network should possess the tools it needs to analyse the music. A success on this milestone will result in a neural network that uses information from other bars in the music to make choices about chords. Use of cadences in correct positions would be a sign that this has been achieved.

Core Objective 3: Improving Structure

The third core objective will be achieved when the neural network makes use of motif in its choice of chords. The neural network will be required to search the given melody for patterns. These patterns represent a motif in the melody. If a motif occurs once at the start of a passage, and then again at the end of the passage with a slight alteration, the network's choice of harmony should reflect this.

Further Objectives

> Once these objectives are complete, I would like to experiment with other avenues in which the network could be augmented. Once core objective 3 is complete, more objectives may be added down the line, but these depend on the implementation and will be touched on in the final report.

There are a number of challenges associated with this project. As I have not worked with neural networks or machine learning before, I imagine the amount of additional learning, and practical experience I need to acquire could increase my workload. In the following section, my work plan will address this. However, there are a number of more practical problems that must be accounted for before this project is completed.

A notable challenge that will need to be dealt with right away will be that of acquiring a suitable dataset. I have located two datasets, that being a repository of Johann Sebastian Bach's chorales[2], and another being a repository of solo jazz piano pieces[3]. The use of two separate datasets, as well as a combination of the two should hopefully give some breadth of knowledge as to the capability of the neural network. It is possible that the datasets may not be sufficient however, and the risk of requiring more data is something to be considered, as locating data will take time.

I believe the larger challenge during this project, however, will be the implementation of LSTM architecture. As such, I feel that this core objective is going to take the longest amount of time to complete and have reflected this in the work plan below.

[2] http://www.kunstderfuge.com/bach/chorales.htm
[3] http://bushgrafts.com/midi/

Work Plan

In order to maintain a level of productivity, I have chosen to structure my work around a weekly task-based schedule. Due to the fact that supervisor meetings will occur weekly, work will also be able to receive weekly evaluation, and re-direction can be undertaken if required.

Week 1	The work in the first week will consist of background research and the creation of this initial report. A schedule between myself and the supervisor will be established during this week. By the end of this week, I will have submitted the initial report.
Week 2	This week will consist of additional background research, and gathering practical experience with the tools I will be aiming to use for the rest of the project. Additionally, the collection of data should be completed, or at least initiated during this week. By the end of this week, I will have selected a set of tools to use, and will be ready to begin working toward Core Objective 1 in the following week.
Week 3	This week will consist of developing a neural network suitable for chord analysis. Initial versions of this neural network will need to be documented in the draft of the final report. Collection of the training data should be completed by the end of this week at the latest. By the end of this week, decent progress towards Core Objective 1 should have been made.
Week 4	Work this week will revolve around iterating on work from week 3. Ideally, basic chord prediction will have been achieved by the middle of this week, and time will permit me to analyse the performance of the neural network at this stage. Ideally By the end of this week, Core Objective 1 should be accomplished, and work on Core Objective 2 can be initiated.
Week 5	The beginning of implementing of a LSTM architecture will be undertaken during this week. I expect this to take a considerable amount of my time. Ideally a scheduled review meeting would take place during this week, to allow for review of the basic neural network. By the end of this week, decent progress towards Core Objective 2 should be accomplished.

Week 6	Implementation of a LSTM architecture should be concluded towards the end of this week. Significant tweaking and refinement will be needed to accomplish Core Objective 2 from this point, however. By the end of this week, the implementation of the LSTM architecture will be completed.
Week 7	See above. Final tweaks will be needed in order to reach decent performance of the neural network's ability to analyse musical structure. By the end of this week, Core Objective 2 will have been completed.
Week 8	This week will consist of training the neural network to spot patterns in the melody, thus satisfying Core Objective 3. If Core Objective 3 is completed by the end of this week, work on the final report should be made with the remainder of the week. By the end of this week, Core Objective 3 will be nearing completion, if not completed.
Week 9	Having completed all Core Objectives by this point, I will use this week to explore any further objectives that I may wish to pursue at this point.
EASTER	I will take easter as a break, with the exception of preparation of the final report.
Week 10	This week will consist entirely of completing the final report. Ideally, by the end of the week, the report will be 50% completed.
Week 11	As per the previous week, this week will consist of completing the remaining 50% of the final report.
Week 12	Week 12 will be my overflow week, allowing any work that may have escaped me to be caught up on.

Bibliography

[1] Bernstein, L. (1973). *The Unanswered Question: Lecture 2, "Musical Syntax"*.

[2] Bernstein, L. (1973). *The Unanswered Question: Lecture 3, "Musical Semantics"*.

[3] Roberts, A., Huang, A., Hawthorne, C., Howcroft, J., Wexler, J., Hong, L. and Dinculescu, M., 2019. Approachable music composition with machine learning at scale.

[4] Huang, C.Z.A., Vaswani, A., Uszkoreit, J., Simon, I., Hawthorne, C., Shazeer, N., Dai, A.M., Hoffman, M.D., Dinculescu, M. and Eck, D., 2018. Music transformer: Generating music with long-term structure.

[5] Eck, D. and Schmidhuber, J., 2002, August. Learning the long-term structure of the blues. In *International Conference on Artificial Neural Networks* (pp. 284-289). Springer, Berlin, Heidelberg.

[6] Skúli, S. (2017). *How to Generate Music using a LSTM Neural Network in Keras*. [online] Towards Data Science. Available at: https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5 [Accessed 29 Jan. 2020].

[7] Faria, W. (2018). *MIDI Music Data Extraction using Music21 and Word2Vec on Kaggle*. [online] Towards Data Science. Available at: https://towardsdatascience.com/midi-music-data-extraction-using-music21-and-word2vec-on-kaggle-cb383261cd4e [Accessed 25 Jan. 2020].

[8] Madjiheurem, S., Qu, L. and Walder, C., 2016, December. Chord2vec: Learning musical chord embeddings. In *Proceedings of the constructive machine learning workshop at 30th conference on neural information processing systems (NIPS2016), Barcelona, Spain*.