Real-Time Identification of Simple and Extended Musical Chords using Artificial Neural Networks

**Introduction**

Chords are collections of two or more musical notes, often played together, and are arranged in such a way that they follow the so-called “rules of harmony” (Leino, Brattico, Tervaniemi, & Vurst, 2007). These rules are recognized by humans as a response in the brain that is triggered when they are violated (Leino et al., 2007). Because of this, chords are fundamental to the harmonic integrity of any musical work.

A musical chord is commonly identified by three parameters: root note, chord type, and inversion. The root note serves as a reference point for the other notes which are played in the chord. These other notes are determined by the chord type. The inversion tells which note in the chord is played lowest. For example, a C major chord, 0th inversion has “C” as the root note, and “major” as the chord type. A major chord type includes the 1st (root), major 3rd, and 5th harmonics of the typical Western major scale, and the 0th inversion says that the root note is the bottom note. Thus, the notes of the 0th inversion of a C major chord are C (1st), E (major 3rd), and G (5th), in that order.

Artificial Neural Networks (ANNs) are computational models that use a layered structure of computational units called neurons in order to learn a certain task. Using mathematical functions, the neural network is able to train itself iteratively through a process called backpropagation until its error rate is significantly low (Nielsen, 2015; Sanderson, 2017).

**Statement of the Problem**

Humphrey, Bello, and Cho (n.d.) state that “the general music learning public places a high demand on chord-based representations of popular music” (par. 1). However, complete and accurate determination of these chords by hearing requires the use of both absolute and relative pitch, because chords utilize both an absolute reference point (root note) and a relative configuration of harmonies (chord type).

Absolute pitch is expressed when one can identify a musical note by hearing it, while relative pitch is shown when one can recognize the distances between musical notes (Zatorre, Perry, Beckett, Westbury, & Evans, 1998). It is expressed in a low percentage of the human population and acquired through either favorable genes or music training at a young age, or both (Baharloo, Service, Risch, Gitschier, & Freimer, 2000).

Complete chord identification is thus a rare skill found in those with mastery of both absolute and relative pitch, even though chords play an important role in any musical work. An algorithm that automatically identifies chords from individual notes in real time would be a first step towards addressing this problem.

**Objectives of the Study**

This study aims to develop a neural network that quickly and correctly identifies simple and extended one-root musical chords formed by playing more than two notes on a MIDI input device. Specifically, the program must identify common & extended chords and respond quickly enough to be used in live performance (Stark & Plumbley, 2009). The program must be implemented in programming languages that have MIDI input-output libraries such as *pygame* for Python and and neural network libraries such as Keras and TensorFlow to facilitate ease of coding.

**Significance of the Study**

Such application of real-time chord identification would be used in the field of music education, where a low proportion of music students have absolute pitch (Gregersen, Kowalsky, Kohn, & Marvin, 1999) despite their demand for chordal representations of music (Humphrey, Bello, & Cho, n.d., par. 1). These allow said students to learn to identify the chords they are playing more quickly and accurately, and help them develop their senses of relative and absolute pitch. They are also used in situations when musicians need to verify the chords they are playing for correctness, which usually happens when they are learning or composing a musical piece.

While automatic chord identification programs exist, they are either classical algorithm implementations that run on audio inputs in real-time (Fujishima, 1999; Stark & Plumbley, 2009), or neural network implementations that do not aim to run in real-time and do not include chords outside of major and minor triads (Perera & Kodithuwakku, 2005; Osmalskyj, Embrechts, Piérard, & Van Droogenbroeck, 2012; Zhou & Lerch, 2015). Including extended chords would allow identification of chords from more complex styles of music, such as jazz. The utilization of **neural networks** to identify **extended chords** from **MIDI signals** in **real-time** is largely unexplored and would provide useful data for future research.

**Scope and Limitation**

The study aims to create an artificial neural network that identifies both non-extended (“common”) and extended chords within a time limit of 10ms. The dataset comprises of the following chord types only: simple triads (major, minor, aug, dim), dominant extensions (7, 9, 11), major and minor extensions (M7, m7, M9, m9, M11, m11), suspended triads (sus2, sus4), major extensions with suspensions (M7sus2, M7sus4, M9sus2, M9sus4, M11sus2), dominant extensions with suspensions (7sus2, 7sus4, 9sus2, 9sus4, 11sus2), augmented and diminished extensions (aug7, dim7, ø7, aug9, dim9, aug11), and other extensions (mM7, mM9, M6, m6, M6(9), m6(9)).

MIDI commands will be used as inputs to the neural network. Audio datasets are thus not included in the scope of this study.

Research Design

**Experimental Units and details to ensure local control**

Each chord in the dataset is an experimental unit that is assigned using stratified random sampling to either the training dataset (used during ANN training) or the validation dataset (used during ANN testing).

Latency testing will be carried out using a random sample of 30 chords. The neural network will be tested on manual MIDI inputs, and its response time will be recorded for each.

**Treatments**

The sole treatment in this study is the neural network being developed. Unlike in experimental studies where the effects of treatments on a certain EU are more or less fixed, the configuration of a neural network changes over time by learning how to perform the task it is assigned (in the case of our study, it is the identification of musical chords), changing the way the neural network responds to the same input. This is normal and should result in a neural network with satisfactory accuracy.

**Allocation of treatments to EUs**

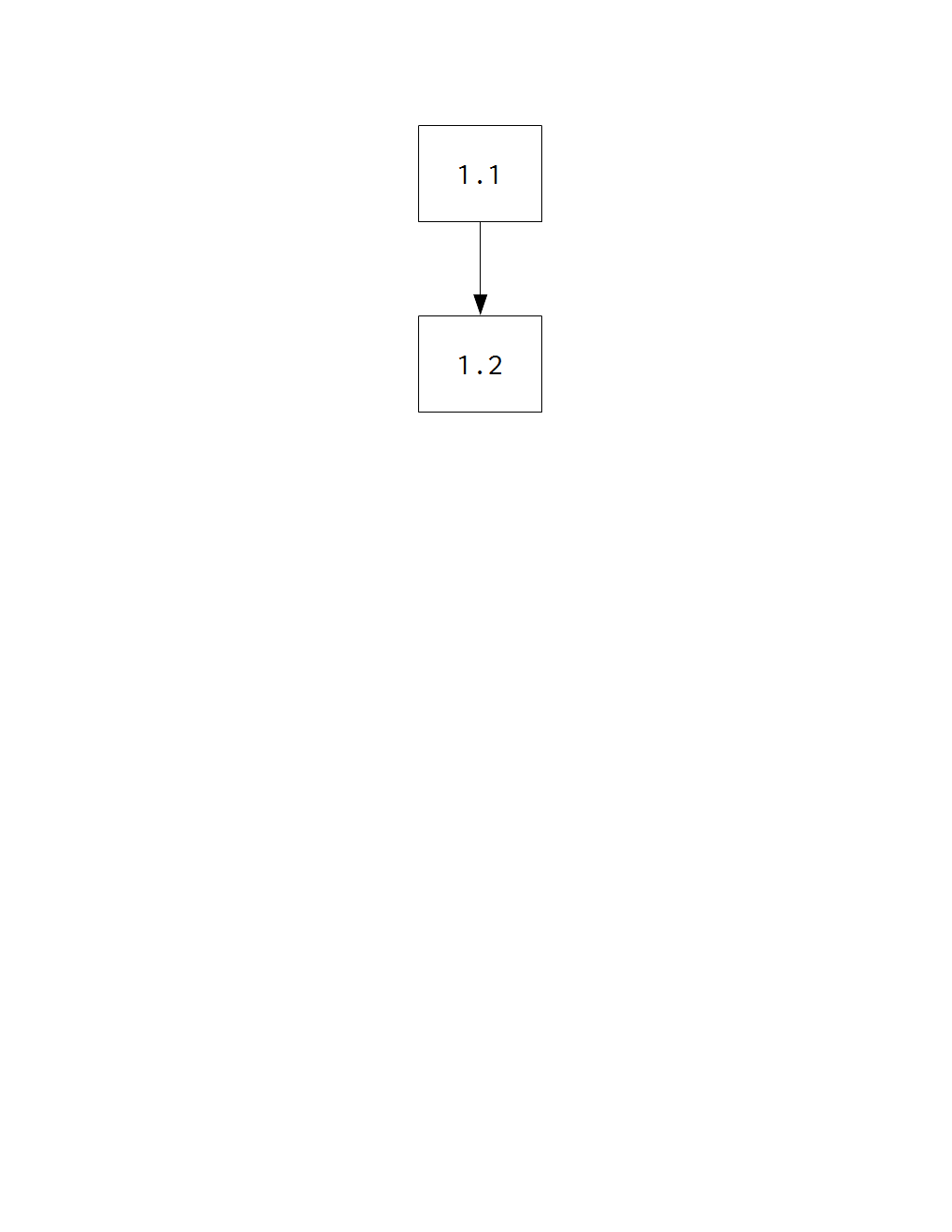
The neural network will be trained on a training dataset and tested on a separate validation dataset.

**Level 0 diagram**



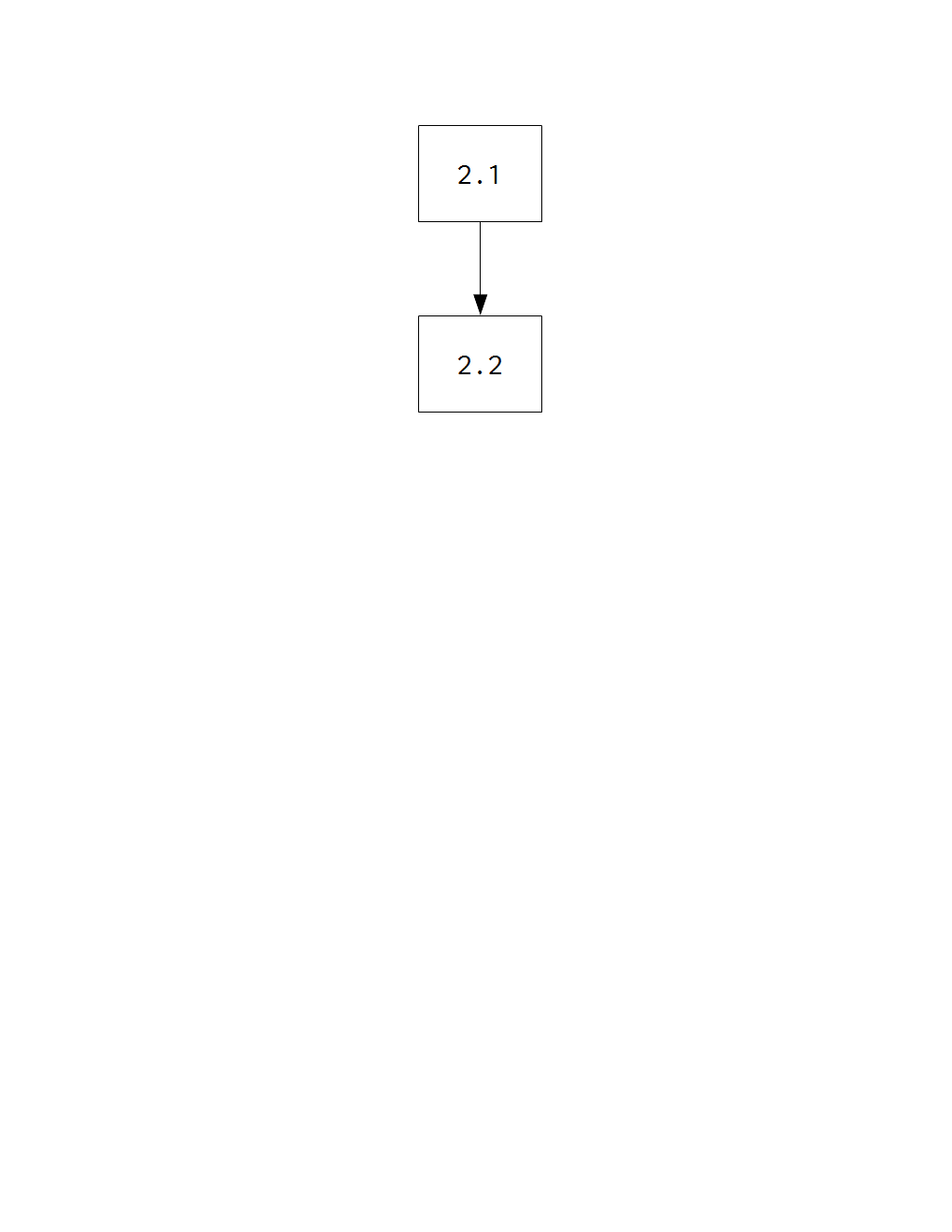
**Level 1 diagrams**

1. **ANN LAYER STRUCTURE REDESIGN**



* 1. Chord type finalization  
     *The chord types to be identified by the neural network will be determined. This step will allow the researchers to add or remove chord types from the list in the Scope and Limitation section of this paper.*
  2. Determination of proper output format  
     *The outputs of a neural network are a series of numbers. In the previous iteration of this study, it was hypothesized that an improper output format rendered the neural network unable to perform its task.*

1. **TEST CHORD DATABASE RECONSTRUCTION**

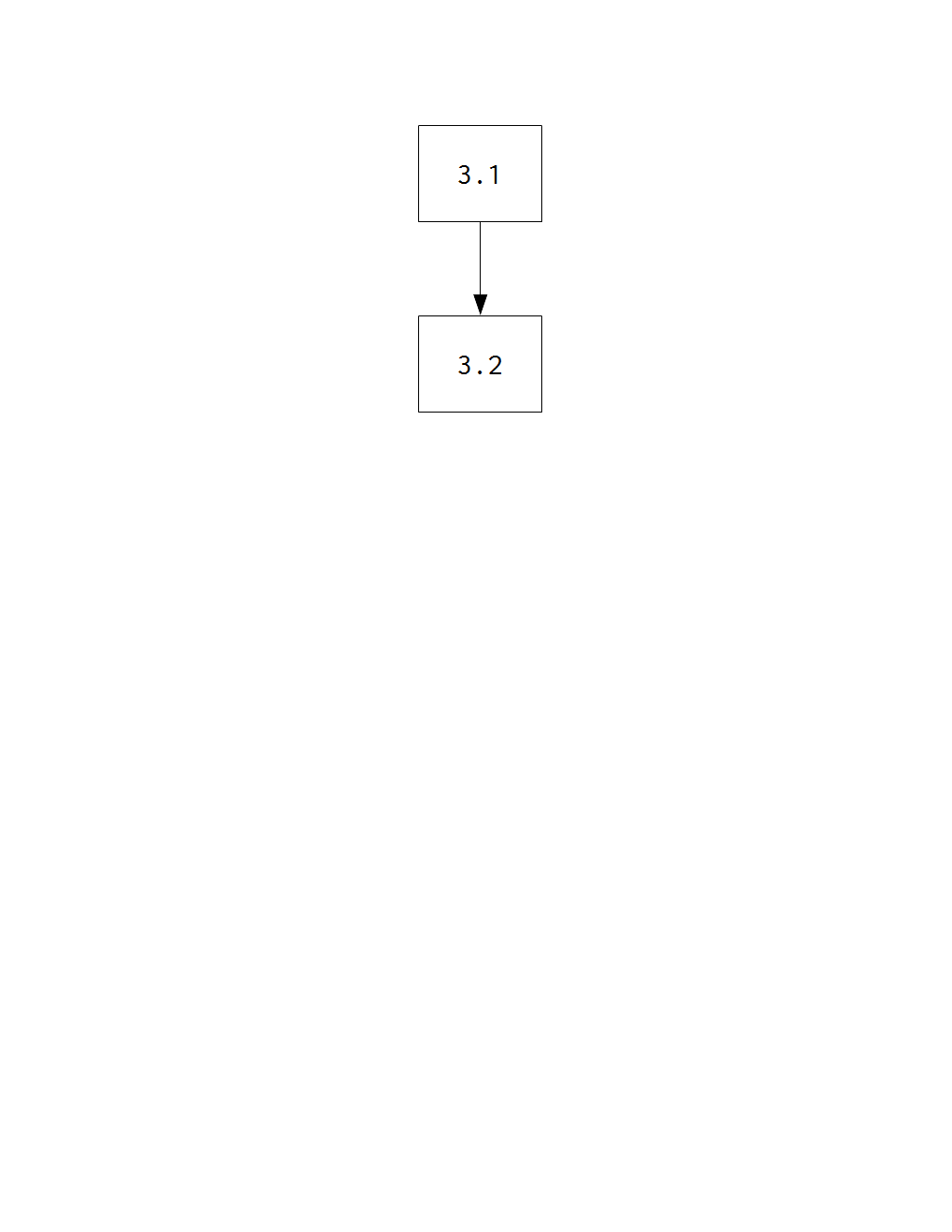


* 1. Implementation of output format revisions

*The changes made to the output format in* **1.2 Determination of proper output format** *will be implemented in the initiator program for training and testing datasets.*

* 1. Stratified random sampling of chords   
     *Two rounds of sampling will be performed. The first round consists of splitting the entire dataset into training and validation datasets, and the second round consists of picking 30 random chords for response time testing.*

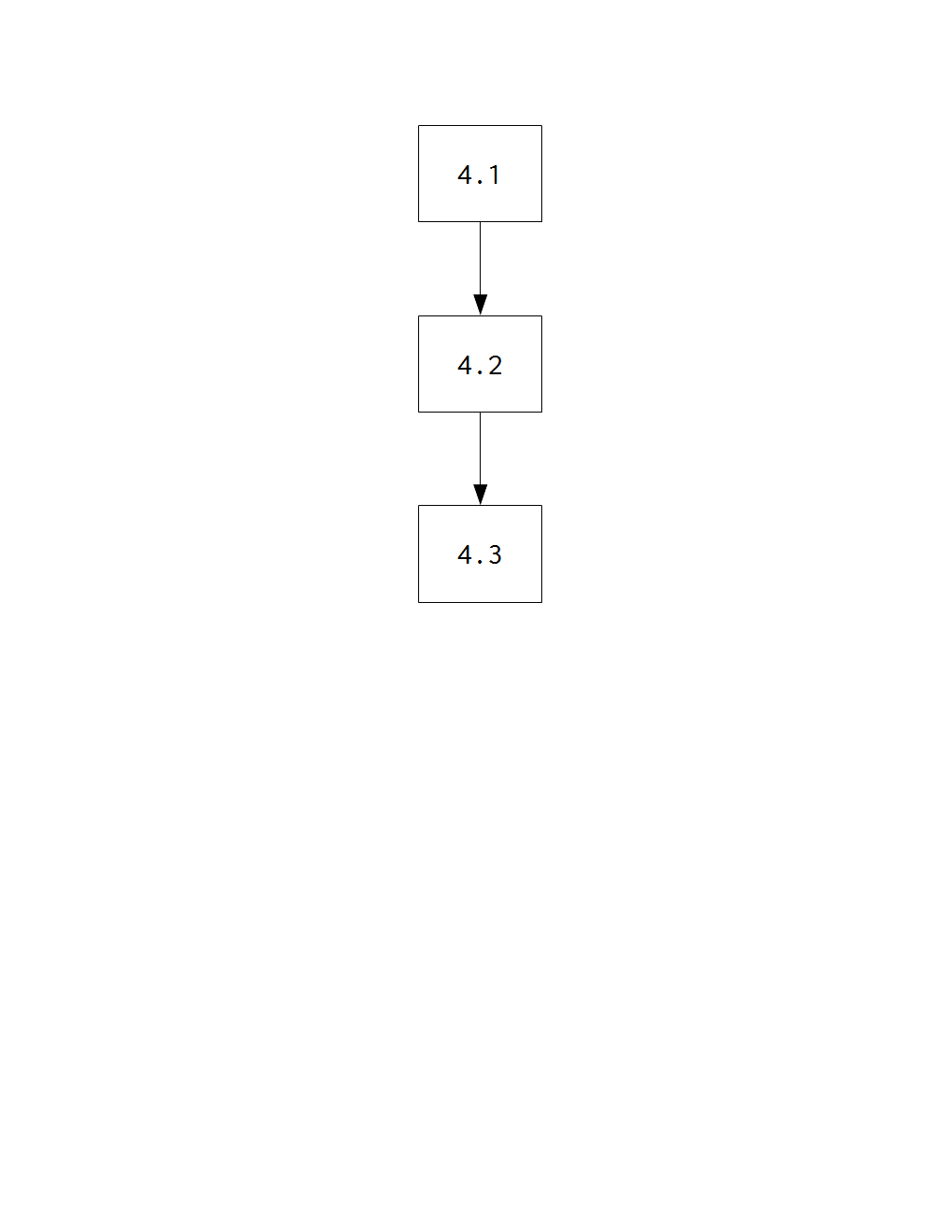
1. **ANN DEVELOPMENT**



* 1. Neural network trainer coding  
     *This involves the use of a programming language and the appropriate libraries to set up and train a neural network.*
  2. Neural network tester coding

*The same programming language will be used to facilitate MIDI input to the network for single-chord latency testing.*

1. **TESTING AND DATA COLLECTION**



* 1. Neural network training and validation data collection

*The ANN will be run for a certain number of epochs until its accuracy reaches a certain satisfactory value.*

* 1. Response time testing and data collection

*The ANN’s response time will be tested using 30 randomly selected chords.*

* 1. Response time data analysis

*A t-test for one mean will be carried out on response time data, with a benchmark of 10ms, a commonly accepted response time for live musical performance.*

**References**

Baharloo, S., Service, S., Risch, N., Gitschier, J. & Freimer, N. (2000). Familial aggregation of  
absolute pitch. *American Journal of Human Genetics, 67,* 755-758. doi:10.1086/303057.

Fujishima, T. (1999). *Realtime chord recognition of musical sound: A system using common Lisp music.* Retrieved from <http://www.music.mcgill.ca/~jason/mumt621/papers5/fujishima_1999.pdf>.

Gregersen, P., Kowalsky, E., Kohn, N., & Marvin, E. (1999). Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component [Letter to the Editor]. *American Journal of Human Genetics, 65,* 911-913. doi:10.1086/302541.

Humphrey, E., Bello, J. P., & Cho, T. (n.d.). Chord Recognition. Retrieved from <http://steinhardt.nyu.edu/marl/research/chord_recognition>.

Leino, S., Brattico, E., Tervaniemi, M., & Vurst, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research, 1142,* 169-177. doi:10.1016/j.brainres.2007.01.049.

Nielsen, M. A. (2015). *Neural networks and deep learning*. Determination Press.

Osmalskyj, J., Embrechts, J-J., Piérard, S., & Van Droogenbroeck, M. (2012, May 9). *Neural networks for musical chords recognition*. Retrieved at <http://jim.afim-asso.org/jim12/pdf/jim2012_08_p_osmalskyj.pdf>.

Perera, N., & Kodithuwakku, S. R. (2005, December 15). Music chord recognition using artificial neural networks. *1st Proceedings of the International Conference on Information and Automation*, 304-308.

Sanderson, G. [3Blue1Brown]. (2017, October 5). *But what \*is\* a Neural Network? | Chapter 1, deep learning*. Retrieved from https://www.youtube.com/watch?v=aircAruvnKk.

Stark, A. M., & Plumbley, M. D. (2009). *Real-time chord recognition for live performance* [PDF file]*.* Retrieved at <https://www.eecs.qmul.ac.uk/~markp/2009/StarkPlumbley09-icmc.pdf>.

Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences, 95*, 3172-3177. Retrieved at <http://www.pnas.org/content/95/6/3172.full>.

Zhou, X., & Lerch, A. (2015). Chord detection using deep learning. *16th International Society for Music Information Retrieval Conference*, 52-58. Retrieved at <http://ismir2015.uma.es/articles/96_Paper.pdf>.