THE UNIVERSITY OF HULL

Creating Persian-like Music Using Computational Intelligence

being a Thesis submitted for the Degree of PhD

in the University of Hull

by

Sahar Arshi

May, 2018

To my Father and Mother,

My angels on this planet,

To my sisters,

The rainbows of my life,

Abstract

Dastgāh are modal systems in traditional Persian music. Each Dastgāh consists of a group of melodies called Gushé, classified in twelve groups about a century ago (Farhat, 1990). Prior to that time, musical pieces were transferred through oral tradition. The traditional music productions revolve around the existing Dastgāh, and Gushe pieces. In this thesis computational intelligence tools are employed in creating novel Dastgāh-like music.

There are three types of creativity: combinational, exploratory, and transformational (Boden, 2000). In exploratory creativity, a conceptual space is navigated for discovering new forms. Sometimes the exploration results in transformational creativity. This is due to meaningful alterations happening on one or more of the governing dimensions of an item. In combinational creativity new links are established between items not previously connected. Boden stated that all these types of creativity can be implemented using artificial intelligence.

Various tools, and techniques are employed, in the research reported in this thesis, for generating Dastgāh-like music. Evolutionary algorithms are responsible for navigating the space of sequences of musical motives. Aesthetical critics are employed for constraining the search space in exploratory (and hopefully transformational) type of creativity. Boltzmann machine models are applied for assimilating some of the mechanisms involved in combinational creativity. The creative processes involved are guided by aesthetical critics, some of which are derived from a traditional Persian music database.

In this project, Cellular Automata (CA) are the main pattern generators employed to produce raw creative materials. Various methodologies are suggested for extracting features from CA progressions and mapping them to musical space, and input to audio synthesizers. The evaluation of the results of this thesis are assisted by publishing surveys which targeted both public and professional audiences. The generated audio samples are evaluated regarding their Dastgāh-likeness, and the level of creativity of the systems involved.

Acknowledgements

The pathway through my PhD was an enriching and growing journey. It was a path full of personal, social, and scientific discoveries. I had the privilege to work in a multidisciplinary area (both science, and music) which both are enlightening and soothing to mind and heart. My viewpoints through life have changes once for all in a positive way. It is very interesting for me how PhD can change people. I am extremely indebted to every single person who have encouraged me and supported me throughout this miraculous journey. I have never been able to do it all by myself.

First and foremost, I am extremely thankful to Dr. Darryl N. Davis, my Supervisor for his kind support, guidance, and patience. I learned a great deal from him and I am always indebted to him for the person that I have become as an individual and as a researcher. He provided me the chance to develop Liquid Brain Music software which was originally designed at the University of Hull. Through Doctor Darryl N. Davis, I was first exposed to Computational Creativity field of research.

I would like to acknowledge Prof. Yiannis Papadopoulos, and Dr. Chandra Kambhampati who provided me with their invaluable advice and feedbacks in the PhD panel meetings. I am so thankful to Prof. Yiannis Papadopoulos who gave me the opportunity for the public presentation and performance at Otley library. I am grateful to Dr. Chandra Kambhampati for setting up the gathering and group meetings of the PhD students for discussing various research papers. Dr. Chandra Kambhampati was also the internal examiner of the thesis and provided me with invaluable comments which I am thankful of.

I am also extremely thankful to Dr. Anna Jourdanous, the external examiner of this thesis who provided me with her invaluable comments to improve the quality of the thesis.

I would like to express my special gratitude to Dr. Saeed Setayeshi whom his wisdom enlightened my mind and soul. His lectures filled me with inspirations. The initial ideas of working with Santur musical instrument came to my mind during one of his machine learning lectures. He supervised me and supported me throughout my MSc. studies and encouraged me for targeting higher education. I would also like to thank Doctor Mojtaba Shamsaii Zafarghandi who provided me with invaluable feedbacks at that time.

Thanks to my beloved family Parviz Arshi, Sepideh Monfared, Yassmin Arshi, and Sadaf Arshi. They supported me through my journey and believed in me through my life. They replenished my heart with happiness and love. They fulfilled my soul with hopes. They are my first teachers on this planet and I owe them my whole existence. Parviz Arshi has provided me with the fundings of my higher education and I am always thankful to him for supporting and encouraging me through my education.

I am indebted to Dr. Fariborz Arbasi for his kind support and encouragement. He helped me throughout the application processes for higher education. I can’t possibly thank him enough for sharing his unique visions towards success subjects and *Rumi* poems (Rumi is a famous Persian poet).

I would like to deliver my special thanks to Ostad Saeed Sabet for granting me the permission to use the musical databases employed in this research. I would also like to acknowledge musicians and Santur players Hossein Khanloo, and Majid Eslami, for their invaluable discussions about Dastgāh music. Those discussions inspired me to come along with the proposal for this PhD thesis.

I am so thankful to friends, and colleagues Seyed Ali Sadeghi, Dr. Mohammad Nasr Esfehani, Lucie Leveque, Dr. Ioannis Sorokos, Athanasios Retouniotis, Luis Torrao, Dr. John Stamford, Dr. Luis Azevedo, Dr. Robert Munnoch, Dr. Sohag Kabir, Dr. John Dixon, Dr. Shylaja Kanaganapalli Ramulu, Dr. Qian Wang, Lee Odiam, Steven Balding, Tareq AlJaber, Dongfei Xue, Dr. Mohammad Al-Khaldy, Brian Peach, Seyed Kazmi and all the other researchers at the Babbage lab, and Turing lab who made the PhD journey more colourful, and memorable.

I would like to express my gratitude to Sirus Arshi, Fereydoon Naghibi, Ehsan Hosseini, Masood Hosseini, Fereydoon Aghaidoost, and Hossein Ahmadi for their kind support.

I would like to acknowledge Lynn Morrell, and Helen El-Sharkawy members of staff who always made the visits for paper works pleasant moments.

Public Outputs

❖ Some of the supporting material used in a number of sections of this thesis were published elsewhere as shown in the following Public Outputs by the author.

- *Sahar Arshi, Darryl N. Davis,* "Capturing the Dynamics of Cellular Automata, for the Generation of Synthetic Persian Music, Using Conditional Restricted Boltzmann Machines*", AI-2017, Cambridge, December 2017 in: Artificial Intelligence XXXIV, Volume 10630 of the Lecture Notes in Computer Science, Springer, 2017.*

- *Sahar Arshi, Darryl N. Davis,* "LBM to LPM: An Investigation into Computational Music", *AISB Quarterly, 2016.*

- *Sahar Arshi, Darryl N. Davis, "*Generating Synthetic Persian music*", Chapter 1 in 3rd International Conference on new Music Concepts, M.D. Ventura (editor), ABEditore s.r.1., Academia Musicale, Milano, September 2016.*

- *Sahar Arshi, Darryl N. Davis, "*Computational Framework for Aesthetical Navigation in Musical Search Space*", AISB 2016 Symposium on Computational Creativity, Sheffield, UK, April 4th 2016.*

- *Sahar Arshi, Darryl N. Davis,* "Towards a Fitness Function for Musicality using LPM*", York Doctoral Symposium, University of York, 28 October 2015.*

- *Sahar Arshi, Darryl N. Davis, "*Capturing the Dynamics of Cellular Automata, for the Generation of Synthetic Persian Music, Using Conditional Restricted Boltzmann Machines*", Poster presented on 9th Annual Departmental Conference for*

Postgraduate Research, The University of Hull, August 2017.

- *Sahar Arshi, Darryl N. Davis, "*A Computational Framework for Aesthetical Navigation in Musical Search Space*", Poster presented on 8th Annual Departmental Conference for Postgraduate Research, The University of Hull, February 2016.*

- *Sahar Arshi, Darryl N. Davis, "*Creating Novel Persian Dastgāh Musical Systems by the Assistance of Artificial Intelligence Tools*", Poster presented on 7th Annual Departmental Conference for Postgraduate Research, The University of Hull, February 2015.*

❖ Santur musical instrument public performance and presentation at Otley Arts Centre, 24th June 2015.

Table of Contents

List of Illustrations

List of Tables

Glossary of Terms

Dastgāh is a modal system in traditional Persian music (During, 2011; Farhat, 1990; Zonis, 1965). There are twelve principle groups of modes in Persian music, namely, Shur, Abou'atā, Bayāt-e-Tork, Afshāri, Dashti, Homāyoun, Bayāt-e-Esfahān, Segāh, Chāhārgāh, Māhour, Rāstpanjgāh, Navā (Farhat, 1990)**.** The Dastgāh concept determines both the title for a group of individual pieces with their characteristic modal identity and the primary mode in each group.

Radif of Persian Music is the repertoire of the ancient melodies in traditional Persian music. Radif consists of the different Dastgāhs. Radif was preserved throughout successive generations by being transfered through oral tradition and been memorized by heart. Therefore interpretations and varieties entered Radif and some melodies have possibly been lost. The Radifs collected and suggested by *'Mirzā Abdollah Farāhāni '* and *'Mirzā Hossein-Qoli'* are the oldest and most famous Radifs in traditional Persian music (During, 2011; Farhat, 1990; Zonis, 1965). The provision of the musical notation for all the pieces of Radif of *'Mirzā Hossein-Qoli'* and *'Mirzā Abdollah Farāhāni'* took place in 2001, and 1970, respectively. One of the characteristics of Radif of Persian music is that most of the pieces do not necessarily follow a tempo; the duration specification of musical motives may vary during the performance depending on different factors. Some of these factors reflect the personal moods of the performer or the singer expressions along the performance, or the audience preferences. The implication derived from the discussion is that the quality of the performance of Radif differs in any single performance.

Santur is a hammered dulcimer. A musical instrument originating in Persia whose invention is assigned to *Fārābi* the Persian scientist. Santur consists of a trapezoidal box, and the strings are stretched parallel on the upper surface of Santur. The strings are excited by means of two light hammers (*mezrāb*) held in three fingers of each hand (Sadie & Tyrrell, 2001). There are different variations of Santur in Iran and other parts of the world. Nine and eleven bridged Santurs are very common versions of Santur in Iran. In a nine-bridged Santur, there are 27 notes present, and there are a group of four strings associated with each of the notes in Santur. In fact, the high number of strings and the resonations occurring in the body of the instrument causes special acoustical characteristics of the instrument. Santur, by changing its tuning, is capable of being played in different Persian musical systems (Arshi, 2012).

Ostad is a title that is often given to the masters or professionals of Arts or Science**.**

Gushé: Each Dastgāh consists of individual melodies called Gushé, which vary in length and importance.

Bemol '♭': This symbol is a musical notation, if placed near a musical note, the pitch of the note is flattened by a semitone. This symbol and the three consecutive symbols are used in Appendix E.

Diese '#': A symbol in musical notation that raises the pitch of the note by a semitone.

Koron '𝄬': This symbol and its counterpart 'Sori' are two Eastern musical symbols, which are used to change the tuning of the musical notes according to the characteristics of Persian music. 'Koron' symbol flattens the pitch of the note by a quartertone.

Sori '𝄲': This Persian musical symbol raises the pitch of the note by a quartertone.

Quarter tone: A musical terminology often applied in the Persian music or generally Eastern music which changes the tuning of the note by a quarter of an interval. In fact, the application of the quarter is erroneous, but it is an established terminology. The changing of the tuning performed by an amount approximating a quarter of a note.

Semi tone: is half an interval, which is the smallest interval applied in Western music. This musical term is common in Persian music.

List of Abbreviations and Symbols

BM	Boltzmann Machine
CA	Cellular Automata
CRBM	Conditional Restricted Boltzmann Machines
DBM	Deep Boltzmann Machine
FN	False Negative
FP	False Positive
GA	Genetic Algorithm
H-Creativity	Historical Creativity
-Inf	Minus Infinity
LBM	Liquid Brain Music
LPM	Liquid Persian Music
MIDI	Musical Instrument Digital Interface
Multimodal DBM	Multimodal Deep Boltzmann Machine
NAN	Not A Number
OpenAL	Open Audio Library
P-Creativity	Psychological/Personal Creativity
RBM	Restricted Boltzmann Machine
SVM	Support Vector Machine
SVR	Support Vector machine Regression
TN	True Negative
TP	True Positive

Chapter 1. **Introduction**

1.1 The Persian Musical Dastgāh

Dastgāh is a modal system in traditional Persian music (During, 2011; Farhat, 1990; Zonis, 1965). There are twelve principal groups of modes in Persian music. Each Dastgāh consists of individual pieces called Gushé. Performing in a Dastgāh begins with Darāmad which has the mode and melodic patterns of the Dastgāh itself (Farhat, 1990; Nettl, 1987). Then the modulation occurs; which can be a move from one Gushé to another or a change in the central tone, or Shāhed note, which may consist of an alteration in the tuning as well (Peyman Heydarian & Reiss, 2005).

The current Dastgāh system is the result of centuries of evolution of Persian music conjoint with historical and cultural transmutations. However, there are still varieties of other possible musical systems and melodies waiting to emerge. Once modulated with Western music it can be considered as a potential source for cross-cultural interactions. Although Persian music has vast musical systems and intervals in comparison to its Western contemporary music counterpart, one of the problems encountered is the entrapment in the existing structures. This makes the composition more reliant on the emergence of great masters whom with their novel creativity and familiarity of the complexities of Persian music are able to take a step forward in this field and add new melodic pieces to different Dastgāh. Therefore, the variety of melodies and *Gushé* in a Dastgāh are limited to what was produced in the past. Perhaps another possible weakness could be the adherence to the current musical frameworks; the creation of a variety of musical systems giving a new essence to the current

Persian music will enable both the composer and the audience to experience novel atmospheres.

1.2 The Emergence of New Musical Instruments and New Musical Forms

Archaeological discoveries, on primary musical instruments, suggest that the human's natural tendency towards music has a long history. These primary musical instruments were usually made from elementary materials (for example, skin, bones and wood) (Rault, 2000). The first humans were affected by natural sounds such as thunderstorm, rain, water flow, sound of animals, etc. They were trying to use the power of sounds to express meanings and communicate. The invention of musical notes are more assigned to a random process, where the first people gained different effects by changing the physical factors in a sound producing media, for example by changing the length of an excited string (Khaleghi, 2013). The formation of musical octaves and tunings for many musical instruments continued through trial and error. Mathematicians, for example *Pythagoras* (Farhat, 1990), helped classify the mathematical proportions between notes and make the process more methodological. This innovation had great impact on the progression of what we have today as musical systems.

The evolution process of musical instruments have also benefited from cultural transmutations. New forms of musical instruments were introduced to different countries and been customized to different cultures and tastes. For instance, Santur the hammered dulcimer is a famous Persian musical instrument. There are different versions of Santur around the world, for example the *Kamboujfi* from India, *Yangqin* from China, *Butterfly* from Britain, *Macper* from *Austria*, *Santorini* from Greece, Zither from United States (Sadie & Tyrrell, 2001). Some of the musical instruments are described as evolutions of previous versions. For example, the Piano follows the same string excitation mechanism as a hammered dulcimer and can be considered as a modern version of the traditional hammered dulcimer. Likewise, musical systems had undergone similar evolutionary processes and gradually became as inseparable parts of various cultural signatures. With the advent of new technologies, electrical musical instruments were introduced where acoustical effects were added to the compositions. Iranian music has benefitted from electronic music as well. Electronic music in Iran was pioneered by *Alirezā Mashāyekhi* (Cont & Gluck, 2008; Gluck, 2013). The contemporary Iranian music uses electroacoustic and sound designing in order to add electronic musical elements to the compositions. With the advent of computational and artificial intelligence tools, the search for new music, and musical forms, can be advanced at a greater pace with the possibility of truly novel forms emerging.

Boden classifies creativity into two main groups according to the novelty of their origins. These are psychological and historical types of creativity (P-creativity and H-creativity respectively). The P-creativity refers to the type of creativity that is new to the person who created it, no matter how many times it has happened before. The historical type of creativity refers to a type of creativity, which has never been manifested throughout history. For instance, undergraduate students implementing the Perceptron algorithm are performing P-creativity; they may come to new ideas and concepts, which were already discovered. However, a PhD student should come up with a new idea in their thesis, whether in methodology or in the nature of the produced knowledge, which is H-creative.

Various musical genres have been revolutionized and continue to be developed; therefore, they can be categorized as H-creativity from many respects. However, once new styles are introduced, the artists or authors are likely to explore the same creativity space. Maintaining the previous established musical concepts, makes the productions likely to be considered as P-creativity types. For instance what are considered as consonant and dissonant chords and the related rules for chord progressions remain the same throughout newly composed music. Similarly, composing Traditional Persian music is followed in one of the existing Dastgāh. The number of musical Gushe have not exceeded the ones collected by '*Mirzā Abdollah Farāhāni'* from old musical references about a century ago (Farhat, 1990). However, many of the *Gushe* have been forgotten or altered through history as well, in a way that only the names of those pieces are available from ancient Persian literatures (Khaleghi, 2013).

The Dastgāh concept embeds ancient spiritual messages and conceptual meanings. This introduces them as complex structures. Musical novices are often encouraged to stick to the available framework as a requirement for advancing to higher levels. Compositions of new pieces are often suspended until the mastery of the Dastgāh itself; at the time when the artist gains a deep perception of Dastgāh concept and Gushe meaning. The descriptions of the Dastgāh are often limited to words in ranges of emotional expressions together with limited general tonal characteristics of the melodies in a Dastgāh. Therefore, composition in Dastgāh music is often reliant on the emergence of great masters who intuitively understand the complexity of traditional Persian music. However, these compositions are still exploring musical forms within quite constrained musical spaces, while there are various other musical possibilities, beyond these constraints, that transcend these traditional forms.

There is the controversy of how the computational creativity experiments in the area of music art can generate quality music, particularly as they are in their early stages of development. This is particularly the case with Dastgāh where most current traditional musical pieces are regarded as masterpieces. The research in computational creativity is strongly interconnected to cognitive sciences and psychology. While there are so many unresolved questions in those areas, there is the hope that these types of experiments unveil some of the answers or at least lead to new approaches for making the underlying nature of creative processes more tangible.

It should be emphasized that the creativity space for producing music is, to all intents, infinite. The aim is to find the proper means for exploring those spaces. The fact is that there are human composers, machines that can create and compose, and human composers who use machines to create. If it is accepted that the computer can produce the desired music, then there would be the debate upon whether the composing machines would supersede their human counterparts.

The process of introducing novel musical systems with their related musical theory can be considered as historical type of creativity. Exploring the possibility of creating novel Persian Dastgāh musical systems by help of computational intelligence tools is one of the ultimate goals of this project. However, a question that comes into mind is how to produce a novel style of music, which can be associated with traditional Persian music. There is always the possibility that the search space gives new types of music, which does not belong to any specific genre, and are quite new in many respects. If the search space is constrained to specific compositional rules, there is the hope that the new type of music generated possesses the essential and key features of a desired musical style or class.

1.3 On Computational Creativity

"What is creativity?" –This can be considered as an open-ended philosophical question. There are no boundaries for creativity, yet binding creativity in a framework for a definition is a necessary but difficult task. However, an artefact has some representative features which describe its qualities to some extent. These qualitative descriptions clarify the attributes an artefact should have to be considered as a piece of artwork. Amongst all descriptions, what is clear is that art and novelty are two inseparable concepts. Sometimes a black circle on a white canvas is defined as a masterpiece and is exhibited in art galleries (Black Circle by artist Kazimir Malevich is kept in State Russian Museum). The work of John Cage in his composition "four minutes and 33 seconds of silence" (Gutmann, 1999) unbounds framed viewpoints towards art and creativity with avant-garde music. In a silent musical performance he lets the energy from audience noise vibrate the strings of a grand piano. The interaction of audience noises and musical instruments made John Cage's performance one-of-a-kind. There are criteria for defining creativity other than novelty, for example quality (Ritchie, 2007). This discusses how the creation is a high-quality instance of its genre. Jon McCormack defines this attribute of creativity as being exhibitable (McCormack, 2007).

Two different viewpoints exist about man-machine creativity. The machines that create art-like productions, and the machines which are autonomous in creating art (McCormack, 2007). The aim of creating could be to satisfy an audience or could involve the exploration of general meaning of creativity, without contributing to human comprehension or appreciation.

Boden (Boden, 2000) defined three types of creativity: combinational; exploratory; and transformational. She stated that all creativity types can be modelled by artificial intelligence. Combinational creativity consists of populating pre-existing materials and linking them in an artistic manner for generating new ideas. Exploratory creativity includes navigating in a *conceptual space* with implicit constraining rules. This exploration can result in discovering new transformed styles which would not have existed before an alteration happening on one or more of their defining dimensions (transformational creativity).

1.4 On Algorithmic Composition

The use of algorithmic composition has been under investigation for many years with different motivations: mechanization of music production; exploration of the behaviour of the algorithms; mathematical models in generating the patterns; studying the cognitive behaviour of creation in human being (McCormack, 2007); and modelling biological patterns in nature in respect to music.

Mechanisation of music generation were done for producing melody, rhythm, harmonization, and counterpoint or imitating a specific genre of music or composition style (Fernández & Vico, 2013). The level of automation varies from generating motifs for inspiration to more complex corpus composition tasks. Computer aided algorithmic composition is the term applied for assisting musicians in the composition process and providing them with new materials. Some available frameworks or languages for making musical software include Csound (Boulanger, 2000), MAX/MSP (Puckette, 2002; Zicarelli, 2002), while some musical software include EMI (Cope, 1987, 1996), GenJam (Biles, 1994), and LBM (Turner, 2008; Woods, 2009)). Higher levels of composition automation target minimal or no interactions with human, Melomics corpus generation (Diaz-Jerez, 2011) is an example of this kind.

Methodologies in algorithmic composition can be categorized, based on the report from (Fernández & Vico, 2013), in four groups: knowledge based systems, machine learning, evolutionary algorithms, and computational intelligence (e.g. cellular automata). They have been widely used both for style imitation and creating novel music. There have been good progress with the research into genre imitation; successful applications include Strasheela (Anders, 2007). All of the aforementioned categories except the last one apply human knowledge in their application. However, cellular automata are able to generate novel material without utilising existing human domain knowledge. This potential for creativity makes them well suited for exploring new dimensions of music composition. Some of the possible future research directions for algorithmic composition includes hybrid methods (Fernández & Vico, 2013) that use cellular automata (CA) as their music generator.

1.5 A Brief Overview of Liquid Brain Music System

This subsection gives a brief overview on Liquid Brain Music (LBM) system (Turner, 2008; Woods, 2009). The underlying LBM mechanism for audio generation has been established as the basis for this research. LBM is a Cellular Automata (CA) based audio software, which was first developed at the University of Hull. In LBM, patterns are extracted from CA progressions. These patterns are mapped to the musical space and are employed for populating the parameters of an audio synthesizer. As will be later discussed in chapter 3, CA are computational intelligence tools which can generate creative materials without using the human domain knowledge. The generative power of CA are established as the main source for obtaining a variety of patterns. LBM system also provides the basic means for interpreting the emerging patterns from CA progressions. The scope of this PhD thesis is refined by the constraint of investigating how the LBM system can be applied to the task of generating Persian Dastgāh-like music. Chapter 4 provides further details on LBM system and the developed version Liquid Persian Music (LPM).

1.6 PhD Hypothesis and Research Objectives

The research hypothesis explored in this PhD thesis is that it is possible to create Dastgāh-like music using appropriate computational intelligence methods. This gives rise to the following questions:

- How to use computational intelligence methods to produce creative (audio) artefacts?
- How to guide this process to produce Persian Dastgāh-like music?
- How to assess our musical productions in terms of aesthetics?
- Is there a measurement for the creativity of the generated materials?

We hope that by applying the main concepts of LBM software and establishing advanced systems using computational intelligence tools we are able to create and experience new dimensions of music composition for generating Dastgāh-like music. The research will see the application of computational intelligence tools (such as cellular automata and Boltzmann machines) for generating creative material without using explicit human defined domain knowledge. On this account the main objectives explored in this thesis are:

- Developing LBM to produce Dastgāh-like music. It should be emphasized that part of producing Dastgāh-like music is to use appropriate timbres compatible with Persian music. On this account, the LBM synthesizer should be updated to produce the sounds of a Persian musical instrument.

- Finding aesthetical criteria, which are in accordance with Dastgāh music. We hope to create Dastgāh-like music by the help of computational intelligence tools and by finding aesthetical criteria, which are in accordance with Dastgāh music. The aesthetics should be measurable in order to be employed in the music generation process.
- Designing systems based on computational intelligence, which enable the navigation of creativity spectrum including combinational, exploratory and transformational types of creativity. Targeting these various types of creativity in algorithmic composition are hoped to be achieved by the application of various computational intelligence tools and techniques.
- Investigating the application of evolutionary algorithms for creating Dastgāh-like music by the usage of appropriate fitness functions. This step can be considered as analogous to the exploration in a musical space for obtaining the desired aesthetical effects. This objectives embeds further investigations for determining an efficient search space traversal. Furthermore, competent designs for the genotypes, phenotypes, and fitness function should be taken into account.
- Measuring the creativity of the designed systems. Evaluating the creativity of the systems should be followed through standard methodologies in the computational creativity area. Evaluating the generated audio materials would also give insights to the extent of the success of this project.

1.7 Thesis Roadmap

This section presents a structural overview of the research presented in this thesis. The next two chapters deliver fundamental backgrounds and literature reviews underpinning this PhD thesis. The background materials are arranged in two different chapters according to the nature of the application of the presented tools and techniques. This would implicitly reduce the volume of the information delivered as background material in one chapter. Chapter 2 starts with an overview on audio synthesis techniques which are used in LBM and LPM systems. Chapter 2 covers the musical information retrieval topics which are employed for extracting features from traditional Persian music and LPM sequences later in chapter 5. The fundamentals of support vector machine regression (SVMR) model is also presented in chapter 2. SVMR was employed as a fitness function for an evolutionary architecture which is presented in chapter 5. Chapter 3 revolves around computational intelligence and machine learning tools used as creative or generative processes. This includes cellular automata as a computational intelligence tool, Genetic Algorithms (GA) for exploring the musical space, and Boltzmann machines as stochastic generative models. The discussion on CA are also motivated by the fact that LBM system is cellular automata based. Chapter 4 is dedicated to describing the LBM software and the developed version, Liquid Persian Music (LPM), in further detail. In chapter 4, Cellular Automata are presented as the main core of creative behaviour for a novel algorithmic composition system ('*Liquid Persian Music*'). Chapter 5 and 6, experiment with the Boden's three types of creativity (Boden, 2009) in the context of traditional Persian music. Chapter 5 portrays a model for performing a meaningful navigation in a conceptual musical space. Genetic algorithms are responsible for evolving LPM voices sequences which are considered as genotypes in the population. The GA search is guided by comparing the individuals' aesthetical characteristics with those of Persian music pieces. The model presented in chapter 5 is responsible for experimenting with exploratory and possibly transformational types of creativity. Chapter 6, benefits from various characteristics of Boltzmann machine families for implementing a simple audio generation model based on Boden's associative creativity notion. The feature extraction powers of Boltzmann machines introduce a novel approach for interpreting the patterns from CA progressions, as well. In the seventh chapter, the results of human evaluation on the produced audio in the experiments in chapters 5, and 6 are presented. The criteria of the designed evaluations are in alignment with the objectives of the thesis. Chapter 8 is dedicated to a discussion on the conducted experiments in terms of computational creativity and the specified targets. Chapter 9 concludes the thesis, and proposes further research directions. Extra detail relating to specific chapters is given in the various appendices.

Chapter 2. Background and Literature Review

This Chapter revolves around overviews on some important and fundamental concepts and computational techniques. These are employed in approaching various agendas in different stages of the conducted research. There are three main sections in this chapter, namely, musical instrument sound synthesis, musical data mining, and some important machine learning concepts. This chapter starts with an elemental discussion on digital sound synthesis of musical instruments. Some of these techniques are employed to design and implement the core of the synthesizer model to be introduced in Chapter 4. The subsequent section brings up a review on musical data mining together with different sorts of features which can be extracted from audio for further processing. These features are later extracted from Persian music and machine compositions, and been post processed using the data cleaning techniques mentioned in this chapter. The Machine learning section covers brief reviews on support vector machine (SVM) and support vector regression (SVR) models which are employed as fitness function evaluators in the first experiment (chapter 5). Support vector machine models are subsequently trained to differentiate between the human compositions and those of the designed machine composer. The machine compositions are produced in an evolutionary environment using the scores given to them by applying support vector regression model. Chapter 3 considers further machine learning techniques where they are used as part of the creative process.

2.1 Musical Instrument Sound Synthesis

There are various methodologies for synthesizing the sound of a musical instrument. Additive synthesis (Manning, 1985; Roads, 1978, 1988) model works based on the addition of a series of signals to produce the final waveform. Additive synthesis works as a fundamental part of other synthesis techniques. The oscillators in additive synthesis technique are parameterized to generate sinusoidal waves. All the generated sinusoidal waves are later added together to produce the final wave.

Other sound synthesis methodologies includes the implementation of the physical model of the musical instrument (Essl, 2002). In this technique, all the physical components concerned with the sound production are simulated. These components include the excitation of the resonating part of the instrument and the instrument body amplification. The process can result in obtaining an assembly of sequenced digital filters known as digital waveguide models (Smith III, 2004). The design of the filters for the constituent parts of a digital waveguide model usually takes place by numerical calculation of a series of partial differential equations which describe the mathematical model of vibration phenomena (Bilbao, 2006; Gustaffson, Kreiss, & Oliger, 1995; Strikwerda, 1989).

2.1.1 Musical Instruments Sound Synthesis based on the Physical Model

The excitation of a musical instrument plays an important role in designing the characteristics of the produced tone (Fletcher & Rossing, 1991). Various classes of musical instruments whether in the groups of stringed, brass, wind, and percussion musical instruments family, follow their own sound production rules. The strings are considered to be the hearts (Bank, 2006) of the stringed musical instrument models. Some of the important factors that determine the timbre of stringed musical instruments are the way the strings are excited, allowing the differentiation into groups of bowed, struck or plucked instruments. Other effective factors are the coupling between the strings and the instrument body, the way the strings are fastened to the instrument, the strings and wood materials, the instrument body patterns, and the various physical aspects that take place between the stringed instrument and the air itself as wave transfer medium. Modelling the musical instruments body are usually possible by measuring the impulse response of the body to a transient signal. A digital filter model would take the responsibility of modelling the musical instrument body (Karjalainen & Smith, 1996).

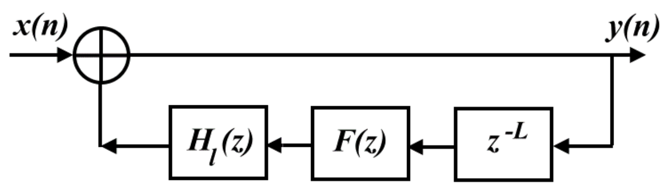


Figure 2-1. Block diagram of a plucked string model.

The delay line $\boldsymbol{z^{-l}}$ is responsible for implementing the loop delay. $\boldsymbol{F(z)}$ is used for simulating the fine-tuning of the model. The loop filter $\boldsymbol{H_l(z)}$ implements the damping of the harmonics (Jaffe & Smith, 1983).

Figure 2.1 shows the block diagram of a plucked string model. The general string synthesis model $S(z)$ consists of a delay line z^{-L}, a fractional delay filter $F(z)$, and a loop filter $H_l(z)$. The string length (L) determines the fundamental frequency (f_0) of the excited string. By having (f_s) as the sampling frequency, one can accomplish the delay line length (L) according to the following equation:

$$L = f_s/f_0 \tag{1}$$

The value obtained for the delay line length may not be an integer value. Therefore another filter is usually introduced to model the fractional part of the string length and compensate for the non-integer part of the division presented in equation 1 (Erkut, Välimäki, & Karjalainen, 2000). Therefore, the delay line and the fractional delay filter both contribute to the value of the fundamental frequency of the produced tone. The loop filter guarantees the attenuation of the harmonics of the generated signal (Jaffe & Smith, 1983). The loop filter is generally implemented as a low pass filter. This low pass filter has a frequency dependent damping effect. The loop filter is designed in a way that different harmonics are attenuated with different rates.

Synthesis tool kit (Cook & Scavone, 2008) is an open source application programming interface written in C++ which includes both low and high level signal processing classes. Synthesis tool kit provides implementations of the general string model. Various classes of stringed musical instruments implemented inside synthesis toolkit are based on the general string model. The parameterizations of the filters vary for different musical instrument cases. In addition, more filters may accompany the general string model for assimilating various physical effects happening with a musical instrument.

2.2 Music Information Retrieval

Music information retrieval is a broad branch of science which focuses on attaining information from musical data (Muller, 2015; Ogihara & Tzanetakis, 2012). Music information retrieval benefits from multi interdisciplinary research in musicology, music theory, digital signal processing, data mining, machine learning, and psychology (McKay, 2010). This field of research is expanding rapidly with the advent of new technologies, facilitating the process of extracting features from musical data. The assumption is that there are various aspects of music yet to be recognized as features, but once retrieved they can give insight to the process of music production itself.

The information attained from musical data are employed in various applications including automatic music transcription (Klapuri & Davy, 2010) and archiving, obtaining musical notation from audio, audio mining (Lerch, 2012), automatic music composition, watermarking, genre/ artist classification (Tzanetakis & Cook, 2002), plagiarism detection, and associating musical features with moods.

Many software tools and libraries have been designed so far for the purpose of music information retrieval. JMIR (McKay, 2010), MIR Toolbox (Lartillot & Toiviainen, 2007, 2008), MIDI toolbox (Eerola & Toiviainen, 2004), Aubio, Essentia, BeatRoot (Dixon, 2007), and Praat (Boersma & Weenink, 1995), are some of the libraries and software used for analysing musical data. JMIR is a suite of open source software tools written in Java for its platform independence advantages. The components include ACE XML, jAudio, jSymbolic, jWebMiner, and jMusicMetaManager aimed to work in isolation or in combination with each other (McKay, 2010). MIR and MIDI toolboxes consist of set of functions implemented in Matlab (Matlab, 2016) for extracting features from audio and symbolic musical data respectively.

2.2.1 Audio Feature Extraction

In this subsection, some of the widely used techniques in audio feature extraction as used in this research are discussed. These include Signal segmentation, Frame Based Analysis, Fourier transform, Envelope detection method, Autocorrelation and Peak picking algorithms.

Signal segmentation: segmentation techniques divide the signal into smaller chunks, for studying their characteristics individually. The signal can be segmented to equal sized fragments or chunks with variable lengths after gaining some more information about the signal. For example, by first obtaining the overall envelope of the musical signal and locating the bursts of energy as note onsets, the signal can be segmented, with each subsequent piece containing of one note (Lartillot, 2013).

Frame Based Analysis: Analysis of a signal on a frame by frame basis provides the chance of studying the changing behaviour of measured features during the signal progression (in time or frequency domain) (Lartillot, 2013).

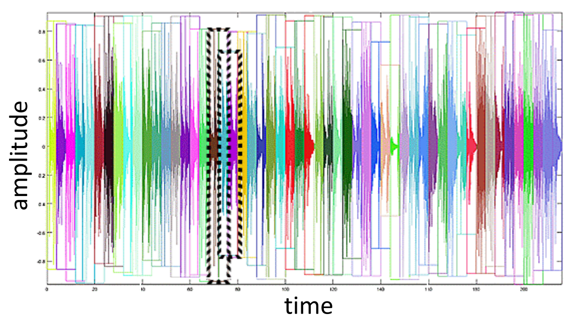


Figure 2-2. An example of framed audio signal.

An example of framed audio signal with overlapping frames. The size of the frames are 8 seconds, and the audio file is about 220 seconds.

This technique includes sliding a window over the original signal in an overlapping or non-overlapping manner determined by the hop size (as seen in Figure 2.2 for the case of overlapping frames). The hop size shows the overlapping between the windows. The length of the framing window is often taken as constant value in number of samples or seconds. Framing is often applied or accompanied by other signal processing techniques as well as transforms.

The **Fast Fourier transform** is the basis of many of the algorithms and toolboxes that were employed in this research for extracting features from Persian music database. Fourier transform decomposes a signal to its constituent sinusoids. Obtaining many of the features in a time domain is a complex task. The Fourier transform converts a signal from its time domain representation to frequency domain (Oppenheim & Schafer, 2013).

Envelope detection methods: The envelope of a signal produces an external contour over a signal regardless of the related temporal details. The envelope curves provide the advantage of detecting some musical events such as note onsets, and duration. Hilbert method and down-sampling are two important approaches for obtaining the envelopes of a signal (Frerking, 1994; Marple, 1999; Tretter, 2008).

Autocorrelation method provides a measurement of the amount of inner similarity of a signal with itself or with shifted versions of itself within specific number of samples (Broersen, 2006; Proakis & Manolakis, 2007). It is often applied for recognizing repeating patterns of events as well as periodicity, and tempo. Autocorrelation methods can also be used to estimate the pitch of a signal.

Peak picking algorithms: Onset detection curves have an important role in measuring the periodicities present in musical signals. This gives clues for estimating the tempo in beat per minute units (Lartillot, 2013).

2.2.2 Zipf's Law

Zipf's law (Zipf, 1949), determines the scaling characteristics of many natural phenomenon and is employed in physics, social sciences, and language processing. Zipf's law states that the frequency of occurrence of an event is inversely proportional to its statistical rank. Zipf's law is expressed by formulations such as $Z \propto\ r^{-a}$. Here Z stands for the frequency of occurrence of an event and r is the statistical rank. The rank is to the power of a in the given formulation. The power (a) in the formulation characterizes the statistical nature of the phenomenon. The more the a parameter approaches 1, the better the distribution approximates to the ideal Zipfian distribution. Another important statement for Zipf's law is through $P(f) = 1/f^n$ equation. $P(f)$ specifies the probability of occurrence of an event with rank f. Zipf's ideal distribution is also known as pink noise when it has the value of $n = 1$ in the formulation. In cases of n=0 and n=2 the noises are called white, and brown noises respectively (Manaris, Romero, et al., 2005).

In order to determine the Zipfian characteristics of a phenomenon, the occurred events within the scope of the studied phenomenon are recorded within a dataset. The elements in the database are then ranked in a descending order considering their importance or prevalence. The ranks and the frequency of occurrence of the events are then taken to a logarithmic scale and plotted against each other. A linear regression is applied to the resultant graph. The slope of the line determines how the distribution follows Zipf's law. A slope of -1 indicates an ideal Zipfian distribution. The R-squared value associated with linear regression has a range between 0 and 1 and indicates whether the linear regression model fits well to the data.

Zipf's law has had applications in audio and music analysis and generation as well. Voss and Clarke (Voss & Clarke, 1978) first observed the existence of 1/f like distributions in audio. They later developed an innovative algorithm for producing music which had spectral densities consisting of white, pink and brown noises. The outcomes suggested that the music produced based on pink noise had more desirable musical attributes and were more pleasing to hear due to their self-similarity characteristics. The white noises produced music seemed too random while the music generated considering the spectral density of brown noise was monotonous due to the high correlation between the musical events.

Other more recent musical applications of Zipf's law can be traced in (Lo & Lucas, 2006; Manaris, Machado, Mccauley, Romero, & Krehbiel, 2005). In the musical domain, the Zipfian characteristics are expressed through a number of metrics. Various events like note numbers and durations occurring in a piece of music are enumerated and stored. The frequency of occurrence of the musical events are extracted and mapped in a log-log scale versus their rankings. The obtained slopes of the linear regression model would vary between $-\infty$ to 0. The more the slope tends to minus infinity, the higher the levels of monotonicity identified within the audio. The experiments in (Manaris et al., 2011; Manaris, Romero, et al., 2005) demonstrate successful applications of Zipf's metrics in music generation. Manaris et al. introduced two sets of simple and fractal Zipfian metrics that were captured from music. The simple metrics were obtained by calculating musical events such as pitch and chromatic tones separately and in relation to each other. The fractal Zipfian metrics were built on top of simple Zipfian metrics. Manaris et al. employed the acquired Zipfian metrics to train a neural network for classifying musical pieces based on the musical style and their composers. The success of this application was reported to exceed by ninety percent. In the context of Manaris et al. experiment, Zipf's law was able to capture useful information from music as well as being able to further specify the Zipfian characteristics of musical pieces. This research forms part of the reasoning why Zipfian metrics play an important role in this thesis for extracting aesthetical characteristics of Persian music and generating audio on that basis.

Music is a subjective matter and there might not be universals in measuring the aesthetics of musical pieces. However, there is still the need to investigate aesthetical criteria applicable to the research at hand. Being measurable is one of the important specifications that the aesthetical criteria should have. Measuring the aesthetics should happen in an automatic way without human intervention. These qualities help in embedding the aesthetical evaluation as a part of the audio generative process. To this date the author have not found any aesthetical criteria other than Zipfian metrics which has these qualities. Zipfian metrics are measurable for Persian music pieces and for Liquid Persian music audio sequences. Some of the other potential aesthetical criteria refer to consonance and dissonance of the audio samples. These qualities may be determined by cultural preferences or through psychoacoustical tests. The consonance and dissonance of the audio samples are measurable and can be quite helpful while working with the ADSR envelopes or harmonizing the musical notes. However, in this thesis the attention was towards generating audio with desirable melody rather than harmony or ADSR envelopes in the first place. Zipfian metrics have been found suitable option for the task at hand. It is worthwhile mentioning that there are some ongoing PhD researches for identifying aesthetics in Eastern music which might be useful for future investigations.

2.2.3 Data Cleaning and Attribute Selection

Data mining (Kononenko & Kukar, 2007) consists of processing data through several phases for extracting information and discovering knowledge from potentially large scale and complex databases. The extracted information can be in the form of patterns, structures, or behaviours which are not explicitly represented in the raw database.

Data cleaning is an important pre-processing stage which is performed on raw data prior to any other data mining tasks (Han, Kamber, & Pei, 2011). Purifying the database from inconsistent data has benefits. It not only saves subsequent processing times and refrains from over-fitting problems by identifying outliers; it also increases confidence in the subsequent data mining processes. The quality of the data base (Rahm & Do, 2000) can be increased by removing redundant, and invalid data and handling noisy, or missing and unknown attributes or records (Han et al., 2011). The reduction process should not affect the integrity of the information (for example loss of information).

The attribute selection methods represented in this section are going to be employed in chapter five for selecting useful and relevant attributes out of the Zipfian metrics databases for Persian music, and the ones related to our system's output. Four attribute selection methods are going to be employed: *ReliefF*, *Information Gain*, *Gain Ratio*, *Symmetrical Uncertainty*. These methods are briefly overviewed here.

ReliefF (Kira & Rendell, 1992; Kononenko, Šimec, & Robnik-Šikonja, 1997) is an attribute selection method which works based on obtaining a worthiness measurement for an attribute. A weight is assigned to each of the features based on their capabilities in discriminating between the classes. The associated weights are obtained based on the calculation of probability values: the probabilities of the items of each of the classes having the same value for a feature are calculated. Likewise, the probabilities of the samples from the different classes having the different values for an attribute are computed. The higher the weights related to each of the features, the more chances the attribute will have for survival in the attribute selection process.

Information gain, gain ratio (Karegowda, Manjunath, & Jayaram, 2010; Novaković, Strbac, & Bulatović, 2011) and symmetrical uncertainty are examples of attribute selection techniques which are based on the concept of entropy. Entropy is a measure used in information theory and contributes to the unpredictability of a system. The information gain related to a feature in the dataset attains a measurement in the reduction of the entropy of the class. The information gained about attribute A by observing attribute B is the decrease in the entropy of A. In other words, the information gain determine to what extent the knowledge gained about an attribute B can predict attribute A or whether the two attributes are completely uncorrelated. The information gain measure is symmetrical for attributes A, and B (Novaković et al., 2011). The bias about information gain method is its tendency towards selecting attributes with large number of values (Karegowda et al., 2010; Novaković et al., 2011). The gain ratio, and symmetrical uncertainty feature selection methods are extensions of information gain method. They aim to resolve the bias associated to information gain method by performing a normalization procedure.

CfsSubsetEval (Correlation-based Feature Subset Selection) (Hall, 1999) is a feature selection method in Weka software (Hall et al., 2009) which is also employed in chapter five as an attribute selection option. CfsSubsetEval evaluates the worthiness of a subset of attributes. The predictive power of each of the features are calculated. The correlation of each of the attributes and the classes are calculated together with the inter correlation of the attributes. The chosen subset is the one which has attributes with maximum correlation with the classes and minimum inter correlation.

2.3 Machine Learning

Machine learning is a field in computer science where the goal is to produce learning algorithms capable of generalizing to new data after being trained on existing data. The trained models can be utilized for decision making and prediction applications. Machine learning methods follow learning processes which fall into broad categories of supervised and unsupervised learning.

Bayesian classifier, rule-based classifiers, and decision tree induction (Kononenko & Kukar, 2013) are some of the machine learning models which are used in chapter 5 (table 5-5) for comparing their performance as classifiers. Naïve Bayes classifiers are probabilistic models which use Bayes theorem. Naïve Bayes classifier treats each of the attributes in the feature vectors independently. A probability is achieved for each of the attributes which shows the independent contribution of the feature for determining the classes. Decision tables are predictive machine learning tools. Decision tables predicts the target classes of the items by following a set of branches related to the features of the items. J48 is a Weka (Hall et al., 2009) implementation of the C4.5 algorithm (Quinlan, 1993) which is widely used for generating decision trees. Rule-based classification utilizes a collection of logical rules (IF-THEN-ELSE) to determine the target classes. Detail on support vector regression, and Boltzmann machines is supplied later in this and the next chapter.

2.3.1 Confusion Matrix

A confusion matrix (Witten, Frank, & Hall, 2011) allows the portrayal of the performance of a classifying algorithm and is widely utilized in machine learning applications (Powers, 2007; Tom Fawcett, 2006). The confusion matrix (Witten et al., 2011) clarifies how a classifying algorithm positions the data with regard to the class labels. The confusion matrix table is built upon true positive, true negative, false positive and false negative terminologies as shown in table 2-1.

Table 2-1. Confusion Matrix.

	Predicted : True	Predicted : False
Actual : True	True Positive	False Negative
Actual : False	False Positive	True Negative

True positive (hits) specifies the number of instances the classifier predicted to be true and they are actually true. The true negative items determine the number of instances the classifier expects to be false and it correctly classified them. The false positive individuals are those which are expected to be false but they are mislabelled as true items. The False negative (misses) determines the number of items predicted as false but are actually true. Overall true positive and true negative instances show the number of items being correctly classified (Witten et al., 2011).

Table 2-2 depicts some of the terminologies derived from confusion metrics and their implications (Witten et al., 2011). Accuracy alone does not necessarily show a good performance of the classifier. One should take the other measurements (for example, sensitivity, and specificity) into account to see how the classifier performs in predicting the individuals in the pre-specified categories. These metrics are used in the experiment described in Chapters 4 and 5.

Table 2-2. Some terminologies and concepts inferred from confusion matrix which are used in this thesis.

Terminology	**Concept**	**Formula**	**Formula Number**
Accuracy	Identifies the proportion of successfully classified items over the total number of instances.	$\frac{TP + TN}{TP + FP + TN + FN}$	(2)
Sensitivity	Shows the rate of correctly classified items over the total expected true items.	$\frac{TP}{TP + FN}$	(3)
Specificity	Shows the rate of the correctly excluded non-true items over total actual false items.	$\frac{TN}{TN + FP}$	(4)
Negative Predictive Value (NPV)	Shows the rate of true negative results over all the items predicted as false.	$\frac{TN}{TN + FN}$	(5)
Positive Predictive Value (PPV)	Shows the proportion of true positive outcomes over all the items classified as true.	$\frac{TP}{TP + FP}$	(6)

2.3.2 Support Vector Machines

Support Vector Machines (SVM) are popular machine learning tools which are used widely for both classification and regression applications (Vapnik, 1995). Support vector machines have high performance in solving classification and regression problems with small numbers of data points, while effectively avoiding high number of dimensions present in the database (Yin, Wu, Luo, & Gao, 2015). One of the advantages of SVM over some other machine learning tools is that they are trainable using a small number of training samples. SVM can be employed for both linear and nonlinear classification problems. Kernel functions are an important property of support vector machines, employed for data which are nonlinearly separable. The kernel functions are designed to project the data samples to a higher dimensional space, where they can be linearly differentiated.

2.3.2.1 Support Vector Regression Model

Support vector machines can be used for regression applications as well as classification. A support vector machine which performs regression tasks is called a Support Vector Regression (SVR) model. In this subsection, some major theoretical concepts for solving the optimization problem for the case of SVR are discussed which are adapted from various sources (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Smola & Schölkopf, 2004; Steinwart & Christmann, 2008; Theodoridis, 2009; Vapnik, 1998). Support vector regression model maintain the same principles as support vector machines with slight differences. In traditional linear regression, the problem is to map a linear function $f(x) = w^T X + b$ to the data samples in least squares sense as:

$$min \sum_{n=1}^{N} (y_n - w^T X_n - b)^2 \tag{7}$$

This is only valid where a linear function can be well mapped to the data. If data are not linearly distributed, they can be transformed to a higher dimensional space, where a linear function can be mapped to them. Since the training samples are not necessarily placed on the regression line, the concept of margins are introduced to the problem (Wang & Gao, 2012). The bigger the margin, the higher the generalization capability of the trained support vector model to the new data.

In SVR a small boundary with the distance of ε from the separating hyperplane is assumed. All the points inside this boundary are ignored as having any contributions for determining the separating hyperplane. A function of this type is called an ε-intensive loss function.

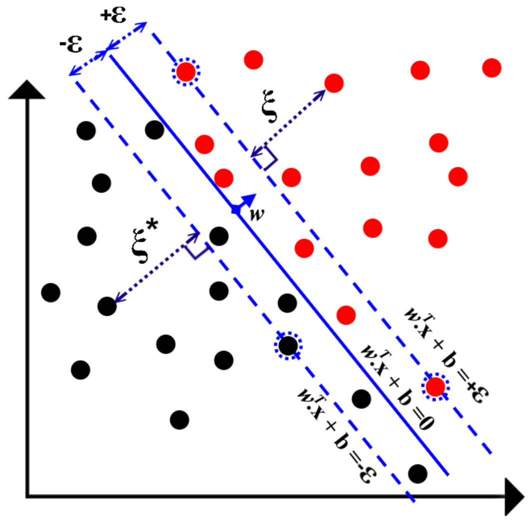


Figure 2-3 A schematic example of the support vector regression model mechanism.

Figure 2-3 illustrates an ε boundary condition for the case of regression. The solid blue line represents the optimal hyperplane surrounded by an ε tube. The solid blue line works as a regression line which has a deviation allowance of ε. The data samples outside of the tube are penalized proportionate to their distance from the marginal line. ξ_i and ξ_i^* are slack variables which are the orthogonal distances of training points from the ε boundary. Slack variables determine the deviation of the training samples from the ε boundary. ε-intensive loss function for SVR model denotes zero loss for data points inside the ε tube. The slack variables are zero inside the tube. The SVR problem can be solved through a constrained optimization problem in its primal form formulated as:

$$Primal\ problem: \begin{cases} minimize: \dfrac{1}{2}\|w\|^2 + C\displaystyle\sum_{n=1}^{N}(\xi_n + \xi_n^*) \\ subject\ to: \begin{cases} y_n - w^T X_n - b \le \varepsilon + \xi_n \\ w^T X_n - y_n + b \le \varepsilon + \xi_n^* \end{cases} \end{cases} \qquad (8)$$

Slack variables $\xi_n, \xi_n^* > 0$ are introduced to the optimization problem to handle the cases where the ε precision assumption is neglected. C>0 is the trade-off between putting the focus on minimizing $\|w\|$ or tolerating the deviations bigger than ε. The primal problem is converted to a dual optimization problem by the use of Lagrange function. The solution to the optimization problem has a saddle (minimax) point. By taking the partial derivatives of the parameters and equalling them to zero on the saddle point, a basis is achieved which helps finding the optimal w^T, b.

2.4 Previous Experiments of Music Information Retrieval and classification of Persian Music

Some of the previous musical information retrieval experiments for the case of Persian music can be found in (Abdoli, 2011; Dārābi, Azimi, & Nojumi, 2006; Peyman; Heydarian & Reiss, 2005; Peyman Heydarian & Reiss, 2007; Heydarian, 2016; Lāyegh, Haghipour, & Sarem, 2013). In (Peyman; Heydarian & Reiss, 2005) the pitch frequencies for an 11-bridged Santur were obtained. In (Peyman Heydarian & Reiss, 2007) the Empirical Mode Decomposition technique was used to extract features from long-term rhythmic structures. In the suggested approach, the signal can be decomposed to a collection of oscillations which can be later applied for further feature extraction applications.

In (Dārābi et al., 2006) Darabi and his colleagues, performed an experiment for recognition of Dastgāh Persian music with detecting skeletal melodic models. They used statistical measures on the Fast Fourier Transform spectrum of music pieces to obtain the prevalent mode and Shāhed tones of the pieces.

Abdoli in (Abdoli, 2011) suggested a model for classifying Iranian traditional musical Dastgāh based on Fuzzy logic. The Fuzzy logic was responsible for modelling the uncertainty of tuning the scale steps of each Dastgāh. He used scale steps for each of the Dastgāhs which were later weighted using fuzzy similarity and distance measures. His dataset consisted of 210 tracks from different Dastgāh vocal pieces.

Layegh in (Lāyegh et al., 2013) performed a classification of the Radif of Mîrzā Abdollāh Farāhāni using support vector machines. The used database consisted of 1250 samples played using Tar or Sitar instruments. The extracted features mainly included frequency domain attributes such as: pitch; mean and standard deviation of spectral centroid; inharmonicity; and Mel frequency Cepstral coefficients.

In (Heydarian, 2016) Heydarian provided a recorded database using different musical instruments (most of the samples were played on a Santur; either 11 bridged or 12 bridged). He used Gaussian Mixture models to estimate the modes of different Dastgāh musical pieces. The features included were spectrograms, chromagrams, and pitch histograms as popular attributes for mode recognition. He suggested some parameterization for the classification of Persian music intervals, which gave better performance.

In (Beigzadeh & Koochesfahani, 2016) they employed a multi-layer perceptron for classifying Dastgāh music played with different musical instruments including Ney (similar to Flute), Violin, as well as vocal Dastgāh pieces. The extracted features included the top twenty peaks from the frequency spectrum of each of the musical pieces. Some older experiments of classifying Dastgāh music using neural networks can be found in (Hajimolāhoseini, Amirfattāhi, & Zekri, 2012; Mahmoodan & Banooshi, 2012).

As have been discussed in this section, the previous experiments for the case of Persian music focus on feature extraction and classification of Persian music. The previous research does not consist of algorithmic music generation studies. In this thesis, there will be an in-depth study on the possibility of creating Persian Dastgāh-like music using computational intelligence tools. This makes the work in the thesis a novel contribution.

2.5 Chapter Summary

In this chapter, the technical fundamentals which are used for the modelling and implementations in the thesis have been introduced. The current chapter consists of three main bodies: musical instrument sound synthesis, music data mining, and a section dedicated to some machine learning tools. The signal processing techniques are employed for synthesizing the sound of a musical instrument in the context of Liquid Persian Music software. The music data mining approaches are applied for extracting features from the Persian musical pieces. These features are later employed for training a support vector regression model. The SVR model is applied as the fitness function in evolving the sequences of voices of the Liquid Persian music audio generator as discussed in chapter 4.

In the next chapter an overview of creativity and algorithmic music composition is given, covering cellular automata, genetic algorithm fundamentals and their applications in computer music. This also includes the background on Boltzmann machine families to be employed as stochastic tools for generating audio and extracting features from Persian music and cellular automata progressions.

Chapter 3. Creativity and Algorithmic Composition

The first part of this chapter studies Cellular automata (CA) as a computational intelligence tool in musical applications. The most important properties of one-dimensional CA rule space and the basins of attraction concept are discussed. CA can generate creative patterns. This vast possibility allows CA to be considered as a source for generative behaviours in different areas of computational creativity and computer music. CA are the main core of pattern generation in the proposed compositional machine in this thesis. The second part of this chapter revolves around genetic algorithms (GA) as a member of the family of evolutionary algorithms. They have important applications in algorithmic composition, since they can be applied as a navigational tool to explore musical space.

The remainder of the chapter describes Boltzmann Machine families and their usage in generating musical artefacts as deployed for the experiments in chapter 6. Boltzmann machines as a type of associative memories are described. Restricted Boltzmann Machines (RBM), Deep Boltzmann Machines (DBM), Conditional Restricted Boltzmann Machines (CRBM), and Multimodal Deep Boltzmann Machines (Multimodal DBM) are the models which are employed in this thesis. Each of them has their specific benefits in the musical application in chapter 6.

3.1 Creativity and Synthetic Music

Creativity can be considered as a universal principle. In this work, we are talking about more specific types of creativity. As has been mentioned in chapter 1, three types of creativity have been identified: combinational, exploratory, and transformational. In combinational creativity new links are discovered and established between elements which are not directly related to each other. In exploratory creativity, a conceptual space is navigated in the hope for finding forms within the predefined constraints. Transformational creativity alters one or more of the governing dimensions of the forms yet to be created. Transformational creativity can happen during the navigation in the conceptual space.

In algorithmic composition, these types of creativity are hoped to manifest themselves through the application of various tools and techniques. Some of the tools establish human domain knowledge in music as a guideline for generating musical material. This knowledge, such as music theory, are often embedded in the algorithmic composition system. In another perspective the music information are retrieved from musical pieces databases and applied with various techniques to the system. There are also computational intelligence tools which are capable of generating creative materials without the contribution of human domain knowledge. One of the targets of algorithmic composition system is generating music within the pre-specified musical styles, or achieving styles and musical pieces which has not been encountered throughout history.

Combinational creativity is a type of creativity which is very difficult for an algorithm to achieve, yet it is easier for humans to find associations between matters that are not directly connected to each other. Samples of this can be found in collage painting. Music is an arena which has various governing dimensions. Interpreting the various musical motives within a musical piece needs more complicated analysis. Likewise, enabling a computer to find the hidden associations between existing musical pieces seems to be a labour intensive research. Achieving this type of creativity for computational music seems difficult to obtain at the moment. Exploratory creativity is a type of creativity which is more easily achievable by computers. Evolutionary algorithms are machine learning tools which enable this type of exploration in the space of all possible musical forms. This kind of navigation needs to be constrained by the usage of musical critics in order to get admissible results. Transformational creativity are easier for computers to perform. Achieving this type of creativity within a musical context in an algorithmic composition system seems to be a complex task. Such a system should be able to go beyond the existing musical rules and current expectations in a way that the outcomes would be the result of discovery of new possible domains for music.

Enabling a computer program to perform an act of creativity might not be that easy. The more advancement achieved in the arena of understanding human creativity, the better the clarifications would be towards the required steps for building a machine which can be creative in a stand-alone manner. In this chapter, a computational intelligence tool (cellular automata) together with genetic algorithms and Boltzmann machines are described which are employed in this thesis to enable the manifestation of creativity within our domain of synthesized Persian music.

3.2 Cellular Automata and Music

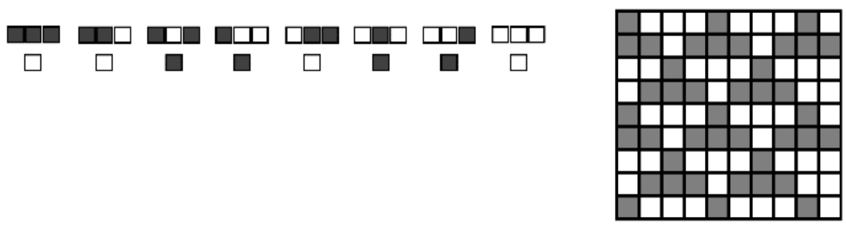
The advent of cellular automata goes back to 1940s. CA was studied as a specific discrete dynamical system in 1960s (Peragine, 2013). Stephan Wolfram started a methodical study of CA from 1980s, his research contributed to various areas considering CA (Weisstein, 2016). According to Wolfram, CA is not only of particular significance in computer science. Wolfram emphasized that CA can be applied in other domains such as physics and chemistry to represent some problems. (Wolfram, 2002).

Cellular Automata consists of simple regular elements, with identical configurations. Each cell has *k* finite states at time *t*, and all the cells evolve simultaneously. The time-space evolution of a one dimensional CA can be represented by a two dimensional mesh (Beyls, 1988; Burraston & Edmonds, 2005; Wolfram, 2002) .

(a) (b)

Figure 3-1. Von Neumann (a), Moore (b) CA neighbourhoods

An elementary CA consists of an assembly of n cells arranged in a one-dimensional array which can be represented by $A^t = \{x_1, \dots, x_n\}$. Each cell takes s pre-specified states: $x_i^t \in \{0,1, \dots, s\}$. In each time step, the colour (state) of the cell is determined by the colours of the same cell in the previous iteration together with the states of its adjacent left and right neighbours. The states of neighbourhood of cell i at time t can be depicted as $NH_i^t \in \{0,1, \dots, s\}^{2r+1}$, where r is the neighbourhood radius. Possible specifications for NH_i are s^{2r+1}. The sum of all CA rules can be represented as $s^{s^{2r+1}}$.A more specific generalization considering the updating of the state of a cell is computed as s^n. The possible number of transition rules are s^{s^n}, where s is the number of possible states for a cell, and n is the number of cells involved in determining the final state of a cell in each time step (Wolfram, 2002). Two of the most popular neighbourhoods are Neumann and Moore's neighbours shown in figure 3-1. If the n cells in the neighbourhood vicinity have binary values, then there would exist 2^n mixture of cells as specifications. For such a system there would exist 2^{2^n} number of rules. The rules determine the quality and nature of the CA dynamics. Each of the rule numbers (from 0 to 2^{2^n}) are represented in their binary format. Each bit from the binary representation of the rule number are arranged from low bit to high bit. Each of the possible neighbourhood configurations (from 0 to $2^n - 1$) are also arranged in order and are assigned to their associated bits in the binary rule number.



(a) (b)

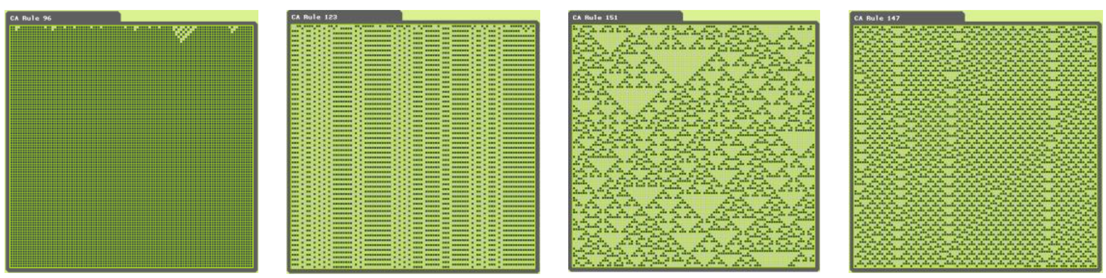
Figure 3-2. Transition rule and evolution for CA rule 54.

(a) Transition rule table for rule 54 (b) An elementary cellular automata progression in eight iterations after applying rule 54. The random intitial state in this example is : 1-0-0-0-1-0-0-0-1-0

In the case of elementary CA, the initial state consists of a series of cells each depicting a binary state (0 or 1). This one-dimensional sequence will generate a two-dimensional presentation during CA progression. Each cell is in one of the two finite states at time t, and all the cells evolve simultaneously. The state of a cell at time *t* depends on its state and its neighbours' states at time $t-1$. In the one-dimensional elementary CA (which is the subject of this study), the permutations of each cell with its two adjacent left and right neighbours defines $2^{2*1+1} = 8$ specifications. Once allocated to binary states, the selection of one of the

256 local transition rules specify the CA progression (Wolfram, 2002). For instance, the binary representation of 54 as a rule number is displayed as 00110110 (bits from 0-7). The transition rule table for rule number 54 is obtained as in figure 3-2:a. A simple mapping is performed between the cell neighbourhood configuration and the binary representation of the rule number. The cell's state, and its left and right neighbours' states determine the cell state in the next time step. Starting from a random initial state, the CA progression for 8 iterations are achieved as depicted in figure 3-2:b.

3.2.1 Rule-space Classification



(a) (b) (c) (d)

Figure 3-3. Four Wolfram classes.

Class 1: invariant, e.g.: rule 96 (a), class 2: cyclic/periodic, e.g.: rule 123 (b), class 3: chaotic, e.g.: rule 151 (c), class 4: complex, e.g.: rule 147 (d) behaviours.

Different studies have been performed for investigating the dynamical behaviour of CA. The ones conducted by Wolfram, suggest four classifications of the applying rules on the pattern propagation of CA (Burraston, Edmonds, Livingstone, & Miranda, 2004; Wolfram, 2002). Much of this classification from plain uniform repetitive structures to patterns with nearly no regularity is done by visual inspection. According to Wolfram, an approximate fourteen percent of CA reveals complicated sort of patterns.

Wolfram classifications on CA emergent patterns can be studied in four categories as (The behaviour of four Wolfram classes are illustrated in figure 3-3):

1. Fixed (or invariant): Patterns become permanent or annihilate each other and disappear: resulting to a homogeneous state.

2. Cyclic: patterns evolve to repetitive fixed structures; oscillating between steady combinations forever.

3. Chaotic: pseudo-random patterns emerge, which do never alternate. Any stationary patterns are perished rapidly; they never find the chance to survive.

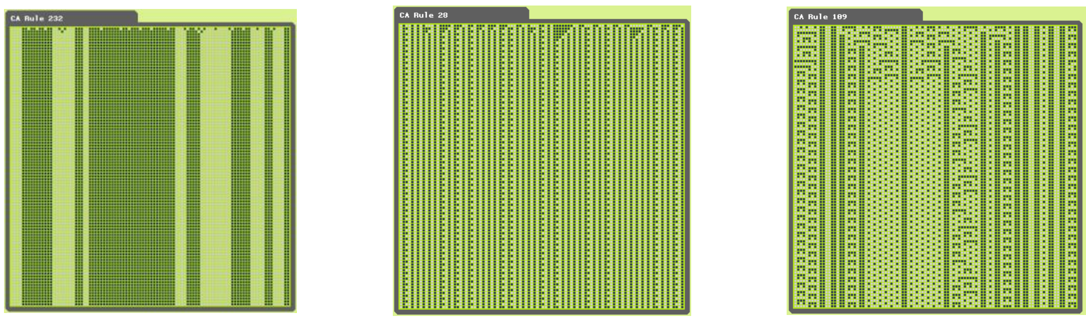
4. Complex: the yielded patterns are in complex formation which may or may not maintain their self-similarity in successive steps. These complicated assembly can be in the form of fractal or nested patterns (Wolfram, 2002), distributed both spatially and temporally (Burraston et al., 2004; Wolfram, 2002). This class is often named as the "edge of chaos" (Langton, 1990); the boundary between classes 1, 2 and class 3; consisting of both ordered and irregular organizations (Wolfram, 2002).

Table 3-1. The one-dimensional elementary cellular automata Wolfram classes percentages.

Class	Percentage	Rules examples
1	9.3 %	0,8,32,40,128,136,160,168
2	76.1 %	1,2,3,4,5,6,7,9,10,11,12,13,14,15,19,23,24,25,26,27,28,29,33, 34,35,36,37,38,41,42,43,44,46,50,51,56,57,58,62,72,74,76,77, 78,94,104,108,130,132,134,138,140,142,152,154,156,162,164, 170,172,178,184,2,204,232
3	11.7 %	18,22,30,45,60,73,90,105,106,122,126,146,150
4	2.3 %	54,110

The percentages of rule numbers belonging to each of the four Wolfram classes are depicted in table 3-1. For each of the classes a set of rule numbers have been given as examples. Most of the CA rules belong to the second class of behaviour. Class 4 has the least number of rules belonging to it.

Li and Packard (Li; Packard, 1990) propose five classes of behaviour. This kind of classification were determined based on rules with periodic and non-periodic dynamics, and more specifically whether they have short or long periodicity.



(a) (b) (c)

Figure 3-4. Examples of second, third, and fourth Li-Packard classes .

Fixed point, e.g.: rule 232(a), periodic, e.g.: rule 28 (b), locally chaotic, e.g.: rule 109(c).

The subdivision of the CA behaviours as proposed in (Li; Packard, 1990) is reported as:

1. The first class consists of rules with homogeneous patterns. This class is the same as Wolfram first class and is called Null rules class.

2. The Li-Packard second class consists of cellular automata rules which yield heterogeneous patterns.

3. The third class contains CA rules with periodic patterns with intervals greater than unity.

4. The rules in the fourth class have locally chaotic behaviour which are entrapped between fixed walls.

5. The fifth class consists of global chaotic rules. The CA progressions have random looking spatiotemporal patterns, and/or the periodicity cycles have highly divergent lengths. The fifth class is the combination of Wolfram's third and fourth classes.

The second, third, and fourth Li-Packard classes are subdivisions of Wolfram's second class. Different behaviours of the second, third, and fourth Li-Packard classes are illustrated in figure 3-4.

Langton introduced Lambda parameter (Langton, 1990) as a tuner for moving between the classes of CA rule space. He investigated the characteristics of various CA behavioural dynamics by reflecting those to the state-space of attractor basins, which helps recognizing the evolution of CA in terms of stationary, oscillating, chaotic, or complex realms. Langton's Lambda parameter would become an important criteria in determining the behaviour of the system in more complicated CA configurations (Burraston & Edmonds, 2005).

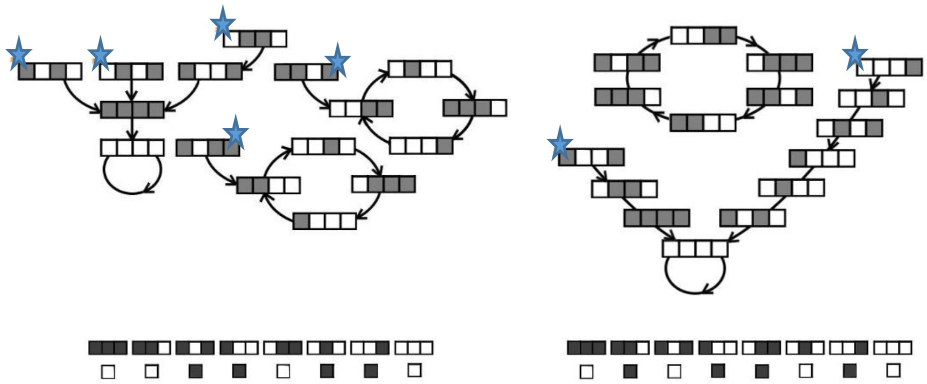
Wolfram conducted a series of experiments on the level of complexity of the underlying rules and the emergent behaviour of CA. What he found out was that increasing the complexity of the initial rules and conditions does not necessarily contribute to more complex behaviour and does not essentially add more characteristic properties to CA (Wolfram, 2002).

An important property of CA is that the rules can be categorized into classes with essentially the same dynamic behaviours (Li; Packard, 1990; Powley, 2009; Wolfram, 1986, 2002). Equivalent rules can be obtained by application of reflection, conjugate, and both transformation operators (shown in formula 9). One-dimensional elementary CA has 256 rules. 88 distinct behaviours can be obtained by enumerating the essential dynamics using formula 9 (Table B-7 in Appendix B depicts the equivalent CA rule numbers in elementary one-dimensional CA). The representatives of each of these 88 behaviour classes are the equivalent rule with the lowest values (Powley, 2009). The transformation rules for the elementary CA rules have the following forms (Li; Packard, 1990; Powley, 2009; Wolfram, 1986, 2002):

$$\begin{cases} \text{Reflection} & f(a,b,c) = f(c,b,a) \\ \text{Conjugate} & f(a,b,c) = 1 - f(1-a, 1-b, 1-c) \\ \text{Conjugate and Reflection} & f(a,b,c) = 1 - f(1-c, 1-b, 1-a) \end{cases} \tag{9}$$

3.2.2 Basin of Attraction in Cellular Automata

The dynamics of CA can be better illustrated in structures known as basin of attraction (Wuensche, 1999, 2004; Wuensche & Lesser, 1992). In order to clarify this notation, the meaning of state space is briefly overviewed. A state space consists of all possible configurations or patterns. For instance, a binary vector of size N has 2^N patterns, and a binary matrix of size 4*4 has 2^{16} different configurations. For a CA rule number, this state space can be divided to sections with each section having a structure relating its constituent patterns. These structures are known as basins of attraction. They often consist of a central point or a collection of patterns arranged on a circular path. There might be branches with so many sub-branches connecting to these figures. The links between patterns in a basin of attraction is formed according to the accessibility by previous observed patterns. These outer branches are pre-images of the inner ones. The leaves are Garden of Eden states, which are not accessible by any previous states (Wuensche, 1999, 2004; Wuensche & Lesser, 1992). Two examples of basin of attraction models (Arshi & Davis, 2017) are depicted in figure 3-5.



(a) (b)

Figure 3-5.Examples of Basin of Attraction model .

Two basin of attractions for Cellular automata configuration for (a) rule 54, and (b) rule 90. The star signs shows the Garden of Eden states. These graphs are manually obtained, however, there are algorithms available for obtaining basin of attraction models for any CA rule space with different configurations (Wuensche, 2009).

3.2.3 Cellular Automata in Music Composition Systems

Cellular Automata are discreet dynamical systems. They have a global behaviour which is influenced by the local behaviour of identical elements. These reciprocal components are the cells in the lattice of one-dimensional elementary CA. Cellular automata demonstrate various genres of behaviour. This specific feature has brought CA into the attention of artists as a creative tool. By changing the configurations of the neighbourhoods and/or increasing the number of possible states for each of the cells, the state space would expand exponentially. The normal life span of a human is not adequate for navigating through all the generated patterns. Therefore CA have emergent behaviour and been looked as a source of creative material for artists. They have been employed for MIDI sequencing, structuring the compositions, and sound synthesis (Burraston & Edmonds, 2005).

Beyls' Cellular Automata Explorer, and CAM developed by Millen are two of the early models of musical CA. Beyls investigated the application of time-dependent rules. He included the neighbourhood states from previous and future time steps from CA progression. Beyls investigated broad criteria of configurations for CA rules, and cell neighbourhood (Burraston et al., 2004). Cellular Automata Explorer targets the formation of complex musical patterns in the output (Fernández & Vico, 2013). Dale Millen exploited two and three-dimensional game of life cellular automata and projected the results to the melody structure. He later navigated the formation of musical organization from CAM (Burraston et al., 2004).

CAMUS and Chaosynth are two other famous CA music generation systems (Miranda, 2001, 2002). CAMUS employed Game of Life and Demon Cyclic Space. A Cartesian space is applied for mapping the patterns to MIDI domain for obtaining the musical triplets. The propagation of the musical patterns in CAMUS are inspired by the similar effect happening during CA progression (Miranda, 2002). Chaosynth is another CA sound generator which is based on the model of chemical reactions of a catalyst. The sound production process in Chaosynth is based on additive synthesis technique. The sound granules generated by the system are added together to obtain the results. The produced tones do not often resemble the acoustic sounds found in the musical instruments. They are often reminiscent of the natural sounds flow as well as the sound of waterfalls, or insects swarms (Miranda, 2001). The interested reader is referenced to (Burraston et al., 2004; Fernández & Vico, 2013) for a thorough review on previous research on the application of CA in generating electronic music.

CA as computational intelligence tools are usually accompanied by other artificial intelligence models to produce hybrid music composition systems. In isolation CA do not presently produce melodic sounds. CA can be embedded in assisted composition systems used as a source of creativity for helping musicians (Fernández & Vico, 2013). Patterns generated by CA have also been used independent of any frameworks and as raw material for inspiring artists. Xenakis is one of the pioneers who used CA for structuring his composition ("Iannis Xenakis webpage".). Xenakis (Georgaki, Solomos, Zervos, & Proceedings, 2006; Hoffmann, 2002; Solomos, 2005) also stated that the generated sounds may need heavy editing by the composer to conform to results being musically pleasant (Hoffmann, 2002; Lo, 2012). Similar issue have been reported by Miranda, the creator of CAMUS, who considers the results as not being very musical (Miranda, 2007).

3.3 Genetic Algorithm and its Applications in Music Composition

Genetic Algorithms (GA) are a type of Evolutionary Algorithms (Goldberg, 1989). Genetic algorithms are inspired by natural selection which was first discovered by Darwin (Darwin, 1906). Natural selection guarantees the survival of the most competent (fittest) genes. GA have applications in various areas for finding optimal solutions to problems. Previous applications of GAs demonstrate their success in problem solving for domains with widespread solution spaces (Buckles & Petry, 1992). There are infinite possible musical combinations and structures. Genetic algorithms are good candidates for exploring the music space and be applied in music composition domain. The fitness function guides the GA exploration in the search space and constrain the musical productions. For instance one can tailor fitness functions which fulfil musical aesthetical aspects or adhere to certain musical tastes or styles (Burton & Vladimirova, 1999).

The classic genetic algorithm pursues the following structure: An initial population of individuals are randomly generated in a mating pool. The individuals are coded as genotypes and are progressively evolved in each of the consecutive generations. Each of the individuals inside the mating pool are potentially a solution candidate. The individuals are rated according to the level of their conformation to the fitness criteria. The fitness function task in the reproduction process is to evaluate the solution candidates. The fittest individuals in the population are selected as parents for breeding. The parents undergo crossover and mutation operations. In crossover, individual parents are selected and their genes are transmitted to each other by swapping (mostly in a meaningful manner). The mutation operator involves the changing of a random gene in the genotype (Goldberg, 1989). The mutation operation takes place with a low probability and is designed to avoid the search being trapped in local solution spaces. There are also elites in the population who are the fittest individuals in the population. Elites are transferred to the next generation without alteration. Their good genes might be contributed in the evolutionary process for elevating the quality of the individuals in the population. This can appear as the increase in the mean fitness of the population.

The algorithm continues until a pre-specified criterion has been satisfied. Another possible way for stopping the algorithm is to limit the time spent for execution or to constrain the number of generations (Burton & Vladimirova, 1999). The parameterization of GA is often performed in a way that raises the expectancy of convergence to optimal solutions.

3.3.1 Genetic Algorithms in Algorithmic Music Composition

Genetic algorithms have been widely applied for composing melodies, and harmonizing pre-specified melodies. Genetic algorithms were applied independently or as hybrid models accompanying various self-governing artificial intelligence methods as well as knowledge-based models, hidden Markov models, and artificial neural networks.

Horner and Goldberg (Horner & Goldberg, 1991) were pioneers who presented the application of genetic algorithms in algorithmic composition. Thematic bridging is a composition methodology; starting from an initial pattern, the system undergoes a series of transformations to manifest the final target pattern. In the GA system proposed by Horner and Goldberg the individuals are the transformation operators. The fitness function calculates the distance between the individuals and the target pattern. The output of this system were the sequences of the generated patterns.

The fitness function can be interactive or autonomous. Interactive fitness functions work based on a human user assessment on the competence of the candidate individuals in the population. The existence of some dynamics as well as musical expressions make them seem to rely more on human intelligence, which allows a composed musical piece to be rated more as human-like rather than machine-like. Adding fitness functions which benefit from human scrutinized evaluations may result in outcomes which are more similar to artefacts generated by human or at least cover the areas which are very difficult to implement using a machine. However, there is a bottleneck problem often associated with interactive fitness functions. Evaluating a high number of individuals in consecutive generations often contribute to user fatigue and should be used in domains where autonomous fitness functions are unable to gain the desired results. The autonomous types of fitness functions are based on machine learning tools. In the following, some examples of both types of fitness functions are described.

Jacob (Jacob, 1995) proposed a composition system with three phase modules: the *Ear*, the *Composer* and the *Arranger*. The *Ear* and *Arranger* are modules where their qualities and nature are determined by human user. The *Ear* was trained by the user and acts as an assessment module in the process of creating musical motifs according to authorized intervallic combinations. The *Composer* is the core of music generation system. The *Arranger* reorders and assembles the output with the structure of musical phrases.

In GenJam (Biles, 2014), Biles devised an evolutionary algorithm for generating Jazz melodies. The initial version of GenJam used to have an interactive fitness function. Biles later applied an artificial neural network to automate the evaluation task and overcome the interactive fitness function bottleneck. The artificial neural networks did not demonstrate success in evaluating new cases. In fact the trained artificial neural network failed to expand the assessments to cases other than what have already been specified in their training dataset (Fernández & Vico, 2013).

The fitness function in an evolutionary algorithm agenda can be designed based on some standard evaluation criteria for creativity (which are discussed in chapter 7). On this account, human-evaluation on the levels of creativity of the system can be embedded as part of the design of the GA system for criticising the iteration of artefacts during the production process. Jordanous in (Jordanous, 2010) designed a system to evolve creative entities based on their levels of creativity. The system employed a GA to evolve the musical parameters for algorithmic music. The system relied heavily on human interventions since it worked based on Ritchie’s empirical criteria for creativity (Ritchie, 2007).

Some of the designed genetic algorithms with autonomous fitness functions for algorithmic composition are presented in the following:

One of the simple approaches for designing fitness functions is the calculation of the weighted sum of distances to a target melody. This method needs a strong selection of musical features for reaching satisfactory results (Dahlstedt, 2007; Laine & Kuuskankare, 1994).

In a series of applications, neural networks were used as fitness functions. Neurogen (Gibson & Byrne, 1991) is an algorithmic music composition system which possesses two different artificial neural networks as fitness functions. One of the neural networks evaluates the intervals between pitches, and the other one is used for assessing the overall musical structure.

Manaris and his colleagues proposed one of the successful hybrid models consisting of neural networks and genetic algorithms (Manaris, Romero, et al., 2005). Manaris et al. trained neural networks as a fitness function trained with Zipfian metrics to identify individual compositions which contributed to the target Zipfian distribution property.

In (Lo, 2012; Lo & Lucas, 2006), N-gram models were applied as trainable fitness functions in a series of experiments where Zipf's law, and information entropy were used as musical aesthetics measurements. An N-gram classifier was trained with musical samples consisting of three consecutive notes (i.e. N =3). The designed fitness function was employed for assessing the sequences of pitches. The genetic operators worked as tools for navigating the search space. Later in the same project, evolutionary algorithms evolved CA progressions in a music generator system. In the survey held by Lo, on average, the human compositions were preferred to machine generations and only one piece was recognized to be indistinguishable from human composition.

3.4 Boltzmann Machines and their Generative Nature

Deep learning is based on a collection of machine learning tools which are sourced from neural networks terminology and structures. The general trait in deep architectures is that they often consist of a hierarchy of layers which are stacked on top of each other. This type of stacked architecture serves the purpose of achieving several representations of data in various layers. Some instances of deep learning architecture are Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), Convolutional Neural Networks, Deep Boltzmann Machines (DBM), and Deep Belief Networks (DBN).

DBM are capable of modelling the probability distribution of the input data. These new representations of the observable data, are accumulated in the intra-layer weights. Consequently, these networks are widely used in feature extraction, dimensionality reduction, and classification applications. Some examples of these kinds are available in (Gehler, Holub, & Welling, 2006; Larochelle & Bengio, 2008; Salakhutdinov & Hinton, 2009b). For instance, once this network is trained, the features which are now embedded in weights in the DBM network can be used as data for further machine learning applications (Hinton & Salakhutdinov, 2006; Salakhutdinov & Hinton, 2012).

This section provides fundamentals about some Boltzmann machine families including DBM, Multimodal Deep Boltzmann Machines (Multimodal DBM), and Conditional Restricted Boltzmann Machines (CRBM), as they are going to be utilized in chapter six of this thesis.

3.4.1 Boltzmann Machines and Restricted Boltzmann Machines

Boltzmann Machines (BM) are stochastic networks invented by Hinton and Sejnowski (Ackley, Hinton, & Sejnowski, 1985). Boltzmann machines can be considered as extended versions of Hopfield networks. Hopfield networks are a type of associative memory with the storage capability (Rojas, 1996). Hopfield networks' underlying memorization process is known to resemble that of human. For example, memorizing names according to people's faces features is a sample of associative memories (Freeman & Skapura, 1991). BM have stochastic decision rules for determining the states of units opposed to binary decision rules as in Hopfield networks. The dynamics of Boltzmann machines are governed by an energy concept. The global energy function in BM is similar to its counterpart in Hopfield networks as denoted in equation 10.

$$E = -\sum_{m<n} s_m s_n w_{mn} - \sum_{m} a_m s_m \tag{10}$$

where s_m, s_n are states of units m, and n respectively. w_{mn} is the weight between units m and n. a_m is the bias connected to the m^{th} unit. The states of the units are denoted as $\mathrm{s} = \{0,1\}$. The $m < n$ condition is for assuring that the states are not counted twice in the formula. The minus sign before each of the terms in the formulation indicates that the minimization of energy is a desirable trait in the training process of Hopfield and Boltzmann machine families (Coughlin & Baran, 1995).

Training a fully connected Boltzmann machine is an intractable process in a network with large number of units. On the other hand, constraining the connectivity of BM units to be solely between visible and hidden neurons would provide a base for inferring tractable training formulas and procedures (Carreira-Perpiñán & Hinton, 2005). Limiting the connectivity between visible-visible, and hidden-hidden neurons, produces a bi-partite graph, where the connections are only between visible and hidden units. Therefore, the units will be grouped in two sections (a bi-partite graph): visible units, and hidden units. Restricted Boltzmann machines (RBM) (Smolensky, 1986) are constrained versions of Boltzmann machines which only retain the pair-wise connections between the neurons in the visible and hidden layers.

RBM are capable of modelling the probability distribution of the presented data in their visible units. Restricted Boltzmann machines build new representations of data and accumulate them in the weights in the bi-partite graph model. RBMs employ hidden units for achieving a model for the distribution of the binary units presented in the visible layer. RBM were used in various applications for extracting features from data, dimensionality reduction and classification (Gehler et al., 2006; Larochelle & Bengio, 2008; Salakhutdinov & Hinton, 2009a). For instance, the new features which are embedded in the weights of the RBM network can be employed for other machine learning classification tasks. Some other examples of this kind can be found in (Hinton & Salakhutdinov, 2006; Salakhutdinov & Hinton, 2012). RBMs have also been employed in music applications (Lauly, 2007).

$$E(v,h) = -(h^T W v + \ a^T v + b^T h) \tag{11}$$

$$E(v,h) = -(\sum_m \sum_n w_{mn} v_m h_n + \sum_m a_m v_m + \sum_n b_n h_n) \tag{12}$$

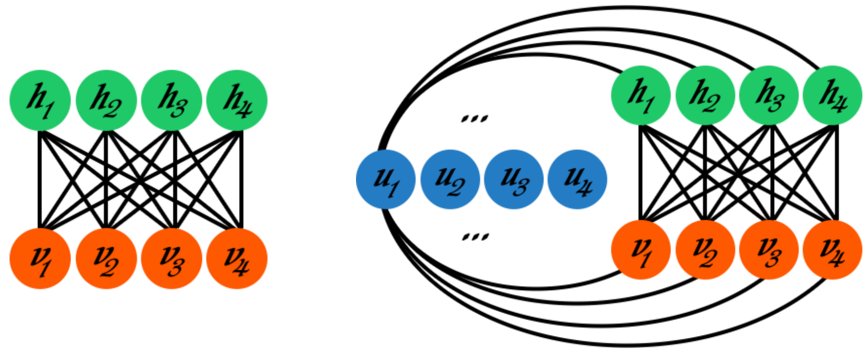
The behaviour of the RBM system is governed by the energy function specified in equations 11 and 12. In which v, and h terms stand for the observable and hidden units. The bias parameters a, and b, respectively contribute to visible and hidden units. The energy function in RBM model is achieved by summing the linear products of the visible units, hidden units and their associated connecting weights. The additional terms in the following formulas stand for the products of the visible and hidden units and their related biases.

The general type of RBM has binary random variables in their hidden and visible layers. Other extensions of RBM provide other types of variables in the visible units as well. For example, the units in the visible units can accept unbounded real number observations to form a Gaussian-Bernoulli RBMs. Multinomial observations or binomial ones are other examples. It is noteworthy that in these extensions, the elements in the hidden layer remain Boolean. The energy function is a hyper-dimensional surface. The energy of the system is minimized on the learning patterns to be stored. Appendix A.1 provides further details for training RBM by the usage of contrastive divergence algorithm.

3.5 Conditional Restricted Boltzmann Machines

Conditional restricted Boltzmann machines are a variety of RBM systems, which are employed for representing the dependability or relationship of data which inherit some sort of inner-attachment (Mnih, Larochelle, & Hinton, 2011). For instance, the data may appear in time-frame sequences. Some successful applications of CRBMs can be traced in simulating human motion (Taylor & Hinton, 2009; Taylor & School, 2012), and pigeon behaviour (Zeiler, Taylor, Troje, & Hinton, 2009).

An architecture of CRBMs is illustrated in figure 3-6 (Arshi & Davis, 2017). In comparison to RBMs, CRBM model encompasses an additional layer which supplies temporal information to hidden, and visible units.



(a) (b)

Figure 3-6 The architectures of (a) RBM, and (b) conditional RBM,

In a RBM there are pairwise connections between the units in the visible layer, and those of hidden layer, in conditional restricted Boltzmann machine there are additional connections provided by (visible) units from previous time steps (u_i in 3-10b).

The visible and hidden layers are conditioned on the data from previous steps in a temporal format. The energy function of the system can be represented as:

$$E(v,h,u) = -v^T Wh - u^T U_{uv} v - u^T U_{uh} h - v^T a - h^T b \qquad (13)$$

In which v,h,u stand for observables, hidden neurons, and the additional layer, respectively. The additional layer is associated to visible units in the previous time steps. The units in the additional layer is represented by u. W are the weights between visible layer, and hidden layer, U_{uv} corresponds to the weights between visible layer and additional layer, and finally, U_{uh} represents the weight matrix between conditional layer, and hidden units. a, b are bias matrixes. The training in the CRBM models are performed by the application of the contrastive divergence algorithm.

3.5.1 Deep Boltzmann Machines

A Deep Boltzmann Machine (DBM) is a multi-layered restricted Boltzmann machine with several hidden layers (Salakhutdinov, 2010; Salakhutdinov & Hinton, 2009a, 2012; Salakhutdinov & Larochelle, 2010). Each of the hidden layers is built on top of the previous hidden layers. An example of a two-layered DBM is presented in figure 3-7.

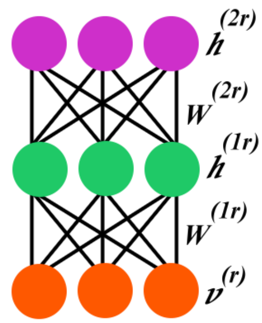


Figure 3-7. Deep Boltzmann machine architecture with two hidden layers.

By having v as a set of units in the visible layer and $H = \{h^{(1)}, h^{(2)}\}$ as a set of units in the first and second hidden layers, the model parameters are expressed as: $\tau = \{W^{(1)}, W^{(2)}, b^{(0)}, b^{(1)}, b^{(2)}\}$. $W^{(1)}, W^{(2)}$ are the weights between visible to first hidden layer and the first hidden layer to the second hidden layers. The $b^{(0)}, b^{(1)}, b^{(2)}$ parameters are the biases of the visible, first hidden and the second hidden layers. The energy of the joint configuration can be obtained as (Salakhutdinov & Hinton, 2009a, 2012):

$$E(v,h;\tau) = -\sum_{l=1}^{N_0}\sum_{m=1}^{N_1} W_{lm}^{(1)} v_l h_m^{(1)} - \sum_{m=1}^{N_1}\sum_{r=1}^{N_2} W_{mn}^{(2)} h_m h_n^{(2)} - \sum_{l=1}^{N_0} b_l^{(0)} v_l - \sum_{m=1}^{N_1} b_m^{(1)} h_m^{(1)} - \sum_{n=1}^{N_2} b_n^{(2)} h_n^{(2)} \tag{14}$$

Appendix A.2 obtains the procedure for obtaining the weights in the above formulations. The algorithm presented for training a DBM is a basis for training the Multi-channelled DBM.

3.5.2 Multimodal Deep Boltzmann Machines

Multimodal Deep Boltzmann Machines (Multimodal DBM) are employed for data which have more than one data modality (Srivastava & Salakhutdinov, 2012). For instance, the video data consists of a sequence of frames accompanied by their related audio signal. The image and audio frames from video are two different data modalities with their specific statistical properties. If a DBM is used for working with such data then the multiple modalities are representable by separate Boltzmann machines.

A Multimodal DBM consists of multiple channels where each channel represents one data modality. Each of the pathways in the model can be trained separately and then be reunited in a conjoint section. In a Multimodal DBM, the intersection part embeds a joint representation of various modalities of data. By clamping the data on one side of the pathway, the related modalities can be retrieved on the other sides of the pathway.

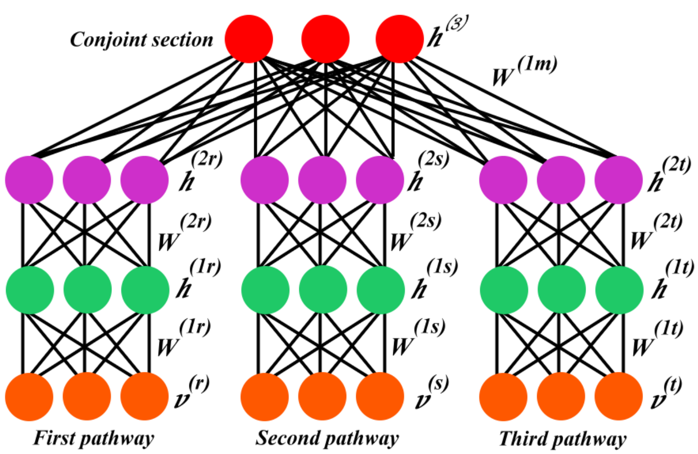


Figure 3-8. Multimodal Deep Boltzmann Machine with three sections.

Figure 3-8 demonstrates the architecture of a Multimodal DBM. This structure consists of three pathways each of which has two hidden layers. The presented Multimodal DBM architecture is going to be employed in chapter six. The accompanied formulations in this chapter are tailored for the Multimodal DBM with three pathways. The original formulation can be found in (Srivastava & Salakhutdinov, 2012) where the Multimodal DBM were introduced. The pathways are tagged with English alphabets, which assist in the presented formulations in this subsection. The set of units in the visible layer and the first and second hidden layers of the first pathway are denoted as: $path1 = \{v^{(r)}, h^{(1r)}, h^{(2r)}\}$. These sets for the other existing pathways in the presented structure are $path2 = \{v^{(s)}, h^{(1s)}, h^{(2s)}\}, path3 = \{v^{(t)}, h^{(1t)}, h^{(2t)}\}$. The model parameters including the weights and biases related to the three pathways are expressed as: τ^r, τ^s, τ^t. The parameters of the conjoint layer is depicted as τ^m. Appendix A.3 obtains the required formulations for designing a training algorithm for a three-channelled Multimodal DBM, which is employed in chapter 6.

3.5.3 Applications of Boltzmann Machine Families in Creating Artefacts

Boltzmann machine families are stochastic generative artificial neural networks. Unlike other types of artificial neural networks, which have deterministic nature, Boltzmann machines have stochastic nature. Moreover, they do not have separate output layer. The input and output layers for Boltzmann machines overlap with each other. In order to achieve an output from this system, one needs to populate the input layer. In the next phase the sampling procedure, takes place which may include bottom-up and top-down passes from several layers (depending on the depth of the architecture in a deep Boltzmann machine structure). The stochastic nature for Boltzmann machine means they are able to obtain different results providing the same input. RBM were applied for generating chords, and producing music.

RBM were first utilized for generating polyphonic music in (Boulanger-Lewandowski, Bengio, & Vincent, 2012; Briot & Pachet, 2017). In (Boulanger_Lewandowski, 2015) sampling technique (like Gibbs sampling) were applied for music generation by the help of RBM with units accepting real values instead of binary numbers. In (Lattner, Grachten, & Widmer, 2016) RBM learned local musical structures. Further global structures were imposed to the music generation process by performing constrained sampling techniques. The constrained sampling technique consisted of gradient descent optimization for imposing the desired global structure and the Gibbs sampling as contrastive divergence training technique. The constraints that Lattner and his colleagues imposed to their generated music were tonality constraints and meter constraints (Briot & Pachet, 2017).

CRBM have had applications in music classifications as well as composition. CRBMs were used for auto tagging music in (Mandel, Eck, & Bengio, 2010; Mandel, Pascanu, Larochelle, & Bengio, 2011). The proposed CRBM model in Mandel's papers were outperforming support vector machines as a machine learning tool. Loeckx and Butheel designed a CRBM structure for improvising music in the desired musical style. The music was reconstructed based on the first musical note provided to the system (Loeckx & Bultheel, 2015). In (Lauly, 2007) CRBM were used for learning melodic sequences and long term dependencies between notes. Their model outperformed n-gram, and artificial neural networks.

3.6 Chapter Summary

This chapter provides information on three of the tools which are employed in this thesis for the design of algorithmic composition systems. CA are a computational intelligence tool. CA offer a diverse range of behavioural patterns as a source of non-human based creative material. They have the potential to directly or indirectly accompany the process of creating artefacts. On the other hand, evolutionary algorithms provide a range of tools for exploratory and transformational creativity. They navigate the space of possible musical solutions, accompanying the chance of new forms to appear. In this thesis, CA and genetic algorithms are combined as a hybrid tool for an algorithmic composition system. The requirement of such system includes targeting proper architectures of the fitness function and the right definition of the criteria for evaluating the emerging forms. The next chapter investigates a system which makes this idea possible. Chapter 5 proposes a design for the fitness function based on an aesthetical criteria (Zipfian metrics) to evolve the productions from a CA based system.

Boltzmann machines are extended versions of Hopfield networks. They have stochastic nature. This attribute makes them desirable as a generative machine learning tool for producing artefacts. Moreover, Boltzmann machine families have the capability of storing data and obtaining new representations of the data. In this thesis, Boltzmann machine families are employed as a tool for extracting features from CA progression. Chapter 6 provides further means for extracting patterns from CA iteration using Boltzmann machine families. Later in chapter 6, multimodal deep Boltzmann machines are used for associating Zipfian features from CA and Persian music. This system is also used for generating audio samples.

Chapter 4. Methodology for producing Synthetic Persian Music

In this chapter, the fundamental ideas behind the main designs of the software experiments in this thesis are explored. The experiments in this chapter are important in the flow of the thesis since they illuminate the direction as further explored in the next two chapters. The heart of this chapter revolves around analysing some aesthetical aspects of patterns extracted from Cellular Automata (CA) progressions. Later in the chapter, patterns extracted from CA are employed for producing musical forms. Liquid Brain Music (LBM) system as a cellular automata based audio generator is presented and later developed to Liquid Persian Music (LPM). Once the outputs of LPM system are considered as musical motives (known as voices in this chapter), the concern would be to produce musically meaningful sequences of the voices. An initial experiment is proposed for performing competency measures on the output voices from LPM system. The competency criterion considers Zipf's law as an aesthetical measure. Subsequently, an evolutionary framework is suggested for evolving the sequences of LPM voices. The proposed evolutionary framework establishes the foundation for further experiments in chapter 5.

4.1 Liquid Brain Music System

Liquid Brain Music (LBM) is a cellular automata based tool developed at the University of Hull (Turner, 2008; Woods, 2009). LBM explores the idea of artificial life systems in generating audio. The first and second versions of LBM are based on additive sound synthesis (Manning, 1985; Roads, 1978, 1988) for producing the sounds. The LBM software takes advantage of the synthesis toolkit (Cook & Scavone, 2008) for performing tasks such as parameterizing the signals and adding them together for yielding the results. The nature of the sounds generated by the system changes with CA progressions. Moreover, the user can alter the configuration of the system to feed the signals' parameters with different CA rules and pattern-matching rules.

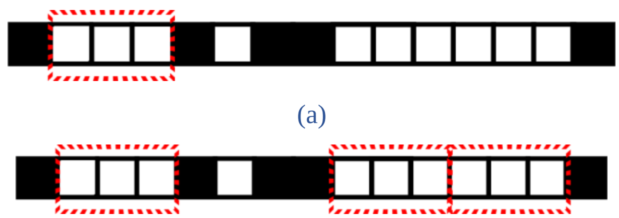
In every time step of the CA progression, the pattern-matching rule obtains values for populating the synthesizer parameters. In one-dimensional elementary CA, the progression in each iteration is representable by an array of white or black cells. The suggested pattern-matching rules in (Turner, 2008; Woods, 2009) work by extracting patterns from each CA iterations or by computing the difference between the specifications of two consecutive iterations. The twenty pattern-matching rules proposed in LBM are presented in table 4-1.

Table 4-1. The twenty pattern-matching rules used in LBM project.

BS stands for block size, and RL stands for run length encoding. The numbers show the size of the white or black blocks. The size of the consecutive white or black cells were chosen to be from 1 to 4.

1. BS_1_White	2. BS_2_White	3. BS_3_White	4. BS_4_White
5. BS_1_Black	6. BS_2_Black	7. BS_3_Black	8. BS_4_Black
9. RL_1_White	10. RL_2_White	11. RL_3_White	12. RL_4_White
13. RL_1_Black	14. RL_2_Black	15. RL_3_Black	16. RL_4_Black
17. Hamming_Difference	18. Jaccard_Similarity	19. Jaccard_Difference	20. Dices_Coefficient

Sixteen of the pattern-matching rules are based on specifications of consecutive white/black cells in only one CA iteration. Block-size and run-length encodings calculate the number of sequences of white or black cells. The lengths of the white/black sequences to be detected by pattern-matching rules were chosen to be $N = 1, 2, 3, 4$ consecutive cells. This would provide sixteen pattern-matching rules based on block-size, and run-length encoding techniques. The difference between block-size, and run-length encoding is the way these two methods behave towards the surrounding cells of the target sequence to be calculated. In block-size encoding for N white/black cells, a block sequence of exact size N will be counted. Therefore, longer sequences of all white or black cells are not broken into smaller chunks of size N consisting of the same-coloured cells. In run-length encoding for N white/black cells, the array is divided into chunks of size N, and the chunks with same-coloured cells are calculated. Figure 4-1 illustrates the difference between the underlying procedures in block-size, and run-length encoding methods.



(a)

(b)

Figure 4-1 Calculating the Block-size run-length encoding.

One 3 sized block is shown in (a), and three chunks of run-length encoding 3 in (b) from pattern-matching rule 3 BS_3_White in (a) and rule 11 RL_3_White in (b)

One of the features associated with the pattern-matching rules based on block-size and run-length encodings is that they calculate the specific number of patterns in just one binary string of cells at a time (one CA progression). If two consecutive strings in the CA progression happen to have the same number of detected patterns, the auditory result would produce the same result after being mapped to the musical space. The remaining pattern-matching rules work based on calculating the difference between the specifications of two consecutive CA iterations. Hamming distance, Jaccard similarity, Jaccard difference, and Dice's coefficient perform binary string comparisons between the consecutive CA iterations. Hamming distance calculates the number of cells with opposite colours in the two strings.

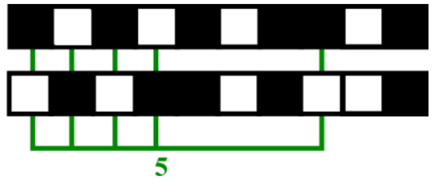


Figure 4-2 Finding Hamming distance between two binary strings.

The values of the cells are compared in the same locations over the 10-bit binary array using pattern-matching rule 17, giving result of 5.

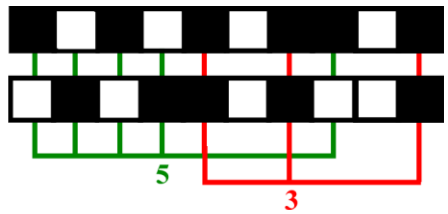


Figure 4-3 Calculating the Jaccard similarity and difference.

In this example two 10-bits binary strings are the subject for finding Jaccard similarity/difference. The Jaccard similarity (pattern-matching rule 18) is obtained as : $\frac{5}{5+3}$, and the Jaccard difference (pattern-matching rule 19) is calculated as $1-\frac{5}{5+3}$.

Figure 4-2 illustrates an example for calculating Hamming distance between the two strings. Jaccard similarity and difference calculate the number of cells which happen to be black in the same locations in both strings (this number is specified by p). Afterwards the number of cells with opposite colours on the two strings in the same array locations are calculated (this number is specified as q). The Jaccard similarity is calculated as $p/(p+q)$. The Jaccard difference is yielded as $1-(p/(p+q))$. Computing the Dice's coefficient is similar to the case of Jaccard similarity and is obtained by $2p/(2p+q)$. Figure 4-3 shows an example for computing Jaccard similarity/difference.

4.2 From Liquid Brain Music to Liquid Persian Music

Liquid Persian music (LPM) is a version of LBM software developed for the research presented in this thesis. The physical model of a stringed musical instrument (Sitar) is responsible for sound production in the early versions of LPM system (Arshi & Davis, 2015). (In a later version of LPM system an alternative synthesiser model, based on the Santur musical instrument, is responsible for generating audio). Some of the synthesizer parameters in LPM include ADSR (Attack, Decay, Sustain, and Release) envelopes, loop gain, and the musical instrument string length for defining the notes frequencies.

Figure 4-4 The Initial Liquid Persian Music user interface.

The defined twenty pattern-matching rules in LBM are also employed in LPM system. (Later experiments looked to replace the pattern-matching rules with Boltzmann machines). The obtained values from the pattern-matching rules are then fed as parameters into the synthesizer for producing the sounds. Configuring the system with a CA rule and a pattern-matching rule (for each of the synthesizer parameters) produces a collection of notes over CA iterations. These collections are referred to as *voices* throughout the thesis. Each of the voices aggregates a collection of notes or sounds (in the case of LBM).

The code related to LBM and LPM are written in C++ and the OpenAL library is responsible for propagating the produced voices. Further information about the software (and audio samples) can be found in (Davis, 2016). Figure 4-4 illustrates the LPM user interface. A manual for working with LPM user interface is presented in the Appendix L.

A one-dimensional elementary CA consists of 256 rules. The musical behaviour derived from one-dimensional (1D) CA does not require examining the 256 rules' behaviours. As discussed in the third chapter, the rule sets with inherently equivalent behaviour can be identified by the application of conjugate, reflection and both transformations together (Powley, 2009). In fact, the rule space of one-dimensional elementary CA can be reduced down to 88 fundamental behaviours (Li; Packard, 1990; Powley, 2009; Wolfram, 2002).

In the remainder of this section, the number of the produced voices by LPM system is calculated. There are $88 * 20$ possible configurations for CA rules and pattern-matching rules. The synthesizer in LPM takes 7 parameters (contributing to pitch and timbre of the notes). The 88 1D CA rule behaviours, 7 defined synthesizer parameters related to pitch and timbre, together with 20 pattern-matching rules expand the number of voices to $88^7 * 20^7$. Considering the number of CA progressions involved, the number of voices would expand to $88^7 * 20^7 * t$, where t is the number of CA iterations. This defines the base auditory search space for the computational framework being developed. By counting the number of other parameters including speed of music (similar to tempo), note durations, note onset times, and note intervals, 4 more parameters will be involved in generating the melodic and rhythmic structures. This extends the search space to $88^{11} * 20^{11} * t$. Considering all the possible initial CA configurations on an array of 100 cells would add a factor of 2^{100} to the search space figure. The constituent elements in the search space are depicted in figure 4-5. In the next chapter, this space is constrained to be $88^{3} * 20^{3} * t$ for reducing the complexity of the system to be developed. Only one of the 7 synthesiser parameters (pitch frequency) plus note duration and note onset times are selected as the varying parameters. Further alterations on this search space are used in Chapter 6, where only pitch frequency and note duration are used.

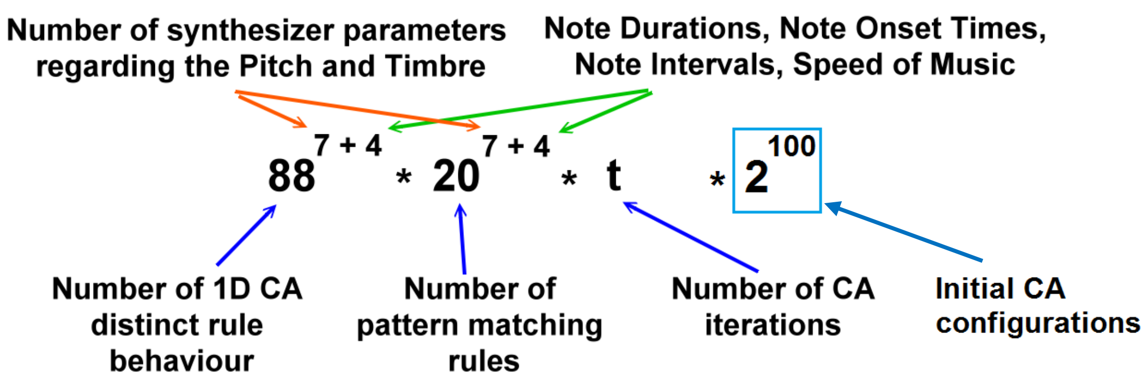


Figure 4-5 Constituent components in the maximal LPM search space.

4.3 Discussion on the LPM Musical Output

After describing the LPM underlying methodology for producing audio, the output results from LPM were analysed in terms of musicality. The musicality of LPM outputs were investigated by the combination of two approaches. In the first approach, the decision of a human subject was taken into account by performing auditory tests. In the second approach, the musicality property was explored through studying the plots of pattern-matching rules outputs through visual investigations. Both the auditory tests and visual tests were performed by the author of this thesis. In the visual tests, the values of 20 pattern-matching rules for 88 CA rules over 10000 iterations were extracted. In this chapter, only one out of 2^{100} possible initial CA configuration was selected. This initial seed for the CA progressions was selected randomly.

Table 4-2 Examples of human evaluation on the musicality of LPM outputs by auditory test.

In this table nine CA behaviours were studied. The equivalent CA behaviours are also demonstrated in this table (obtained by conjugate, reflection and both transforms as described in chapter 3). Ten pattern-matching rules in this table are responsible for extracting features from CA progressions. In the table, cells show the constant (yellow), oscillatory (dark blue), and disordered fluctuations (red) behaviour categories. 'y', and 'n' are abbreviations for 'yes' and 'no' to indicate the musicality or nonmusicality of the related output from CA, and pattern-matching rule. 'l' stands for low levels of musicality.

CA Rule number	Equivalent Rules			Human evaluation											
	Conjugate	Reflection	C, R	PM1	PM2	PM8	PM9	PM10	PM11	PM12	PM17	PM18	PM20	Melodic	Rhythmic
0	255	0	255	n	n	n	n	n	n	n	n	n	n	n	y
2	191	16	247	n	n	n	n	n	n	n	n	n	n	n	y
5	95	5	95	n	n	n	n	n	n	n	n	n	n	l	y
6	159	20	215	n	n	n	n	n	n	n	n	n	n	l	y
11	47	81	117	n	n	n	y	y	y	y	y	y	y	l	y
14	143	84	213	y	y	y	y	y	y	y	y	y	y	l	y
90	165	90	165	y	y	y	y	y	y	y	y	y	y	y	y
105	105	105	105	y	y	y	y	y	y	y	y	y	y	y	y
54	147	54	147	n	n	n	n	n	n	n	n	n	n	n	y

In fact, the derived values would normally be employed for populating the parameters of the synthesizers. The auditory tests were performed by listening to short audio pieces of lengths up to 30 seconds each. In the auditory experiments, the pattern-matching outputs were used for populating the frequency parameters of the synthesizer. The rest of the synthesizer parameters were kept constant. Therefore, the timbres of the notes were left unchanged and only the pitch frequencies of the notes were varied. Table 4-2 demonstrates some examples of result of the human evaluation. The CA rules with equivalent behaviours are shown in this table; meaning that the LPM outputs were not separately analysed for the equivalent rule behaviours. Ten out of twenty pattern-matching rules are shown in this table. The last two columns show the opinion of a human subject on the LPM musical outputs in terms of melody, and rhythm. All the voices were considered to be rhythmic, due to the sound synthesizer changes in each CA time step. Tables B-1 to B-6 in appendix B illustrate the complete tables consisting of 88 CA rules behaviour. The tables in Appendix B were classified based on their constituent four Wolfram classes and Li and Packard (Li; Packard, 1990) extensions on the second class.

The visual investigations on the nature of the LPM outputs took place by studying plots similar to figure 4-6. Figure 4-6 illustrate some examples of specific pattern-matching rules outputs behaviours for CA rule numbers 168, 11, 38, 110, 22, 27, 51 (more examples are shown in Appendix C). In fact, the visual investigations were complementary to auditory tests. The visual investigation of LPM outputs using graphs (as in figure 4-6) enabled us to study the behaviour over larger number of CA iterations while saving time. The graphs in figure 4-6 suggest that for some CA rules, there are oscillations occurring at the beginning of the CA progression before the CA reaches a stable state. Converging to stable states requires more iteration in some of the graphs (e.g. CA rule 110, pattern-matching 5) while this happens very quickly for other rules (e.g. CA rule 11, pattern-matching 2).

The behaviour of CA were studied after a certain number of progressions. After becoming stable, the CA progressions for various rules show different behaviours. Some stay on the same value while others start fluctuating between different values. In the cases where CA progressions converge to a value, the result would be a monotonous audio. The oscillation ranges from two values to more than thirty in some cases. In the auditory tests, these phenomena show themselves as oscillations between two or more pitches.

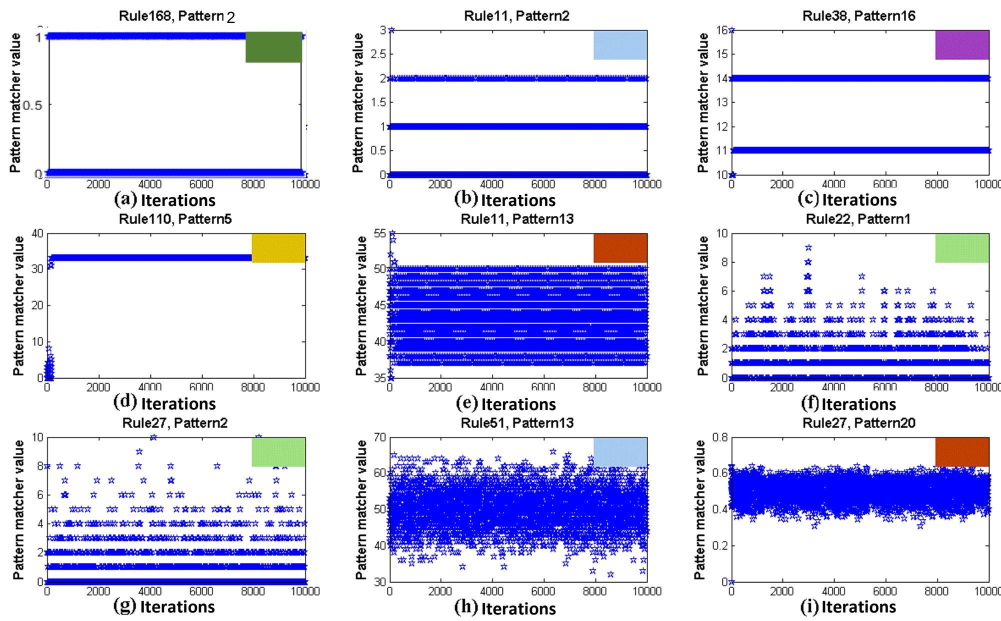


Figure 4-6 Pattern-matching values for CA over 10000 iterations

The nine examples show (a) rule 168 : pattern 4, (b) rule 11 : pattern 2, (c) rule 38 : pattern 16, (d) rule 110 : pattern 5, (e) rule 11 : pattern 13, (f) rule 22 : pattern 1, (g) rule 27: pattern 2, (h) rule 51: pattern 13, (i) rule 27 : pattern 20.

The investigations over LPM outputs determine three groups of LPM audio outputs (table 4-3). Investigating the audio output of the LPM over all the graphs for 88 rules have explicitly shown that the musical behaviour of the designed system can be categorized in the following groups:

- In the first group, after a few number of iterations, the CA progression converges to a homogeneous state, in which no differences can be measured over consecutive iterations. The output values of pattern-matching rules remain constant during the CA progression resulting in a uniform sound.

- The behaviour of the identified second group demonstrates the fluctuation of pitches between two or more frequencies. Oscillations between two or more different voices occur periodically.

Table 4-3 The investigations over LPM outputs determine three groups of behaviour.

The table shows the three groups obtainedby studying the different pattern-matching outputs over CA progressions with one random initial seed. In this table R stands for rule number, and P stands for pattern-matching rule number (Appendix C provides supporting illustrations for this table).

Wolfram Classes	W-class 1	W- class 2			W-class 3	W- class 4
Li-Packard Classes	LP- Class 1	LP- Class 2	LP- Class 3	LP- Class 4	LP- Class 5	
Group 1	R168:P7	R184:P5	R38:P20	R73:P1	R18:P1	R110:P20
Group 2	R168:P1, R168:P9	R184:P2	R38:P8, R11:P9, R51:P3, R27:P1	R73:P12	R22:P1, R146:P6	R110:P1
Group 3	none	None	R27:P20	none	R22:P20, R146:P19,20	none

- The third group has audio in which the frequencies just wander between a diverse range of values. The third pattern of behaviour is observed as a disordered fluctuation between a large number of values. In this case, the previous values might be met; however, fluctuations happen through diverse ranges of values. This means that this behaviour can hardly be considered as a periodic behaviour.

The three groups based on audio and visual analysis of pattern matching over the CA iterations cuts across the Wolfram and Li and Packard groupings. The pattern matching has affected the interpretation of the CA classes of behaviour, and therefore the input to the synthesiser is now independent of the original (4 or 5) CA groupings (as shown in table 4.3). (It should be mentioned that the author of this thesis examined further initial seeds and studied the pattern-matching outputs through visual investigation. The obtained result on the three grouping behaviours persisted for those cases as well.)

4.4 Investigation on LPM output and Zipf's law

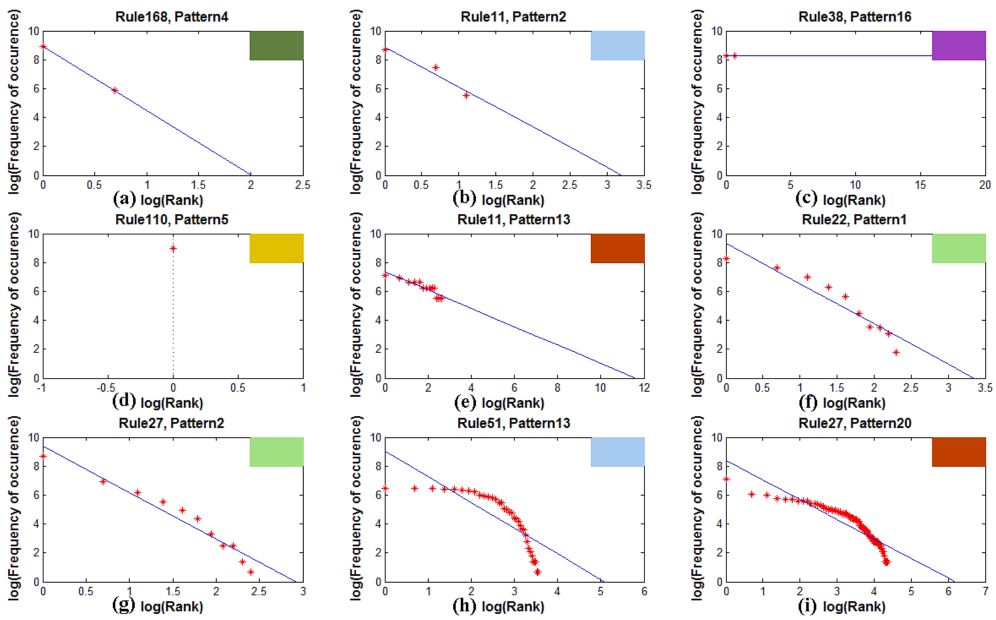


Figure 4-7 Various examples of Zipfian distributions for different CA and Pattern-matching rules.

Zipfian distributions for (a) rule 168 : pattern 4, (b) rule 11 : pattern 2, (c) rule 38 : pattern 16, (d) rule 110 : pattern 5, (e) rule 11 : pattern 13, (f) rule 22 : pattern 1, (g) rule 27: pattern 2, (h) rule 51: pattern 13, (i) rule 27 : pattern 20.

In this section, Zipfian metrics are employed as an aesthetical measurement to study LPM outputs. The values achieved from the pattern-matching rules for the 88 CA rules were applied for investigating the behaviour of LPM in terms of Zipfian distribution. The pattern-matching rules outputs for each of the CA rules were ranked in compliance with their redundancy (a stage in the procedure for obtaining Zipfian slopes). Linear regression was applied on the rank and frequency of occurrence of the pattern-matching rule values. The obtained slopes and R-squared measurements characterize the Zipfian distribution and the precision of the linear regression fit respectively, (the procedure for determining Zipfian slopes are described in chapter 2).

Table 4-4 Zipfian slopes for some CA and pattern-matching rules examples.

The first column on left depicts CA rules and the first row stand for pattern-matching rules (please refer to table 4-5 for colour coding) (this table was originally presented in (Arshi & Davis, 2015)). In this table PM followed by numbers 1 to 9 stands for pattern-matching rule numbers. (The complete table can be found in Appendix D)

CA rule	PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9
168	-2.53	-4.43	-3.12	-2.87	-2.53	-2.53	-Inf	-4.43	-0.61
11	-2.33	-1.21	-2.52	-2.78	-1.94	-2.19	-1.96	-2.00	-0.75
27	-1.54	-3.37	-1.57	-3.15	-1.60	-3.30	-1.61	-3.13	-1.99
38	-3e-4	-1.18	-0.11	-1.74	-1.22	-0.93	-1.36	-1.19	-3e-4
51	-3.30	-3.18	-3.19	-2.84	-3.28	-3.28	-3.34	-3.00	-1.95
22	-2.60	-2.86	-2.96	-3.07	-3.26	-2.94	-3.44	-3.29	-1.64
110	-3e-4	-3e-4	-3e-4	-3e-4	-Inf	-Inf	-Inf	-Inf	-3e-4

The values of the twenty pattern-matching rules were extracted from 10000 iterations of CA progression. The Zipfian distribution characteristics of LPM outputs were studied from the five-hundredth to ten-thousandth CA iterations. This time delay is to let the CA progressions to reach stability after the initial state. Figure 4-6 and 4-7 (and figures in Appendix C) illustrate the linear regression lines fitted to Zipfian data distribution of LPM outputs for specific CA rules. Table 4-4 depicts the obtained Zipfian slopes for a range of CA rules and pattern-matching rules. The subfigures in figure 4-6, and 4-7 have colour-coded labels in their upper left corners. The cells in the table 4-4 are colour coded as well (the colour coding are defined in table 4-5).

The remainder of this section provides an analysis on the colour coding given for table 4-4 (the complete table can be found in appendix D). The colour coding were determined based on a scrutinized comparison over the results from auditory tests together with visual study performed on the LPM outputs, and the Zipfian data. Most of the parameters attained from the Zipfian slopes are in accordance with the investigations performed in the previous section on the musicality of LPM outputs. This means that the Zipfian results are as expected according to the output graphs and auditory tests from the previous section.

Table 4-5 Colour coding interpretations in terms of confusion matrix.

Colour Coding	Colour interpretation	Wolfram CA Classes	Number of occurrences
	FN	2, 3	207
	FP	1, 2-2, 3	30
	TN	1,2,3,4	1031
	TN	1, 2-2, 3	71
	TN	1, 2-2, 2-3, 3, 4	167
	TP	2-2, 3	254

In the remainder of this section, the confusion matrix is built to investigate how the predicted and actual classes overlap with each other. The target classes were yielded by human decision as described in the previous section. The human labelling was performed according to the auditory tests on LPM outputs and by investigating the LPM outputs presented in graphs as in figure 4-6. The LPM outputs were obtained by varying CA rule numbers and pattern-matching rule numbers. The predicted classes were obtained by studying LPM outputs musicality in terms of Zipf's law. The slopes of the linear regressions were used to categorize the musicality of LPM outputs. The slopes which were mostly near to the Zipf's ideal (-1) govern the LPM output as being musical. In this experiment, it was decided that Zipfian slopes between -2.1 and -0.6, are expected to have musical LPM outputs (This range was obtained empirically and for the case of LPM outputs). The colour coding in the tables 4-4, and 4-5 have the following meanings:

- The rusty orange cells show the cases (LPM outputs) where the distribution follows Zipf's law (slopes between -0.6 and -2.1). The auditory tests and studies on the LPM output graphs demonstrated musical LPM outputs for these cases.

- The yellow cells indicate that the LPM outputs have minus infinity slopes in their Zipfian distributions. The auditory tests and studies on the LPM output graphs had also shown non-musical results for these cases.

- The dark green cells stand for monotonous LPM outputs according to Zipfian slopes (slopes less than -2.1 excluding minus infinity). The auditory tests and studies on the LPM output graphs had also demonstrated non-musical LPM outputs for these cases.

- The light green cells depict the situations in which the author found the LPM outputs pleasing to hear and/or the LPM outputs graph demonstrated their musical characteristics. Therefore, the author expected the LPM output to have Zipfian ideal parameters, however, the obtained Zipfian parameters were found to be far from ideal.

- The light blue cells demonstrate the cases where the Zipfian distribution is nearly ideal (slopes approximating to -1 and slopes between -0.6 and -2.1); however, the auditory tests and the study on the LPM output graphs proved to be the contrary.

- The dark purple indicate the LPM outputs having Zipfian slopes near to zero, and according to the auditory tests and graphs the outputs are not expected to be musical. The author found the LPM outputs for these cases tedious and monotonous. However, according to Zipf's law, zero value for slope govern the cases with high random events. A study on the LPM output graphs clarifies the reason for these situations (where the Zipfian slope is zero but the data has monotonous characteristics). In fact, in these graphs the logarithm of the frequency of occurrence of the events has almost the same values. Since there is limited number of such occurrences for these cases, the slope of the linear regression will tend to be zero. A sample of this case is depicted in figures 4-6:c, and 4-7

:c where the LPM outputs oscillate between two values, and the number of oscillations for the two values are equivalent. The logarithm of the frequency of occurrence of events would have almost the same values and the regression line fitted to these values will have zero slopes. The important thing to note is that although the Zipfian slopes do not show the true situation for these cases, they show that the LPM outputs are not musical. The items in the purple category are not expected to be musical (according to auditory tests and study of LPM output graphs) and the Zipfian slopes demonstrate that they are not musical.

The confusion matrix in table 4-6 is defined in the following (Table 4-6 was originally presented in (Arshi & Davis, 2015)):

- TP (True Positive) items have musicality according to human evaluation and their Zipfian slopes affirm their musical attributes (rusty orange cells).
- FP (False Positive) are items that are not musical (the human investigations show they are not musical) but their Zipfian slopes indicate that they have musical attributes (light blue cells).
- TN (True Negative) indicate items not labelled as musical by human and are correctly classified outside of the musical group according to their Zipfian distributions (yellow, dark green, and purple cells).
- FN (False Negative) refer to cases with musical output (according to human critics), however, Zipf's metric is not showing that (light green cells).

The Accuracy, sensitivity, and specificity of the classifier are computed as 87%, 55%, 98% respectively. High accuracy indicates that the Zipfian values for LPM outputs are likely to predict the musical and non-musical samples correctly. Middle range rate for sensitivity suggests Zipfian classifier average ability for identifying musical elements. The specificity demonstrates its success rate in correctly excluding non-musical individuals.

Table 4-6 Confusion Matrix.

Confusion Matrix	**Musicality =True**	**Musicality=False**
Zipf's musicality (Positive)	TP (254 items)	FP (30 items)
Zipf's musicality (Negative)	FN (207 items)	TN (1269items)

4.5 Applying Zipf's Law on a Crafted Sequence of Voices

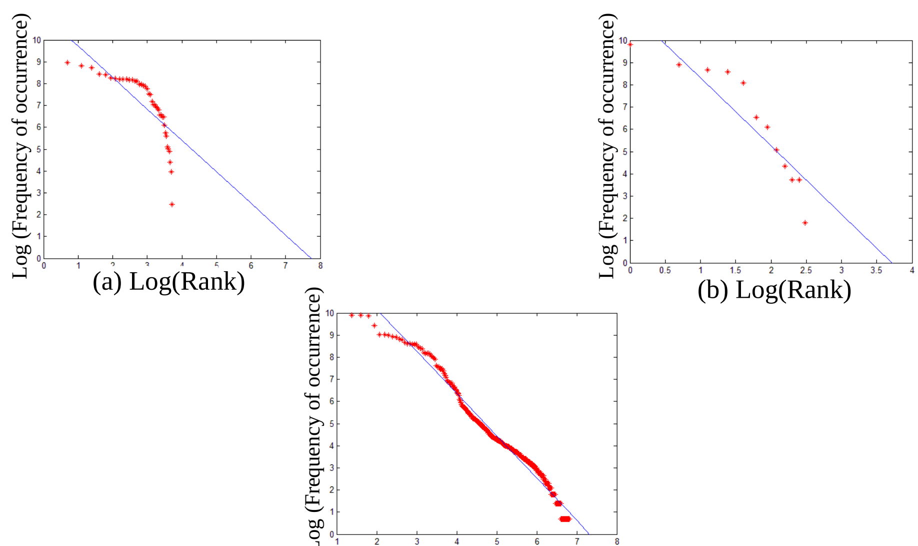


Figure 4-8 Examples of Zipfian distributions for the first experiment

Voices of random length, up to 10000 CA iterations,with obtained results of (a) $s = -1.44$, r^2= 0.66, (b) $s = -3.06$, r^2= 0.81, (c) $s = -1.91$, r^2= 0.96.

In the previous section the musicality of LPM outputs were studied individually for all the combinations of 88 CA rules, and 20 pattern-matching rules. In this section, two further experiments are conducted, where sequences of LPM outputs were investigated for their musicality based on Zipf's law. This section tests the assumption that sequences of LPM voices are more likely to achieve Zipfian ideal distributions rather than single LPM voices.

In the first experiment, 100 sequences were produced by randomly juxtaposing LPM outputs consisting of lengths up to 10000 iterations. The LPM outputs were selected randomly from all the possible configurations for CA rules and pattern-matching rules. In the first stage, the length of voices were randomly chosen amongst the total number of CA iterations (10000). The obtained Zipfian slopes range from -3.06 to -1.44 and their respective r-squared values are 0.81 and 0.66. Figure 4-8 illustrate the Zipfian distributions for three sample sequences.

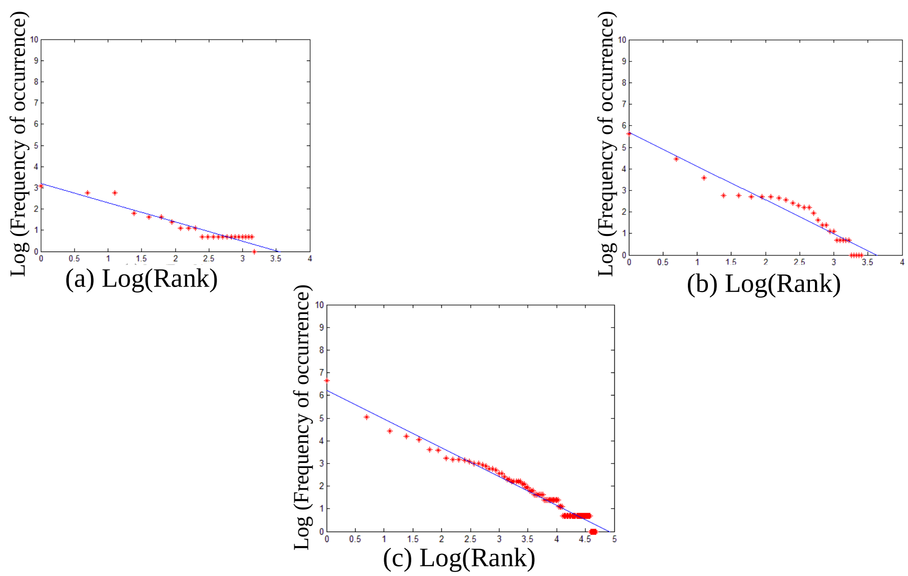


Figure 4-9 Examples of Zipfian distributions for the second experiment

Voices of random length, up to 20 CA iterations, with obtained results of (a) $s = -0.89$, $r^2 = 0.91$, (b) $s = -1.56$, $r^2 = 0.96$, (c) $s = -1.36$, $r^2 = 0.97$.

In a subsequent experiment the length of each of the voices were limited to consist of maximum 20 CA iterations. It was assumed that by decreasing the length of each of LPM voices in the sequences, the monotonicity of those characteristic elements would be lowered, while the diversity of the LPM voices presented in the sequences might resulted in having sequences with acceptable ranges for Zipfian slopes. In this last experiment, the minimum and maximum achieved Zipfian slopes were reported to be -1.56 and -0.89; and the R-squared values were 0.92 and 0.91. Some examples are depicted in figure 4-9. The experiments imply that sequencing voices with more appropriate characteristics are more likely to achieve better Zipfian results.

4.6 Why Evolve LPM Output

Tailoring the aesthetically pleasing combinations from the possible space of LPM emerging voices is a task, which seems to be consistent with the usage of evolutionary algorithms. The pattern-matching rules over CA define structured sounds known as voices. The applications of genetic algorithms have the possibility to tailor sequences of LPM output voices to make them aesthetically acceptable to audience. The search for finding optimal solutions is guided by assigning higher credits to more competent sequences.

The application of genetic algorithms has special necessities in search and optimization of musical sequences. Defining the search space; specifying the search space constraints; and the choice of appropriate fitness functions (Burton & Vladimirova, 1999) are some of the requirements for designing an evolutionary framework. There are infinite possibilities for generating music; therefore, it is necessary to specify suitable constraints to limit the search space. An idea could be to keep those melodies that conform to ideal Zipfian slopes and to discard the pieces that do not conform to the pre-determined aesthetical standards. Associating the music evolution to GA progression is another criterion that needs to be addressed. In fact, we need to clarify how the audio compositions are influenced by genetic algorithm operators.

4.7 A Design for Evolving LPM Output Using Genetic Algorithm

In the beginning of this chapter, LPM outputs were defined as a set of voices. These voices can be assimilated to musical motives of varying lengths. Crafting competent sequences of voices was recognized to be a suitable problem for evolutionary algorithms. In this section, sequencing LPM voices is taken as a search problem for producing aesthetically pleasing melodic structures. Designing such a system gives raise to the following questions (Arshi & Davis, 2016):

- How to design an efficient search space traversal, which resolves the sequencing problem within the constraints of given hardware resources?
- What are the possible approaches for sequencing voices in an aesthetically pleasing manner?
- What are the possible designs for the genotypes and phenotypes of a musical sequencer based on LPM?
- How to define musical critiques in order to criticize the musicality of LPM sequences?

The number of possible LPM voices are specified as $88^{7+4} * 20^{7+4} * t$, in which t is the number of CA iterations involved in generating LPM voices. It was previously stated that this number is reducible to $88^3 * 20^3 * t$ to simplify the current evolutionary problem at hand. Evolving LPM sequences based on their melodic structure and acoustics require the definition of a multi-optimization framework. The first search problem would target competent structures for the melody, including pleasing combinations of pitch frequencies, and note durations. The second search problem contributes to the optimization of the synthesizer parameters like ADSR envelopes. This division would provide different categories for exploration. Aesthetical critics required for evaluating the different dimensions of the produced audio would have different natures. However, in this thesis the focus will be kept on the $88^3 * 20^3 * t$ number of LPM voices. The evolutionary framework would be based on evolving sequences of LPM voices considering their melodic characteristics.

The definition of a search space consisting of all of the possible sequences of LPM voices is nearly impossible. Therefore, other approaches should be taken into account in order to reduce the search space by a notable amount. The definition of the search space should also consider our current hardware, and software facilities. Selecting a limited number of sequences of voices and evolving them at each time step of the evolutionary algorithm would be a feasible solution. During the evolution of the LPM sequences, all the involving parameters change dynamically to fulfil the predetermined musical expectations. Gradual changes of musical parameters provide general improvements in each generation. In the proposed evolutionary framework, there are no guarantees for having unique solutions to musical problems. In fact starting from the same initial conditions, the exploration may result to differing sets of LPM sequences in every execution.

A suitable design of genotypes and phenotypes are necessities for an efficient search. The genotypes are the underlying codes, which manifest a more tangible level of behaviour or appearance that is known as the phenotype. For example, the hair colour is coded in genes. The hair colour seen as brown, blonde, etc., are known as phenotypes. In the LPM system, the genotypes are the set of genes coded whether as binary or integer representations. The phenotypes are the audio sounds, which are heard as the characteristics of the individuals. In the beginning of the next chapter, the genotypes are defined as sequences of values, which embed the required information for producing audio pieces.

Zipfian aesthetical measurements were selected as critics for determining the musicality of the generated LPM sequences. In the next chapter, a fitness function will be defined with a core based on Zipfian metrics extracted from Persian music. This fitness function is responsible for giving more credits to LPM sequences with more acceptable Zipfian distributions based on Persian music.

4.8 Chapter Summary

This chapter presents one of the fundamental experiments of this thesis. LPM was introduced which works on the basis of pattern-matching rules. The pattern-matching rules extract features from CA progression and feed a Persian musical instrument synthesizer. The emergent behaviour of CA is the heart of our machine composition system. Different choices of CA and pattern-matching rules once given, define the parameters of the synthesizer, and therefore they produce voices. These voices can be assimilated to musical motives.

In the next step the output of the LPM system were studied according to their musicality and in terms of Zipf's law as an aesthetical criterion. It was illustrated that the musicality of the produced audio can benefit from sequencing different CA rule behaviours and pattern-matching rules. This effect was studied through contrasting limited CA and pattern-matching rule output behaviours, which did not produce desirable Zipfian characteristics.

It was demonstrated that sequencing those voices might need to undergo many tailoring processes to produce the final desirable outcome. Since LPM system deals with a large number of possible voices, evolving them in an evolutionary environment might be the first choice in such a design. In the next chapter, an architecture is proposed for performing such experiments on the output of LPM voices.

Chapter 5. Evolving LPM Voices using Genetic Algorithms

A computational framework for evolving LPM voice sequences was proposed in the previous chapter. In this chapter, an experiment is performed based on the idea of exploring the LPM audio space by the application of evolutionary algorithms. The audio space is a conceptual space consisting of all the combinations of LPM voices. There are countless possible sequences of voices available. Exploring this space is assimilated to exploring the space of creativity in the hope for finding new musical forms or artefacts. In this chapter, aesthetical measurements based on Zipf's law are employed as a base for designing fitness function for an evolutionary algorithm. An evolutionary environment is developed to enable the search problem contributing to the melodic structure of LPM output.

5.1 The Design of LPM Sequencer based on Zipfian Metrics

5.1.1 Genotype Representation

The evolution of LPM sequences by the application of genetic algorithm (GA) requires the design of a competent evolutionary computational environment. This involves the establishment of the representation of genotypes, and phenotypes. The design of the genetic operators and fitness function are other important requirements of an evolutionary framework. Two possible representations of genotypes are suggested in this chapter.

The evolution of first level of genotypes can be assimilated to evolving the musical motives in a sequence. The altering of the musical motives in the first level can consist of a dramatic change to the musical motive as well as replacing it with a complete different musical motive. The second level of evolution gives access to elements of musical motives including musical notes frequencies, note durations, and note onset times. The second level of evolution alters a small proportion of the musical notes and/or durations and note onset times in a musical motive.

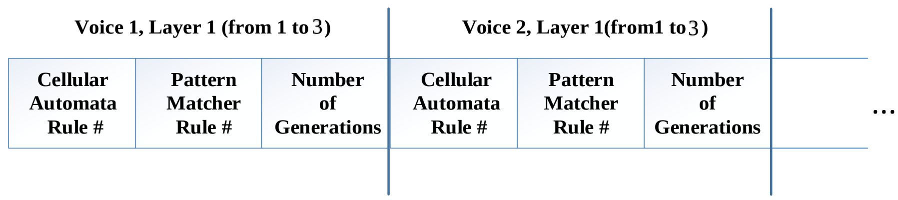


Figure 5-1. The genotypes in the first level for each of the three constituent layers.

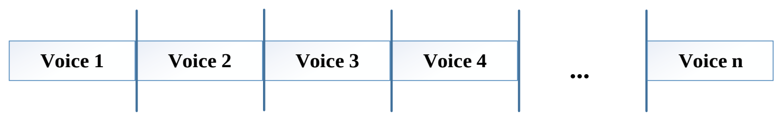


Figure 5-2. The genotypes in the second level, after expanding the genotypes in the first level.

In the first proposed representation, the genotypes consist of sequences of triplet blocks. The values in each block stand for the CA rule number, pattern-matching rule number, and the number of CA iterations. A schema of the described genotype representation is presented in figure 5-1. The alternative representation of genotypes incorporates the expanded version of the triple-block schema, demonstrated in figure 5-2. In this level of presentation, the genotypes are the outputs of pattern-matching rules applied on t iterations of CA progressions for each of the triple-blocks in the first level genotype format. The sequences of blocks are expanded in their representational form in a way that individual elements in the sequence are subject to evolution. Genotypes in their expanded format allow for genetic operations which have more detailed impact on the sequence. Therefore, their associated crossover and mutation operators cover alterations at an elemental level (for instance pitch frequency or duration of a musical note).

The two designs for genotypes can be considered as representations prepared for two different levels of evolution (which can occur sequentially or independent of each other). In the first level of evolution, the nature of a whole triple blocks in a sequence are subject to evolution. Each of the triple-blocks represents one voice (or musical motive). For instance if the CA rule number and/or the pattern-matching rule number changes for a block as a subject of evolution, the whole voice associated with the triple block will change. Extending the genotypes provides the potential for defining additional operators for evolving the sequence in a more scrutinized manner and having access to the elements of the musical motives. The two different genotypes require the design of distinct set of crossover and mutation operators regarding their natures.

In the proposed genotype architectures, the length of the genotypes may be different from each other. Therefore, the musical pieces produced by this method would have varying lengths. Nevertheless, due to potential computational overload, this length has a maximum constraint. It is notable that the allowance of having various numbers of musical motives would produce sequences with variable lengths. Additionally, the number of CA iterations for each of the voices governs the length of the sequences from the aspect of number of notes in the musical motives. However, the sequences are subject to evolution and their lengths may vary throughout the execution of the algorithm according to their competence.

5.1.2 Genetic Algorithm Operators

Evolutionary algorithm operators, are tools which assist in the navigation of search space (Burton & Vladimirova, 1999). In this section a series of genetic algorithm operators are presented. The selection, crossover and mutation operators are inspired from (Goldberg, 1989; Lo, 2012; Matlab, 2016) and justified for evolving LPM sequences. Each of the chromosome blocks consists of the required information for producing musical motives. Then there are sequences of blocks which would be translated to a sequence of musical motives.

Apart from traditional or classical genetic operators (Goldberg, 1989; K.F. Man et al., 1999), additional sets of operators are proposed in (Lo, 2012; Manaris et al., 2007) and are specifically designed for working in the musical search space. These types of operators are inspired from compositional techniques as well as transposition, retrograding, rearranging, and swapping the musical motives. These sets of operators are customized for evolving LPM sequences as described below.

5.1.2.1 Mutation Operators

A proportion of the individuals are selected for performing mutation operator on them. Mutation fraction parameter settles the fraction of the population which are nascent by mutation operation. Mutation rate parameter determines the probability of a genome in an individual chromosome undergoing the mutation process. Various mutation operators examined in the implementation are:

- Classic Mutation: Mutation probability determines whether each of the genes along the chromosome would undergo the mutation operation. Each of the genes are allocated a random number. In case this number is greater than the mutation probability, the gene will mutate, otherwise it will be left untouched. The rest of the mutation operators are inspired by (Lo, 2012) and are in the following.

- Multiple Element Modification: N random genes are chosen along the chromosome. The selected genes are then subject to mutation.
- Segment Alteration: A portion of elements are selected from the chromosome and the associated genes in the chromosome are altered.
- Segment Placement: Two segments are selected randomly (within a single chromosome or two different chromosomes). The fitter segment is replicated and replaces the weaker segment. Determining the competence for each of the segments follows the same procedure for obtaining the fitness value for the individuals in the population. Alternatively, the fitness of the randomly selected segments can be measured in association with their contribution in the fitness of the chromosomes that they are going to be embedded.
- Segment Retrograde: The elements in a randomly selected segment are sorted in a reverse manner.
- Segment Transposition: A constant value is added or subtracted from the elements in a randomly selected segment. The constant value is chosen in a way that performing the computation does not hurt the acceptable threshold ranges of the element values.
- Segment Ordering: The elements in a randomly selected segment are ordered in an ascending or descending manner according to the elements' values.
- Segment Copy/Paste from musical dataset: A segment from a musical piece in the original database is randomly selected and then replaces a random segment from an individual in the population.

5.1.3 Fitness Function Design

One of the challenges in the design of an evolutionary framework is the determination of a suitable fitness function. In this thesis, aesthetical aspects of the generated audio are the focus of attention. These are aspects which contribute to aesthetical criteria while preserving the novelty of the generated materials and directing the formation of voice sequences to be comparable to the style of traditional Persian music. Fitness function built on the basis of Zipfian metrics are aimed for directing the evolution towards producing sequences which have Zipfian metrics similar to those of Persian music. The aesthetical critic is a support vector machine regression (SVR) model which is trained to differentiate traditional Persian music and LPM sequences. The dataset for training the SVR model consists of Zipfian metrics extracted from traditional Persian music and LPM sequences. In order to train the SVR model, the Persian music pieces and LPM sequences are labelled as 1, and 0 respectively. When the SVR model is used as fitness function, it assigns decimal point numbers to the members of the population. The assigned credits determine the similarity of the generated sequences to traditional Persian music pieces or to LPM sequences. The higher the score attained, the more chances are given to the genes survival and it will become more likely for the genes to pass their information to the next generations. In this chapter, all the required considerations for designing the fitness function are presented.

The reason for applying support vector regression models in this thesis instead of support vector machines lies behind the values that the trained SVM and SVR models achieve by being the fitness functions of an evolutionary algorithm. Support vector machines obtain 0 or 1 values as the label of classes for the newly generated audio samples (as a result of evolution of LPM sequences). However, in the current research, decimal point numbers (as produced by SVR) are more desirable as fitness function values for utilization in the genetic algorithm.

5.1.3.1 Training Dataset

One of the obstacles in the way for doing research on traditional Persian music is the lack of existence of standard MIDI databases. The existence of a standard pre-processed database help different researchers to compare their results while applying various tools. Having such MIDI databases would make the calculation of Zipfian metrics easier and more precise. This gap is usually fulfilled by signal processing tools and working with audio files instead. For this project, a MIDI like database was prepared by the help of signal processing toolboxes. The extracted features from Traditional Persian music audio data includes pitch frequencies of notes, note durations, and note onset times. This section delivers the procedure for preparing the traditional Persian music dataset and also discusses about how the collection of LPM sequences were arranged.

The traditional Persian music audio databases were selected from Radif musical pieces collection from “Ostad Faramarz Paivar” and were performed and recorded by “Ostad Saeed Sabet” (Paivar & Sabet, 2004). The musical instrument utilized for performance was Persian Santur. The permission for academic usage of the music recordings were granted from “Ostad Saeed Sabet”. There are a total number of 110 pieces from 10 different Dastgāhs present in the audio database. The lengths of the musical pieces vary from around 50 seconds to 5 minutes. Signal processing tools and techniques were required for extracting the musical features from the audio files. The pitches and note onset times, and durations were attained for producing a symbolic dataset. Various toolboxes were studied as well as Aubio, and MIR toolbox (Lartillot, 2013). MIR and MIDI toolboxes were widely used in this research for feature extraction purposes. One reason for choosing those toolboxes was that the scientific experiments for studying the nature of Liquid Persian Music (LPM) voices were written in Matlab for its suitability for working with matrices (However, LPM user interface was written in c++). Applying those toolboxes helps to stay in the same experimental environment.

An algorithm was setup for extracting the notes information as well as duration and the pitch frequencies in each of the musical files. Note onset times determine a series of consecutive times when the musical notes were excited. This specification was detected by the application of *mironset* command from MIR audio library which discovers the bursts of energy. The *mironset* command is followed by *mirsegment* and *mirpitch* commands as described below. The segmentation of the audio pieces were performed in the next step in regards to the achieved onset times. In other words, the notes starting times are the beginning of the segments. Each of the segments were analysed for recognizing the note's dominant pitch by the application of an autocorrelation algorithm.

The auto-correlation algorithm may achieve slightly different results for the note pitches as they occur in the musical piece. However, obtaining Zipfian metrics for musical events requires the calculation of the number of occurrence of those events. Therefore, further refinements on the achieved pitch values were required. On this account, the concept of pitch bins was introduced for justifying the pitch values. The midpoints of the pitch bins were selected to be the standard pitches of nine-bridged Persian Santur with their various tunings for performance in different Dastgāh. A table of the 27 Santur notes and their pitches are provided in Appendix E. The obtained pitch values for musical notes were standardized according to the pitches in table E-1.

Extracting the onset times and duration of the musical notes were obtained by following the *mironset* command using MIR toolbox. The existence of ornamental notes and nature of Radif music makes the definition of duration bins complicated. In fact, one of the characteristics of Radif of Persian music is that most of the pieces do not necessarily follow a tempo; the duration specification of musical motives may vary during the performance depending on different factors. Some of these factors reflect the personal moods of the performer or the singer expressions during the performance, or the audience preferences. The implication derived from this discussion is that the quality of the performance of Radif differs in any single performance. Perhaps this reason makes the preparation of standard MIDI databases more difficult. The definition of the duration bins is only possible for the pieces which follow a tempo. The note onset times and durations values were extracted from music pieces and were directly inserted in the MIDI tables without using duration bins. This is unlike the cases of note pitches which were standardized using pitch bins.

Table 5-1.The Normalization value ranges for the musical dimensions in LPM sequences.

Musical Dimension	**Metric**	**Lower Range**	**Upper Range**
Pitch Frequency	Hertz	100	3000
Duration	seconds	1/16	2
Onset	seconds	1/16	2

In the remainder of this sub section, there would also be a look towards how the LPM sequences dataset were prepared for training SVR model along with traditional Persian music. The LPM sequences consist of a consecutive arrangement of voices; each of which were obtained by applying a pattern-matching rule over t number of iterations of a CA rule progression. The length of a sequence of voices equals the total number of iterations applied in such a system. Since each of the voices can have 11 governing dimensions as mentioned in the end of section 4.2 in chapter 4, the production of databases can continue to 11 dimensions. The evolution of melodic structures of LPM sequences is the focus of the research in the first place. Therefore, the pitch, duration, and onset times were taken into account as 3 dimensions out of the 11 dimensions. Evolving the audio sequences based on the remaining dimensions is subject for future investigations and are out of the scope of the thesis. This is also because the Persian music databases at hand are not based on the alteration of ADSR envelopes since a traditional musical instrument is applied. Table 5-1 demonstrates the normalization parameters used for preparing LPM sequences database (The complete table for normalizing all the 11 governing dimensions can be found in table E-2 in Appendix E for the interested reader).

The length of the sequences may differ from each other; however, there should be a correspondence between the lengths of the associated sequences in different dimensions. If each of the sequences occupies a record in the database, the number of columns would be dissimilar; however, the identical records in the databases associated with the data dimensions would have the same lengths since they correspond to the different dimensions of the same musical events.

The parameters should also make sense from the aspect of the structure of the music as well. For example, the time distance between consecutive notes should not take too long in a single composition. For example, ten minutes is not acceptable as note duration. The normalization makes sure that there is rationality in the eventuality of the musical events and the synthesizer parameters. Of course, putting this kind of limitation on the note duration might be in contrast with willing to break the norms and expectations in order to have outcomes which might be considered as creative. For instance, the work of John Cage in his composition “four minutes and 33 seconds of silence” (Gutmann, 1999) is an example of a creative artefact which surprised the audience by breaking the music norms and expectations. However, the productions of the designed system in this project are no longer than a few minutes and we did not intend to produce audio which might have been almost silent. The audio samples were selected to be short for the purposes of the auditory surveys designed in chapter 7.

5.1.3.2 Zipfian Feature Extraction

The Zipfian metrics were extracted from both the traditional Persian music database at hand and from LPM sequences samples. The number of musical pieces are 110 in the database which were later segmented to 20 second pieces without overlapping. This provided 766 training samples from the Persian music database in total. The number of samples in the LPM database have also been selected to be 766. The length of the LPM sequences in time have also been selected to be approximately 20 seconds each. The number of instances in the training dataset becomes 1532 (766 Persian Music segments +766 LPM sequences). A total number of 248 Zipfian features were extracted from each of the available pieces in the databases. The Zipfian metrics can be calculated in three different stages all of which contribute to the final Zipfian metrics. These stages or categories are defined by (Manaris et al., 2005) and are as follows:

- Regular Metrics: The sets of values in regular metrics were achieved by taking the first level musical events directly into account. This includes the counts of pitch, duration, and note onset occurrence in a musical piece.

- Higher Order Metrics: These metrics calculate differentiations of the events from the previous stages. The differentiations are often calculated up to three or four degrees. Further degrees of differentiation may not reveal further information. The differentiations were achieved by taking the differences between the two events in the regular metrics. The derivatives of regular metrics were calculated up to 4 levels. Some of the examples of higher order metrics are *duration_d1*, *duration_d2*, *duration_d3*, *duration_d4*.

- Local variability metrics: The metrics in this category were dedicated to the calculation of the entropy difference of an event from the local mean. These metrics were obtained from the regular metrics and the attributes up to second degree derivatives. For instance, *Duration_LV*, *Duration_d1_LV*, and *Duration_d2_LV* are some of the attributes which were built upon the duration matrix.

5.1.3.3 Data Cleaning

Data cleaning is one of the important steps in the process of preparing the data for training to learning machines. The Zipfian metrics obtained from traditional Persian music and LPM sequences were studied in detail for further refinements prior to the achievement of the final SVR model as fitness function. Data cleaning makes the fitness function more efficient, since only the necessary attributes will be left for training to the model. As will be discussed in this section, only a few number of attributes were selected for having a machine learning model capable of crediting the items. The data cleaning performed in this section was assisted by *Weka* (Hall et al., 2009).

Table 5-2. A list of example attributes which have been removed. Supporting graphs showing their value distribution are available in figure 5-3.

Attribute Name	Feature Category	Reason Removed
Durations	First level metrics	Non- overlapping classes
Melodic_Interval_d2	Higher order metrics	
Pitch_distance2_d1_LV	Higher order Local variability metrics	
Duration_Bigram_LV	Local variability metrics	
Harmonic_Bigram_d2	Higher order metrics	Zero, Nan or minus infinite values for one or both of the classes
Duration_Distance2_d1	Higher order metrics	
Harmonic_Interval_LV	Local variability metrics	
Combined_Pitch_Duration	First level metrics	
Harmonic_4gram_d2	Higher order metrics	
Chord_Progression_d3	Higher order metrics	
Harmonic_Bigram_d4	Higher order metrics	
Melodic_Consonance_d2	Higher order metrics	
Quantized_Duration_Distance1_LV	Local variability metrics	
Harmonic_Consonance	First level metrics	

The data cleaning phase would cut down the burden of computing a large number of Zipfian features for all the population member sequences in each of the evolutionary generations. The process of data cleaning started with identifying attributes with large number of same values. They were counted as redundant data and been eliminated from both of the traditional Persian music Zipfian metrics and LPM Zipfian metrics. The attributes with redundant values came with a large number of minus Infinite (-Inf) and NaN, or zero values. Minus infinite (-Inf) indicates that the measured musical data is monotonous. NaN (Not a Number) shows the unavailability of the related event so that it could not be measured. Despite the meaning NaN and -Inf values have in Zipfian terminology, when they are overly repeated for various records, it makes the related attributes biased and meaningless. There were also few records present in the database which had single minus infinite values. These records were also eliminated due to the fact that the attributes values would make the Weka classifier act abnormally and produce erroneous results. Performing the data cleaning in this manner left the database with 82 attributes out of 248 features.

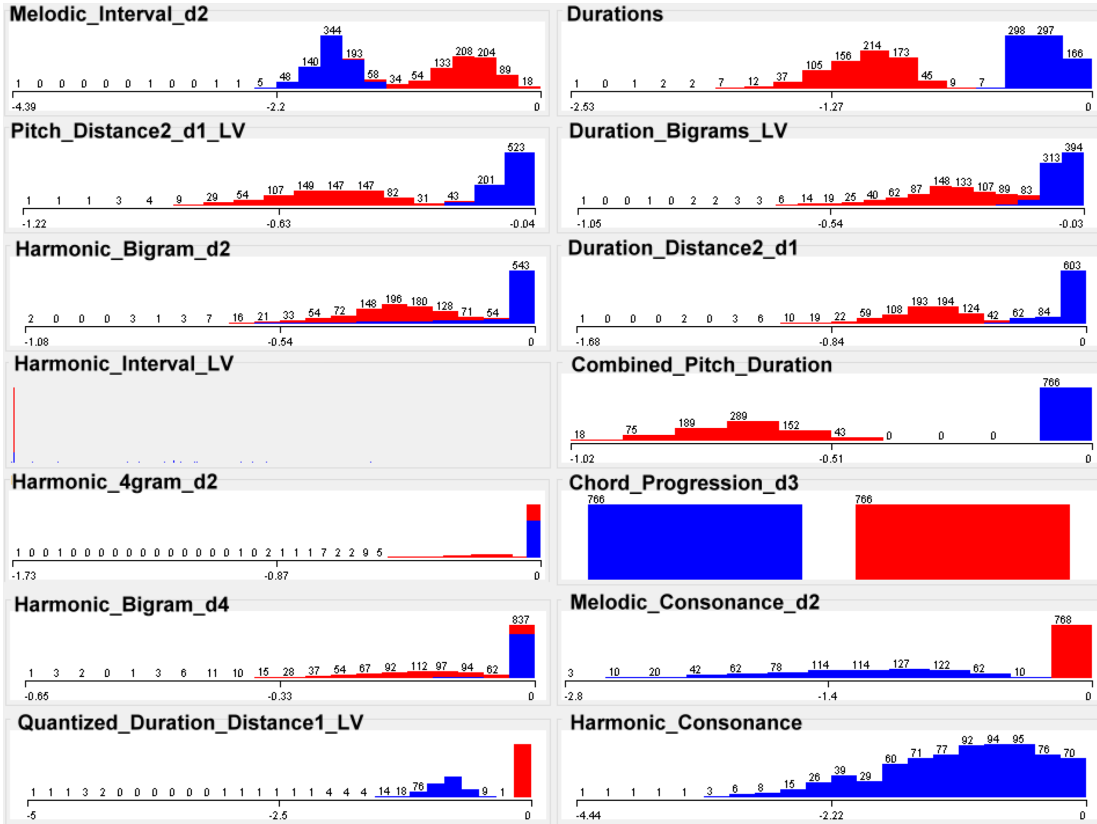


Figure 5-3.Some examples of the discarded attributes.

The red colour stands for samples from LPM sequences class, while the blue color shows the samples from Persian music class. Discarded attributes are those which have undefined, minus infinity, many zeros, or redundant values over one or both classes. This also includes the attributes which do not have spectacular overlapping over the two classes (It is discussed in the main text that we are not looking for these kind of attributes although they have strong dichotomizing characteristics).

Our data cleaning procedure also followed another step, which is not often common in general practices of data cleaning and is specific to this application. In machine learning, the classification usually follows making progress into producing strong dichotomizers between the classes under study. Studying the attributes in the categories may reveal those that are non-overlapping. Obviously, in the classification practices, attributes with such specifications will overcome the rest of the attributes and the machine learning tools will rely on those certain attributes for performing the classification tasks. The point to make here is that, despite the discriminating power of non-overlapping attributes, they are undesirable attributes in this research and the attention was to eliminate those bipolar attributes in the data cleaning phase. Table 5-2 and figure 5-3 provide a list of undesirable attributes examples which were removed from the databases due to their redundancy or having bipolarity characteristics.

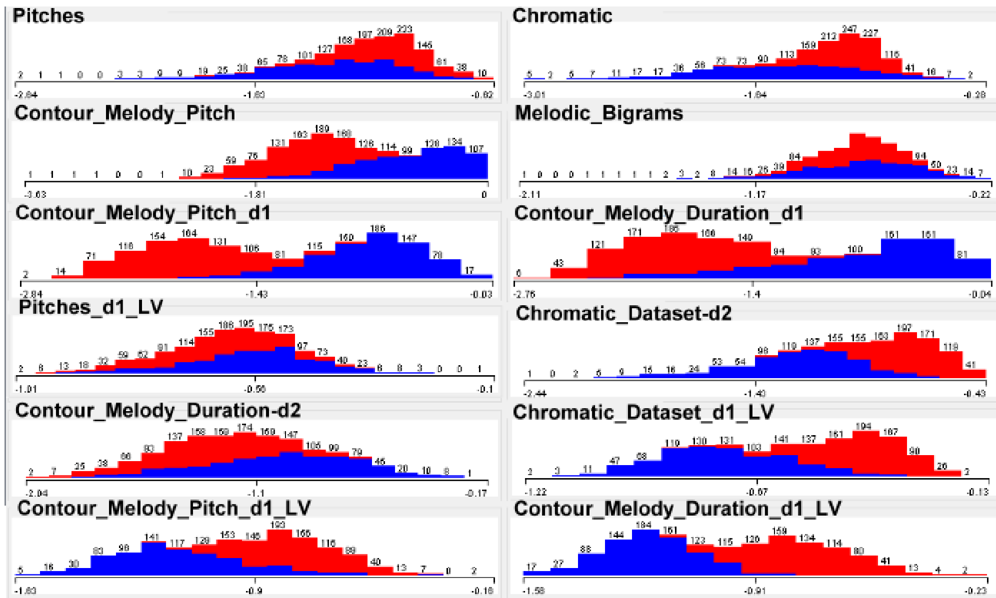


Figure 5-4. Example of attributes with acceptable overlapping classes.

The red colour stands for smaple from LPM sequences class, while the blue color shows the samples from Persian music class.

The overlapping attributes are of interest in the training of the SVR fitness functions. The reason for omitting the bipolar attributes is that a fitness function based on bipolar attributes would make most of the generated audio samples to be discarded in the first few generations without giving them the chance to evolve and contribute in the evolutionary process. On the other hand, the designed fitness function based on overlapping features would give higher credits to those sequences which bear more resemblance to traditional Persian music Zipfian metrics and will try to evolve LPM sequences in the style of Persian music. Taking out the non-overlapping attributes would leave the database with 49 features. Figure 5-4 depicts features that possess overlapping specification over the distribution graphs of records from both classes of LPM and traditional Persian music Zipfian metrics.

5.1.3.4 Attribute Selection

Table 5-3. Example of Attribute Selection procedure. Values less than 0.1 are colour coded. Attributes coded yellow have values greater than 0.1 for all metrics. (Full table in table F-1 in Appendix F)

Attribute Number	Attribute Name	ReliefF	Gain Ratio	Information Gain	Symmetrical Uncertainty
1	Pitches	0.0638	0.0862	0.1304	0.1038
2	Chromatic	0.0854	0.1504	0.2489	0.1875
3	Pitch_Distance1	0.0581	0.2671	0.5951	0.3687
4	Contour_Melody_Pitch	0.1122	0.2683	0.5462	0.3599
5	Melodic_Bigrams	0.0232	0.0524	0.0474	0.0498
6	Melodic_Trigrams	0.0205	0.0984	0.1373	0.1146
7	Melodic_4grams	0.0211	0.1193	0.1864	0.1455
8	Contour_Melody_Pitch_d1	0.194	0.3681	0.7087	0.4845
9	Contour_Melody_Duration_d1	0.1309	0.2843	0.6016	0.3862
10	Melodic_Bigram_d1	0.0298	0.1456	0.2104	0.1721
11	Chromatic_DataSet_d2	0.1428	0.3007	0.5296	0.3836
12	Pitch_Distance1_d2	0.0585	0.346	0.6392	0.449
13	Contour_Melody_Pitch_d2	0.0677	0.2372	0.4624	0.3136
14	Contour_Melody_Duration_d2	0.0311	0.0955	0.1797	0.1247
15	Chromatic_DataSet_d3	0.0425	0.1201	0.19	0.1472

After performing initial data cleaning, the second phase would be the attribute selection phase. The attributes were evaluated in Weka by different algorithms such as ReliefF, Gain Ratio, InfoGain, and Symmetrical Uncertainty by the application of Ranker search method (Kononenko & Kukar, 2007) (as described in chapter 2).

The attributes were ranked in descending order by Weka and were assigned scores between zero and one. The given values assist in the detection of competent features. The underlying attribute selection process started with marking the attributes with evaluation values below 0.1 as undesirable ones. Table 5-3 depicts a small portion of attributes and the accredited evaluation values by different random search methods. The complete table is available in appendix F. Table 5-3 also illustrates the methodology for selecting the attributes consisting of four different ranker searches. The attributes were then sorted by their attribute numbers by the help of Excel custom sort.

Table 5-4.Ten features having the values of greater than 0.1 among all evaluators. Continuing the attribute selection process with CfsSubset Evaluation achieves 8 attributes shown in colour-coded format.

Attribute Number	Attribute Name	ReliefF	Gain Ratio	Information Gain	Symmetrical Uncertainty
4	Contour_Melody_Pitch	0.1122	0.2683	0.5462	0.3599
8	Contour_Melody_Pitch_d1	0.194	0.3681	0.7087	0.4845
9	Contour_Melody_Duration_d1	0.1309	0.2843	0.6016	0.3862
11	Chromatic_DataSet_d2	0.1428	0.3007	0.5296	0.3836
19	Chromatic_DataSet_d4	0.1394	0.2269	0.4473	0.301
26	Contour_Melody_Duration_LV	0.1242	0.2161	0.4945	0.3008
27	Melodic_Interval_LV	0.1235	0.2762	0.4584	0.3447
33	Chromatic_DataSet_d1_LV	0.1653	0.2829	0.5545	0.3747
36	Contour_Melody_Duration_d1_LV	0.1644	0.3564	0.7508	0.4833
42	Chromatic_DataSet_d2_LV	0.1838	0.4275	0.7631	0.548

This would arrange different evaluators' value for each attribute in one row, so that the evaluations can be compared with each other. Colour coding assisted throughout the attribute selection procedure. Those attributes which were not marked as undesirable in any of the evaluators' categories win the selection process and will enter the final phase of cleaning before the training is performed. A total of 10 attributes remain which have the evaluation values greater than 0.1 among all the rankers. The number of the selected attributes shrinks further by the application of *CfsSubset* (Hall et al., 2009) evaluation. Eight attributes remain in the final stage which are colour coded in table 5-4. As can be seen in the table attribute number 4, and 27 were discarded, although they have higher scores than attribute 19 in the symmetrical uncertainty column. The fact is that CfsSubset obtains the most competent set of attributes for classification.

Table 5-5. Classification performance for SVM , J48, and NaiveBayes for attribute selection.

The complete table is provided in table F-2 in appendix F. The first section of the table shows the results of SVM, J48, and naiveBayes algorithm on the original dataset with 248 features. The second section of the table shows the results of the classifiers with 49 features which were left after performing the attribute cleaning phase (The attributes with redundant values, so many Nan or minus infinite, or zero values were removed. In this phase the nonoverlapping attributes for the two Persian music and LPM pieces were also eliminated.) The third phase of attribute selection leaves us with ten features (ReliefF, Gain Ratio, Symmetrical uncertainty, and Information gain methods were used for performing attribute selection). The last phase of attribute selection after performing CfSSubset evaluation leaves us with 8 feaures. Although the attributes were shrinked by a significant amount, the different classifiers maintain their performance. Working with a small number of attributes saves processing time in the computations of our evolutionary algorithm. All parts relate to 1532 instances in the training data.

		Number of Features Attributes	**Correctly Classified Instances(%)**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
1	SVM	248	100	100	100	100	100
	trees.J48	248	98.82	99.09	98.56	98.57	99.08
	bayes.NaiveBayes	248	99.80	99.61	100	100	99.61
2	SVM	49	100	100	100	100	100
	trees.J48	49	98.82	99.09	98.56	98.57	99.08
	bayes.NaiveBayes	49	99.80	99.61	100	100	99.61
3	SVM	10	99.67	99.35	100	100	99.35
	trees.J48	10	98.95	99.22	98.69	98.70	99.21
	bayes.NaiveBayes	10	99.80	99.61	100	100	99.61
4	SVM	8	99.73	99.48	100	100	99.48
	trees.J48	8	98.95	99.22	98.69	98.70	99.21
	bayes.NaiveBayes	8	97.78	99.61	99.87	99.87	99.61

Table 5-5 shows the performance of various machine learning tools trained with different number of attributes in various stages of data cleaning, and attribute selection. The number of features are 248, 49, 10, and 8 in the experiment. The result provided in the table witnesses that the classifiers are still able to perform outstandingly after the dramatic shrinkage of the initial attribute sets. The table shows that the eight subset of attributes has better performance than ten attributes for SVM model.

5.1.4 The Genetic Algorithm Implementation

GA evolution starts by generating the initial population of LPM sequences, each consisting of a level one genotype. The second level genotype is then expanded from each individual of the first type. It has been demonstrated that the types of operators which work on these genotypes are different from each other. However, the concept of GA stays the same for both of the models. In other words, they can be considered as two stages of evolution occurring one after another. This would provide the chance to study the effects of operators on the evolution of chromosomes in different levels. The program is designed in a way that these stages can occur consequently or the user can only choose one level of evolution as well as directly select to evolve the LPM sequences by the application of the extended forms of genotypes. The evolving sequences can be stored in a dataset for various generations. This gives more flexibility since the programmer can delve into the code for manipulating the GA operators for studying their effects and/or resume the evolving of the genotypes from a specific generation.

The first population of genotypes were produced by generating random sequences of LPM in three dimensions. Each of the governing dimensions stands for pitch, duration, and note onset time information. The genotypes in the first level consist of the arrangement of triplet arrays including CA rule number, pattern-matching rule number and the number of CA iterations. In the first level, the sequences are evolved based on the suitability of the triplets. The genotypes in the second stage can be considered as the extended format of the genotypes in the first level genotype. Once this stage is settled and/or specific number of iterations is visited, the second level of evolution starts by expanding the genotypes taken from the latest evolved generation of the first evolutionary stage. This is due to the nature of GA operators working on the genotypes which target the elements in a more scrutinized manner. Once the level one genotypes are decoded to their level two format, the transformation of the genotypes to their former format is irreversible. This means that the triplets in the sequence can no longer be reversed to the format of: CA rule number- pattern-matching rule number- and number of CA progressions. In the second level, a variation of operators targets the sequences in elemental level and operators are based on some aspects of musical theory. For each of the GA iterations, the reduced set of Zipfian metrics were extracted from the individual sequences in the population.

5.2 Results

Table 5-6. GA Configuration Parameters for Graphs shown in Figures 5-5 to 5-9.

Parameter	Fig a	Fig b	Fig c	Fig d	Fig e	Fig f
mutation_rate	0.8	0.8	0.25	0.5	0.5	0. 25
mutation_fraction	0.9	0.9	0.25	0.5	0.5	0.25
crossover_rate	0.5	0.8	0. 5	0.5	0.8	0.9

In this section, the results of evolving LPM sequences are presented and further refinements are suggested. The experiment is conducted with 8 features after final attribute selection process for training the support vector regression model. The performance of the algorithm is studied against different parameterizations and choice of operators, as indicated in table 5-6. Elite percentage and tournament size are fixed at 5%, and 5 respectively. Figures 5-5 to 5-7 present the minimum, mean, maximum of the fitness function over GA progression for both genotype format levels. Figure 5-5 is associated with the first level of evolution and depicts gradual growth in the max of the fitness function values over 100 generations.

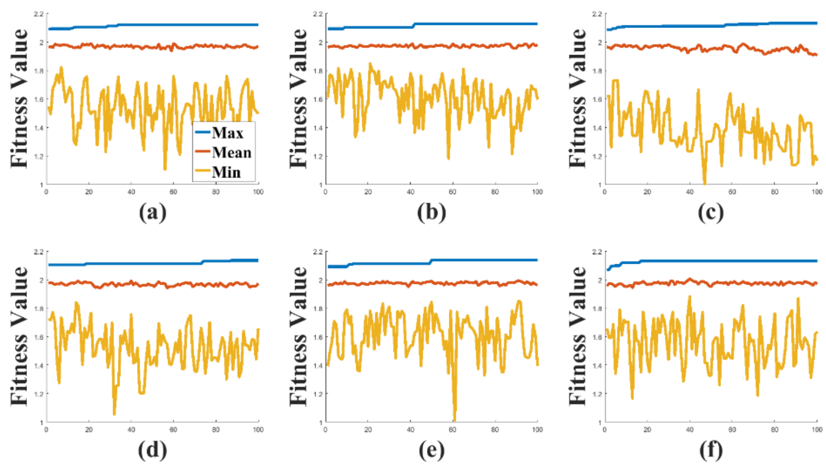


Figure 5-5. Mean, min, and max of fitness values for the first level genotype.

These subfigures illustrate the first level of evolution over 100 generations. This experiment is conducted with fitness function on the basis of 8 cleaned features. The x axis represents the evolution progression, while the y axis shows the fitness values. The 6 GA configurations for different subfigures are presented in table 5-6. The gradual growth in the values of the fitness function suggest the improvement of LPM voices from Zipfian aesthetical critics point of view.

Figure 5-6 magnifies the max fitness function over 1000 generations in the first level of evolution. The first level of evolution shows improvement in the fitness values during the GA evolution. Figure 5-7 shows the result of the evolution after decoding the genotypes and performing the second range of operators (second level of evolution). This, as depicted, does not produce further improvements. This is in contrast with the first level of evolution where the fitness values shows gradual improvements.

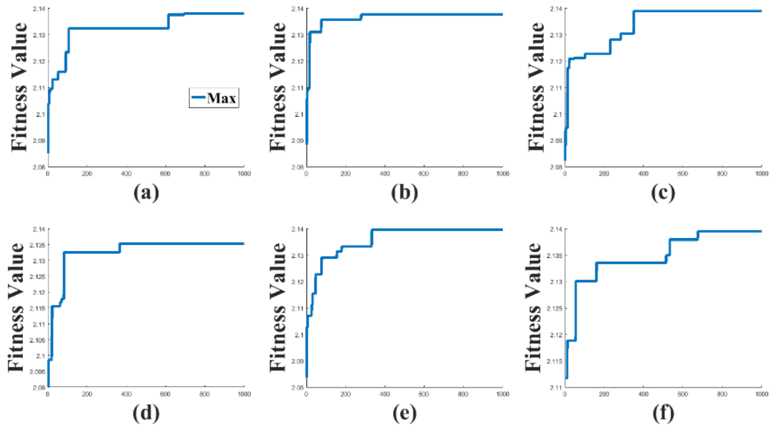


Figure 5-6 Max of fitness values for the first level genotypeover 1000 generations.

These subfigures relate to the first level of evolution over 1000 generations. This experiment is conducted with fitness function on the basis of 8 cleaned features. The x axis represents the evolution progression, while the y axis shows the fitness values. The 6 GA configurations for different subfigures are presented in table 5-6. The gradual growth in the values of the fitness function suggest the improvement of LPM voices from the Zipfian aesthetical critcs point of view.

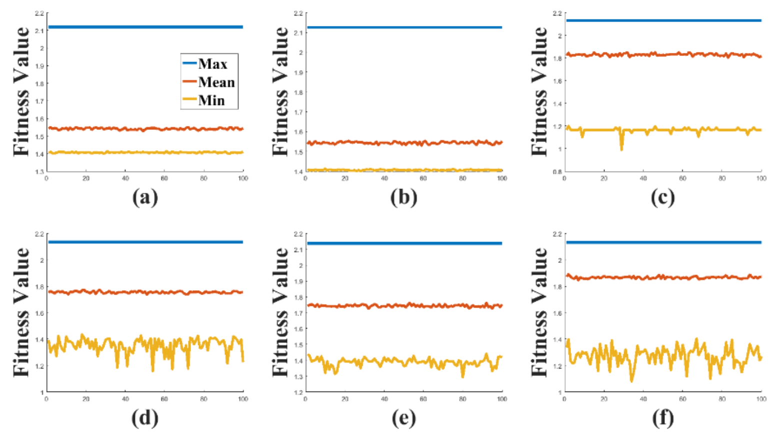


Figure 5-7. Mean, min, and max of fitness values over 100 generation in the second level of evolution continued from the evolution in the first phase.

This experiment is conducted with fitness function on the basis of 8 cleaned features. The x axis represents the evolution progression, while the y axis shows the fitness values. This sub figures illustrate the second evolution stage in the continuation of the first one. The 6 GA configurations for different subfigures are presented in table 5-6. Fitness function values suggest no further improvements in the quality of evolved voices considering the defined aesthetical objective based on Zipfian metrics.

Comparing the resulting figures in the first (figures 5-5, and 5-6) and second (figures 5-7) levels of evolution show an improvement of the fitness values occurs in the first level of evolution, unlike the second level of evolution process.

Figures 5-8 and 5-9 consider the progression of the mean value of the Zipfian slopes metrics in both evolution stages. This shows the effect of GA operators over the underlying metric for the two levels of evolution. High fluctuations in the Zipfian mean values over GA progression is a common characteristic over various parameterizations in the first level of evolution (figure 5-8). However, for the second level (figure 5-9) the mean values follow their alterations in a steadier manner.

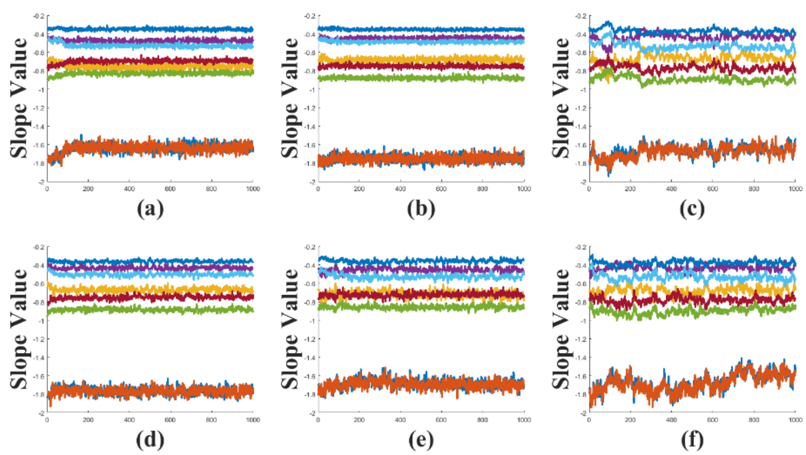


Figure 5-8. Mean of slope values for 8 features over 1000 generations using the first level format for the genotypes.

The six GA configurations (a to f) for different subfigures are presented in table 5-6. The mean values for the slopes undergo more fluctuation which show the effect of evolutionary operators in the first level of evolution.

The reason for this effect is that the Zipfian metrics calculate the number of unique events regardless of the places of their occurrence. For instance, the Zipfian metrics of pitches occurring in the sequence remains the same and independent of the different arrangements of the voices. On this account, a number of musical operations based on altering the order of elements in the sequences or swapping the segments does not have the desired effect on the fitness function values involved. The second level of evolution involves decoding the voices and applying a series of operators consisting of segment swapping and retrogration. However, they may have effects on the musical aspects of the produced audio which can be revealed in auditory tests or by other means of measurement.

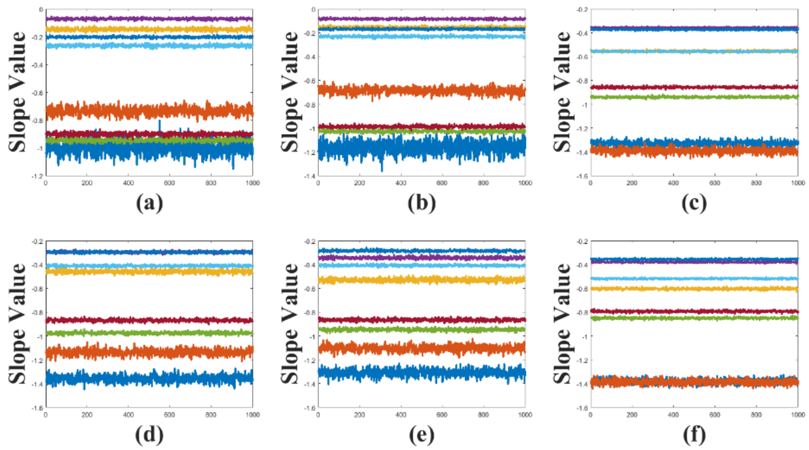


Figure 5-9. Mean of slope values for 8 features over 1000 generations using the second level format for the genotypes.

The 6 GA configurations for different subfigures are presented in table 5-6. The mean values for the slopes tend to fluctuate around steady state values unlike the previous figure 5-8. This result shows that the genetic operators in this stage do not tend to improve Zipfian metrics over time, although they may improve the sequences in the auditory test.

The experiment shows that this series of operators does not affect Zipfian values and therefore fitness values. Therefore, the fitness values tend to stay the same eventually. The defined fitness function based on Zipfian metrics is unable to reflect the changes occurring in the auditory sequences in the second level of evolution.

An auditory survey was performed at this stage of the thesis. The evaluation criteria and the result of the survey are reported in chapter 7. It is worthwhile mentioning that the audience often reported some forms of discontinuity between the musical motives. This jumping effects between the musical motives is sourced from the fact that the LPM sequences were generated by random arrangements of voices in the initial population. It appears that Zipfian metrics were not successful in detecting these effects in the audio pieces. Therefore, some of these effects remained through the future generations.

5.3 Going Beyond Exploratory and Transformational Creativity

The evolutionary algorithm agenda discussed in this chapter was responsible for experimenting with exploratory and transformational creativity. The genetic algorithm was used for evolving LPM voices sequences. The genetic algorithm operators provide the chance to navigate the search space. The members of the population undergo a series of crossover and transformation operators. At the end of each generation, the population individuals are credited with a fitness value. The competency of the evolved forms provide hints of whether some interesting forms were generated. Sometimes this exploration would end up in creating truly novel forms, which is the subject of transformational creativity. It was hoped that this GA model would provide pleasing LPM sequence. The extent of reaching the targets for exploratory and transformational creativity are further discussed in the three final chapters of the thesis.

As have been specified previously, exploratory and transformational creativity are more easier for computers to achieve, rather than combinational creativity. In the next chapter, a model is proposed based on Boltzmann machines to investigate a naïve implementation of combinational creativity. Boltzmann machines are associative memories.

This property of Boltzmann machines are employed for implementing a basic model for experimenting with combinational creativity. Cellular automata progressions would remain as a core for providing raw materials for creativity. CA progressions and Persian music pieces are trained to multimodal deep Boltzmann machine architectures. A novel series of patterns are extracted from the training data. Meanwhile, new links and associations are cultivated between the stored patterns.

5.4 Chapter Summary

In this chapter, an evolutionary framework for evolving LPM sequences was designed and implemented. This experiment was followed by a study on LPM voices regarding their aesthetical aspects, where Zipfian metrics characterized LPM voices individually and in respect to each other. The consideration in the current chapter was towards sequencing LPM voices in a musical manner by navigating a search space. In this regard, random sequences of voices were taken to be genotypes in the proposed architecture.

The fitness function is a machine learning tool trained with Zipfian measurements extracted from both Persian music and LPM sequences. Support vector machine regression model was trained to differentiate these two classes of audio while giving higher credits to the sequences which bear more resemblance to traditional Persian music. A total number of 248 Zipfian features were extracted from traditional Persian music and LPM sequences. However, not all these features were recognized to be necessary for the procedure, nor they were compatible with the purpose of the evolutionary algorithm. Therefore, further pre-processing and data cleaning procedures were involved. The desirable features in this system are those whom do not produce strong dichotomizer for SVR model but those who introduce overlap between the two classes. Strong dichotomizing features are bipolar attributes in the dataset. Preparing a database according to overlapping features between the classes would enable the SVR model not to decide solely based on the bipolar features. Otherwise, the majority of the LPM sequences would have eventually been discarded during their initial phases of evolution without giving them chances to contribute in the evolutionary process.

The experiment performed in this chapter provides guidelines for further investigations and was an important step in this research. However, some issues were associated with this designed architecture. Initializing the evolutionary algorithm by randomizing LPM sequences of voices would leave traces of vivid jumps between musical motives in the sequences, which may not be refined by the application of defined genetic algorithm operators. These dramatic effects cannot be detected by Zipfian aesthetical benchmarks as well.

Although a fitness function based on Zipfian metrics critics plays a successful role in recognizing the sequences which have more similarity to traditional Persian music, the need for covering other aesthetical aspects of the produced audio becomes more evident. To this effect, auditory surveys were performed for evaluating some aspects of the produced audio which is the subject of chapter 7. Some interesting and desirable pieces were produced according to the survey results. The next chapter proposes a design based on Boltzmann machines. The design benefits from the generative and feature extracting powers of Boltzmann machines family for capturing the dynamics of CA, and achieving the internal representations of traditional Persian music. The architecture in the next chapter is responsible for experimenting with combinational creativity.

Chapter 6. Recognising LPM patterns using Boltzmann Machines

This chapter revolves around the design and implementation of an architecture which applies Boltzmann machine families in the Liquid Persian Music composition system. The memory storage, generative and data retrieval capabilities of this class of networks are the focus of attention throughout this chapter. Boltzmann Machine (BM) family models are used as pattern-matching tools for extracting features from CA progression. The information retrieval capability of BM is not limited to CA. It is also applied for extracting patterns from traditional Persian music. The LPM system can then benefit from the generative aspects of BM trained with CA and Persian music data whether individually, and in a connected architecture. The design in this chapter includes a Multimodal Deep Boltzmann Machines (Multimodal DBM) structure which has multiple channels for learning data with various modalities. Two Multimodal DBMs are trained with CA progressions and Persian music data. These two structures are later applied for generating patterns as materials for populating synthesizer parameters. The experiments in this chapter are implemented in Matlab and with inspiration from the software and libraries provided in (Kauffman, 1969; Salakhutdinov, 2012; Taylor, 2010).

6.1 The Approach for Applying Boltzmann Machine in LPM System

Recognition of various phenomena, environments, and objects are possible by their identifying features. This is one of the possible ways that enables the human brain to remember and store information (Freeman & Skapura, 1991). The underlying mechanisms in human brain, systematically recognizes the patterns. The describing characteristics of various phenomena do sometimes overlap and that is often how the similarity or linkage between objects or situations are recognized. One thing that usually happens is that a phenomena or object reminds a sequence of memories of another phenomena or object. This occurs when the features of a memory resonates with associated or similar characteristics of other memories. So many artefacts were created from being inspired from a phenomena and introducing links with other domains. This fact is more prominent in studying an artist's work of art which is directly influenced by the viewpoints and mentality of the creator.

Miscellaneous subjects produce different meanings for people depending on their backgrounds and the types of associations they produce with their prior knowledge. An informal proverb says that learning becomes possible by establishing associations with previous knowledge. There are several research projects about relevance of prior knowledge in learning; for instance please refer to (Hein, 1991). Therefore, the various people's perspectives and viewpoints refer to the associations taking place between upcoming data and previous knowledge. According to Boden, associative creativity takes place when new links are discovered between elements that were not directly linked.

One of the characteristics of deep learning techniques is that they can extract features in various layers (Salakhutdinov, Tenenbaum, & Torralba, 2013). The representations of data in various layers provide means for having further generative powers and further chances for recognizing the possible links between items. In this work, the representations that are available on each layer can be considered as new artefacts or as new creations.

The generative nature of Boltzmann machines in the LPM project application, originates from the fact that Boltzmann machine family models have a stochastic nature. Many types of neural networks have deterministic characteristics meaning that under the same circumstances the output of the system would remain the same (Briot & Pachet, 2017). By providing the same inputs as training data in various implementations, the system outputs would be identical. However, the Boltzmann machine families have stochastic nature which turns out to be beneficial as a part of a system which generates artefacts. There are various possible equilibrium states existing for such systems. These phenomena can be considered to be like crystals being settled on their different facades. Richards in (Wulff & Hertz, 1993) has first suggested the resemblance of memories to be like crystals. The network is like a plasma which undergoes a crystallization process. Respecting the initial configurations of the plasma like system, the associations of the elements of the system are structured during the transition process from plasma to solid. Using the terminology of solid crystals for Boltzmann machines needs to be approached with caution. Once the BM reach their equilibrium status, the states of the neurons may still keep on oscillating (Hinton, 2010). Therefore, the neurons status are not necessarily frozen or solidified. However, the crystallization metaphor is employed for a better clarification over the training paradigm taking place in BM family models.

Anytime the crystal is formed, the association of the elements would be different to each other. The angle of looking into the crystal determines how the proportions and relations of elements are perceived. Looking through the crystal from different angles reveals different associations between the elements of the crystal. Likewise, collecting the connecting weights from each of the hidden units in a Boltzmann machine reveals various patterns. Dissimilar results are achieved depending on the choices of neurons for the output of the system.

Cellular automata display stochastic yet determinable dynamical nature. Therefore, they can be treated and stored as memories. The current viewpoint allows saving the CA dynamics in the selected BM configurations. In this respect, Boltzmann machine families would provide a handful of pattern-matching tools. The pattern-matching outputs are considered as the patterns revealed from looking through a crystal from different angles.

In this chapter, the units in the Boltzmann machine families are used as new types of pattern-matching tools. This would allow a large collection of patterns which can be used for generating audio. Moreover, the resonating characteristics of Boltzmann machine families as associative memories are employed for replicating characteristics similar to some of the possible mechanisms happening in human brain while creating an artefact (based on combinational creativity). A hypothesis is examined in this chapter to see how computational tools might be useful in generating Persian music based on a naïve model of combinational creativity. Boltzmann machine families are employed for testing this hypothesis.

6.2 The proposed architecture

The application of Boltzmann machines in this proposed design is manifold. One revolves around the generative nature of Boltzmann machines which are stochastic. Another is about their feature extraction capability in capturing high level, and complicated features. Extending the architecture of RBM (Restricted Boltzmann Machine) into DBM (Deep Boltzmann Machine) by building further hidden layers would provide additional feature extractors. Therefore, applying RBM model for extracting features has the functionality of pattern-matching rules in the LPM system as described in chapter 4. The features extracted from CA progressions are collectively stored in the memory of BM family models. Each of the hidden units and their connecting weights represent a set of extracted features. The hidden units act as standalone pattern-matching tools. Each of the units in different hidden layers would provide a set of patterns, which could later be used to populate the synthesizer's parameters. The correspondence between the visible and hidden units are representable in the form of a matrix; the dimensions of which are equivalent to that of the initial training data prepared for the visible units. The initial configuration of CA, the number of cells and history of iterations, all have contributions in the final weights. Therefore, various patterns can be generated by altering the initial parameterization and configuration of the system. This would have effects on the generative power of the LPM system in the matter of the number of patterns produced, which are later mapped to the musical domain. In the section related to capturing the dynamics of CA using RBM, and CRBM (Conditional Restricted Boltzmann Machine) the procedure for achieving new pattern-matching tools are explained in more detail.

In a different experiment in this chapter, Persian music are segmented and stored inside Multimodal DBM structure. Multimodal DBMs are mostly employed for data which inherit multiple modalities, for instance the video and audio represent two modalities in movie data bases, where each frame of video is associated with the related audio. In our application, in order to store music inside Multimodal DBM, each of the constituent dimensions of music (note pitches, and duration) are presented to one of the channels in the Multimodal DBM structure. The whole architecture is responsible for providing fused representations of note numbers and durations. The related details about the procedure taken for storing Persian music in Multimodal DBM structures are presented in section 6.2.2 (titled: Storing Persian Music in Multimodal Deep Boltzmann Machines Multimodal DBM).

Cellular automata progressions are also trained to a separate Multimodal DBM structure. There are initially no association between Persian music and CA progressions. However, regardless of the Boltzmann machine family type used for each of those pathways in the multimodal architectures, an additional modality is added as a pathway to each of the CA and Persian music Multimodal DBM structures. This modality is trained with Zipfian metrics of the related data in the other pathways. The Multimodal DBMs trained with CA, and Persian music are later joined with each other via their Zipfian pathways.

6.2.1 Extracting Cellular Automata Features using Boltzmann Machines

The Restricted Boltzmann Machines (RBM) were applied in various ways towards extracting features from CA progressions. One of these is to take 88 unique one-dimensional behaviours over certain numbers of iterations and train them using a RBM model. The CA progressions were taken as static images with white (on) and black (off) cells. The cells were mapped to the neurons in the visible units of RBM model representing on/off states.

There are some queries that need to be addressed about the quality of the training performed; one is about the quantity of CA cells that should be presented to visible units in every visit. In other words, the number of cells in the initial array and the count of CA progressions, which attend the training agenda, should be considered. Further prevailing questions refer to the configurations of the initial seeds which the CA progressions should be built upon. The initial seed influences the upcoming patterns generated in a CA progression. Starting from different initial configurations, various basin of attraction collections are obtained. In this experiment only one possible configuration were taken into account as the initial seed. This was performed to limit the size of the training set and saving the processing time. The increasing in the number of provided patterns for a RBM would require increase of the number of hidden units. Changing the initial configurations of CA gives different training samples. This was assumed to be one of the ways for getting various results in different implementations.

The initial seed was taken to be an array with the size of 25 cells. The total number of CA iterations allowed were taken to be 24. By adding 24 CA iterations to the initial seed, the 25 progressions will be achieved. In the implementations of the experiment with RBM, a total number of 88 main CA behaviours were taken into account. The training set consisted of 88, 25*25 metrics. This produces 625 visible units. The training was performed in batches of 20 with overall number of 200 epochs. It is notable that the same visible RBM units should be associated with the same cell locations in the CA progressions throughout the experiment. The order of the presented cells should be kept constant while visiting all members in the training set.

The number of hidden units was taken to be 400 units (figure 6-1), and 1000 units (figure 6-2) in two different implementations in order to examine the nature of the patterns extracted with different numbers for hidden units. Figures 6-1 shows a portion of the weights between the visible units and hidden units of the RBM with 400 units. Each of the squares represents all the connecting weights to a hidden neuron. Figure 6-2 depicts a portion of the weights between the visible units and hidden units of the RBM with 1000 units. The weights in both cases are in fact real numbers, some of which are negative. The negative numbers are represented as black pixels in the images; however, there is a spectrum of negative numbers, which have not been equitably presented in the figures.

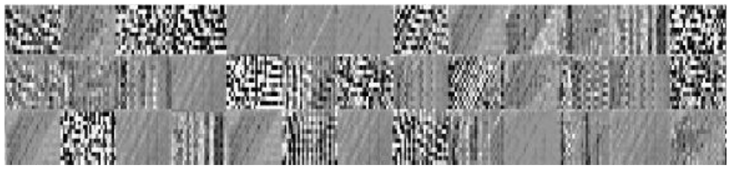


Figure 6-1. Looking into a small portion of the memory of RBM after the training procedure with 400 hidden units.

Each of the small squares are associated to one of the units in the RBM architecture and all its connecting weights. These squares are a portion of 65 units out of 400 hidden neurons.

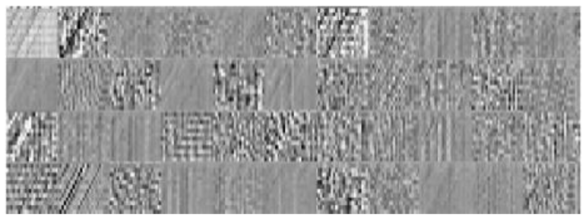


Figure 6-2. Looking into a small portion of the memory of RBM after the training procedure with 1000 hidden units.

Each of the small squares are associated to one of the units in the RBM architecture and all its connecting weights. These squares are a portion of 44 units out of 1000 hidden neurons.

The comparison of figures 6-1 and 6-2 imply that by raising the number of hidden units, the connecting weights to the different neurons becomes more similar. Visual inspections on the extracted patterns show that the patterns extracted from a trained RBM with 400 hidden units have more contrast. However, the patterns extracted from a trained RBM with 1000 hidden units tend to lose their contrast and become more uniform in colour or more blurred. Greater contrast between weights should favour greater variations in the musical patterns. The nature of the extracted patterns will have effects on the generated audio. For instance if the mapping procedure to the musical space involves selecting the connecting weights to one neuron as the musical pattern, then the second configuration for the number of hidden units is more likely to produce a more monotonous (less musical) audio sequence.

6.2.1.1 Extracting Cellular Automata Features using Conditional Restricted Boltzmann Machines

The CA progressions have been applied in the form of certain configurations to another BM model called conditional restricted Boltzmann machines (CRBM). CRBMs are often applied for modelling time series data. Performing this experiment was inspired from the concept of basin of attraction models which are employed for studying the dynamics of CA (there are some literatures available about the attempts made for modelling the dynamics of CA by the help of neural network models (Tanaka, 2016)).

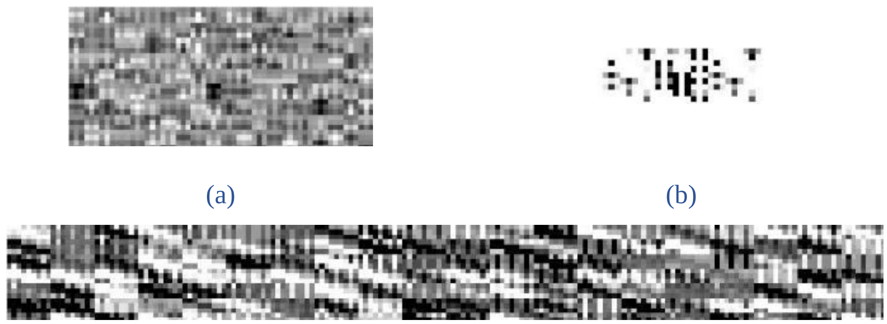
The behaviours of each of the CA rules are studied through a set of basin of attractions. The basin of attraction field consists of all the possible states transitions for a specific CA configuration. Therefore, in a CRBM, an arbitrary but fixed number of CA progressions can be conditioned on an arbitrary but fixed number of their previous transitions. The remaining issue would be the labour intensive task of identifying all the various topologies available in a basin of attraction field for configuring the CRBM model. The truth is that the basin of attraction models do not follow the same structure throughout all the basin of attractions for a CA rule (there might be a basin of attraction which consists of a single node or one with miscellaneous branches possibly connected to other nodes in a circular path). Moreover the topologies vary among different CA rules, since each of them follow a separate structure and that makes the training complicated for CRBM model. Hence, it is not possible for taking one basin of attraction model as a unique standard for establishing it as a base model for training CRBM. Production of basin of attraction graph models are addressed elsewhere in

(Wuensche, 2009). The suggested approach in here would save the computations regarding the construction of the basin of attraction fields for different CA rules.

For this experiment, all the CA initial configurations were explored and the training is based on the history of CA progressions as time series data. All the transients up to 20 iterations were considered starting from the initial seeds. However, this approach raises another issue. Visiting the transitions by this technique would increase the chance that the branches or the basins may be visited several times due to the nature of the CA behaviour in some of the CA rules. However, some of the transitions may be visited fewer times as occurring less in the training data. For example, the rules in the first and second CA categories (described in the section related to cellular automata in the third chapter) will soon enough produce a repeated pattern. Specifically the rules in the first category would converge to homogeneous patterns irrespective of the initial seeds. This invokes the existence of bias in the training set. Therefore, attempts were taken into account to transform the dataset which initially consisted of 100 number of iterations for different CA rules, in order to avoid homogeneous or repetitive patterns in the training set. For instance, the dataset may contain a large number of transients which have fallen into the homogeneous or heterogeneous patterns. This denotes that all the states will evolve to the same pattern or cyclic patterns in the upcoming iterations. For overcoming this problem, the number of iterations involved for the transients should be considered more cautiously.

The determination on the number of CA progressions for training, are based on the CA behaviours and whether they are in the classes of fixed, cyclic, chaotic, or complex. On this account, the number of CA iterations for the rules in the fixed, and cyclic groups were limited.

The number of involved iterations in the first CA category (fixed) was cropped by 75 percent (from the initial 100 iterations), while this number is 50 percent for the second category (cyclic). Instead, more CA iterations (100 iterations) were allocated to the progression of rules in the chaotic and complex behavioural classes.



(a) (b)

(c)

Figure 6-3.Looking into a small portion of the memory of CRBM after the training procedure.

In the configuration of the CRBM for this experiment, Figure (a) relates to the weights between the visible and hidden units, figure (b) demonstrates the weights in between past and visible units, while figure (c) shows the weight values between past and hidden units.

In this experiment, 88 CA behaviours were studied for all the possible initial transitions. Various configurations can be assumed for employing CRBM model to learn CA dynamics. For instance, the state of the cells in each iteration can be trained to CRBM and be conditioned to the state of the cells in previous CA iterations. A more complicated configuration, involves conditioning one or more CA progression to an arbitrary number of previous iterations. Whatever configurations been determined, it will stay constant for selecting the CA cells from different CA rule progressions. The configuration chosen in this experiment uses a series of 3*3 metrics consisting of the cells in 3 consecutive iterations. These cells were conditioned on their three past iteration ancestors, regarding their history of appearance. This would imply a total number of 9 visible units conditioned on 27 past units.

The number of hidden units was selected to be 400. Figure 6-3 demonstrates examples of the weight metrics between visible-hidden, visible-past, and present-hidden units mapped to greyscale images for illustration purposes.

6.2.2 Storing Persian Music in Multimodal Deep Boltzmann Machines

What has been described so far was related to training the CA progressions to Boltzmann machine families. This has attained an arbitrary number of pattern-matching tools depending on the number of units in the layers of Boltzmann machine family architecture. This section represents the details for implementing the Multimodal DBM for learning Persian music data.

The multimodal structure related to Persian music consists of three pathways. The training samples are obtained from segmenting Persian music into short frames. There are three sets of values related to the musical pieces. Two of the sets of values are associated to the occurring musical events including note frequencies and durations. The other pathway consists of Zipfian metrics extracted from the related Persian music segment (As have been described in the previous chapters, Zipfian metrics were applied as aesthetical critics and guidelines in creating artefacts). The pathways were prepared for accepting the frequencies of the occurring notes, the musical duration of the notes and the Zipfian metrics associated with the musical frames. These values were applied for retrieving pieces with similar characteristics, and for generating audio.

The pathways consist of a two layered DBM. Each of the layers provides a higher representation of the data presented in the previous layers. The state of neurons in the hidden layer of each RBM is the input to the upper RBM in the layered stack. A collection of stacked

RBMs in each of the pathways would attain a deep belief net architecture for the pathways. However, the overall structure cannot be considered as a Multimodal DBM unless mean-field training is performed. The pre-training of the Multimodal DBM consists of training each of the individual RBM layers. The pre-training phase initializes the weights and prepares the network for mean-field training phase. This indicates that the mean-field training would converge faster after the pre-training is performed rather than having random initialization (Salakhutdinov & Hinton, 2012). The formulation and algorithm for training Multimodal DBMs are provided in chapter 3, and appendix A. Mean-field inference consists of bottom-up and top-down passes. In a Multimodal DBM architecture, a joint representation is provided for the data in different pathways. This means that the manifestation of patterns in one pathway is conditioned on the occurrence of related modalities in the other pathways. From the training point of view, this condition becomes possible by assigning high probability to the concurrence of the events (training samples) associated to one another. The mean-field training is performed for the overall structure of the Multimodal DBM.

The training samples include frames of Persian music pieces each of which were limited to a number of 20 musical events (musical notes). Extracting such features from Persian music were discussed in the previous chapter, where a MIDI-like database was obtained for Persian music database. The note frequencies and the corresponding note durations for each frame were presented to each of the pathways. The third pathway was populated with Zipfian measurements of each of the frames.

The units in all the Multimodal DBM pathways were selected to be binary units. Therefore, the training data in all the pathways were converted to binary values. Another option would be the application of Gaussian units instead of binary ones. The frequencies of the extracted notes from Persian music as training samples for the system have a range between 158 to 1396 Hz. The note durations vary between 1/16 and 2 seconds. For converting these values to binary domain, the following considerations were taken into account. A total number of 22 bits were allocated to the frequency binary points values. The duration values consists of 14 bit binary numbers. Having training samples consisting of 20 musical notes each, the binarization procedure would result in having 20*22 number of visible units in the note frequency pathway. This number would be 20*14 for the note duration pathway.

The Zipfian metrics from Persian music frames were extracted in the same way as discussed in the previous chapter. The Zipfian features were extracted from frames of Persian music each consisting of 20 notes (unlike the experiment in the previous chapter which extracted Zipfian metrics out of 20 second segments of Persian music). The number of attributes in the Zipfian metrics are 30 features for this experiment. Cleaning on the Zipfian metrics have been performed. The attributes which were related to onset time were discarded. This is due to the fact that only the note duration and pitch frequencies were targeted in this experiment. The attributes which consisted of notable number of zeroes, NaN, or minus infinite among the records were also eliminated. This cleaning left 30 attributes in Zipfian metrics. The Zipfian slope values are negative decimal point numbers. The fraction part needs to be converted to binary with a high precision. 16 bits were allocated for storing the Zipfian binary values. This results in a total number of 30*16 input units for the Zipfian channel.

The number of hidden units in the layers of different channels were taken to be 500 units. However, 500 is relatively a small value considering the high number of patterns to be learned. Therefore the training database was divided into smaller sets each consisting of 100 samples. In each implementation, the user can select which part of the big dataset is going to be used for the training of the Multimodal DBM. Alternatively, the trained weights from various implementations can be stored and later be used for benefitting from the miscellaneous patterns that were generated. This case can be assimilated to the situation where we have several numbers of Multimodal DBMs. It should be mentioned that the mean field training with 500 hidden units in each channel with the limited number of training samples took notable amount of time. The training was performed in batches of 30, and 200 epochs were involved.

Later in this chapter, the Multimodal DBM structure related to Persian music is employed in accordance with another Multimodal DBM trained with CA data. The Multimodal DBM related to CA has a similar architecture to Persian music Multimodal DBM. These two Multimodal DBMs resonate each other for attaining further patterns as will be discussed in the remainder of this chapter.

6.2.3 Design of the Multimodal DBM Trained with CA Progressions

Section 6.2.1 demonstrated the application of RBM and CRBM for extracting features from CA progressions. In this section, an overview will be presented towards training a Multimodal DBM structure with CA data. The structure of the CA Multimodal DBM is similar to the Multimodal DBM used for training Persian music, and has three pathways. The three channels were trained with data associated with CA. The first channel is employed for performing the mappings to note frequencies. The second channel is used for mappings associated with notes durations. The third pathway was trained with the CA data Zipfian metrics.

The Zipfian attributes selected for CA were the same Zipfian attributes calculated for Persian music pieces. In fact, the Zipfian modalities in the CA Multimodal DBM structure are configured identically to the Zipfian pathway from the Multimodal DBM trained with Persian music. One may ask about the possibility of calculating Zipfian musical measures based on musical terminologies from CA which do not presently have any musical characteristics. The approach that is taken for performing the mapping to the musical space is suggested to be a basis for extracting Zipfian metrics associated with CA progressions as well. There are various agendas for performing the mapping, therefore the selection of the inputs made available for extracting the Zipfian metrics would vary. Two different paradigms are suggested for extracting Zipfian metrics from CA progressions in this experiment. Other possible varieties are subject for future work. In the following, these two suggested mappings are further discussed.

In one paradigm an arbitrary hidden unit or group of units were selected. These hidden units play the role of pattern-matching tools. The values of the connecting weights to the hidden units are considered to be pattern-matching outputs. The weights connected to the chosen hidden units were taken into account and mapped to musical note frequency and duration. The Zipfian metrics were calculated according to the combinations of note frequencies, and durations. The frequencies and duration values together with their related Zipfian metrics were stored as training data to be provided to the Multimodal DBM structure. The final musical results in this approach completely depend on the selection of the hidden units mapped to the musical space. In other words, the way the extracted patterns are interpreted and projected to the musical domain determines the quality of the generated audio patterns.

In another paradigm, the CA progressions were trained against their Zipfian metrics obtained from pattern-matching rule values (Twenty pattern-matching rules from LBM). In the study in the previous chapters, it was demonstrated that the pattern-matching rules introduced in LBM system were employed to populate the synthesizer parameters. It has been shown how these values were tailored in a sequence in order to achieve the initial grounds for Zipfian metrics calculations. In this experiment an arbitrary pattern-matching rule number (rule 20 for instance as described in chapter 4) was chosen for extracting the Zipfian metrics from CA progression and performing training on the Zipfian pathway. In this paradigm, the CA Multimodal DBM note frequency and duration branches were trained with CA progressions. The training sets for these two branches contained the same elements, however, the presentation order in the training sets were randomized. This situation conditions random CA rules with each other in the two Multimodal DBM branches. For example, rule 32 will be associated with rule 55. Later the data related to rule 32 will be mapped to note frequencies and CA rule 55 extracted data will be mapped to musical notes durations.

6.2.4 Associating Persian Music, and CA Multimodal DBM structures

Associative memories resemble the particularities of human brain in memorizing things in association to each other. This is similar to the capability of human brain where similar ideas resonate each other and associated phenomena remind of one another. The computational model of this phenomenon is manifested in Hopfield networks with storage ability. BMs can be considered as extended versions of Hopfield networks that inherit this property as well.

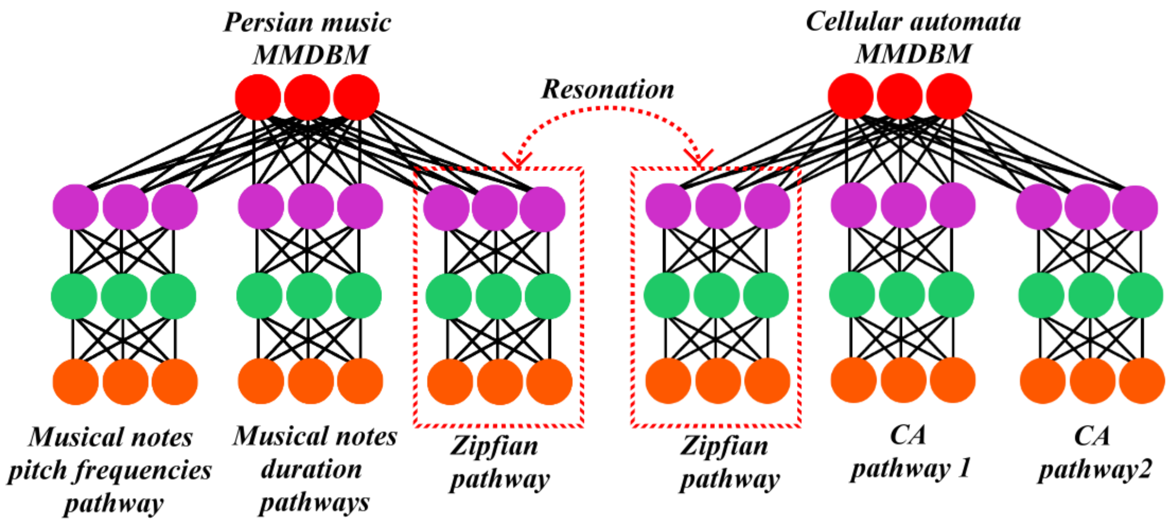


Figure 6-4. Resonation taking place between the Zipfian pathways of two multimodal deep Boltzmann machine structures.

The pathways pertaining to Zipfian metrics are employed for resonating related data in the other pathways. This model is inspired from associative memory architectures. The Zipfian pathways in two Multimodal DBM architectures follow the same configurations in the matter of the number of visible and hidden units and the types of the units. However, the Zipfian pathways are independent from each other. Once the values in each of the pathways are populated, they are capable of evoking the associated modalities in such structure. This phenomenon was employed for triggering the patterns with similar Zipfian metrics in CA and Persian music multimodal architectures. Figure 6-4 illustrates the resonation occurred between the two Multimodal DBMs via their Zipfian channels.

One of the possible paradigms in this architecture is to populate the CA pathway to obtain the associated Zipfian metrics from the related pathway and then use it for populating the Zipfian pathway from Persian music multimodal structure for triggering the related frequency and duration modalities in Persian music Multimodal DBM structure. This procedure can be performed in the alternate direction as well; the Zipfian metrics related to Persian music patterns or training samples can trigger patterns associated with CA progressions with similar Zipfian metrics. This model is used to provide links between items that have not been previously linked before in a direct manner.

6.3 Musical Experiments

Various approaches are possible for applying the RBM, CRBM, and Multimodal DBM structures discussed in this chapter for musical applications. Here four of the possible ways for mapping the results to musical space are discussed. The outputs of the suggested mappings are embedded in the surveys (auditory tests) in the next chapter.

Approach one: Each of the hidden units in the RBM, and CRBM architectures trained with CA progressions can be considered as a pattern-matching tool. The connecting weights to each of the units embed patterns from CA progressions. The values of the weights can be directly used and mapped to musical space. In this approach, the number of pattern-matching tools are determined by the number of visible and hidden units and the initial configurations of CA.

Approach two: the Multimodal DBM model trained with Persian music data are applied for performing the musical mappings. This approach is similar to saying that we have various interpretations and patterns extracted from musical pieces. Binary representations of the musical pieces are stored inside the weights connected to the units. The weight values can be employed for performing the mapping. Moreover, the states of the hidden units in various layers can be used for musical mapping.

Approach three: By clamping musical segments on the visible units in Multimodal DBM, the related patterns can be obtained from other channels. For instance by once providing note durations in one pathway, the related note pitches can be inferred from the other pathway. Since the system has a stochastic nature, the values inferred in the other pathways might be different every time the system is up and running. The provided information to be clamped on different channels can be selected from the training set, or be selected from other musical sources. On this account, different sets of results can be obtained.

Approach four: Data can be provided for the visible units of a channel in one of the Multimodal DBM structures (the Multimodal DBM trained with CA data, or the one trained with Persian music data). The Zipfian metrics related to the activated Multimodal DBM will call patterns with similar associated data in the channels of the other Multimodal DBM. The Zipfian values observed in one of the Multimodal DBM, populates the visible units of the Zipfian modality in the other Multimodal DBM.

6.4 Chapter Summary

Boltzmann machines are associative artificial neural networks, which have stochastic units. RBM achieve binary representations of data provided in their visible layer. Higher-level representations of data can be obtained by building extra layers on top of RBM in order to construct a deep Boltzmann machine architecture. In this chapter, Boltzmann machine families were applied both as pattern-matching tools, and as associative memories. RBM and CRBM were employed for extracting patterns from CA progressions. It was shown that each of the units in the hidden layers provides a pattern-matching tool. These pattern-matching tools were later used for generating audio. In this project the capability of Boltzmann machine families as pattern-matching tools were in the attention. Studying the capacity of Boltzmann machine families and their success in storing the CA progressions as training data are out of the scope of this thesis. They have been looked upon as pattern-matching tools for generating raw materials for populating the synthesizer parameters. Altering the configurations for obtaining various results was emphasized throughout the chapter.

Two Multimodal DBM structures were presented which were separately trained with data from Persian music database, and from CA progressions. Each of the Multimodal DBM architectures consists of three channels, which stand for note pitches, note durations, and Zipfian metrics. The two Multimodal DBM architectures provide further approaches for mapping data to musical space. The pathway associated to Zipfian metrics are used as a resonator for triggering the associated patterns in the other pathways. The idea is that similar Zipfian metrics would trigger the related patterns from CA pathway and Persian music pathway. The suggested architecture provides a pathway for exploring the possibility of combinational creativity (in a small scale). Using Boltzmann machine families as associative memories were made a bold target for assimilating some of the possible mechanisms that happen in human mind for creating artefacts based on combinational type of creativity. In the previous chapter, evolutionary algorithm was responsible for navigating the spectrum of exploratory and transformational creativity. In the next chapter, auditory surveys are designed and published that helps us to estimate the extent of accomplished thesis targets.

Chapter 7. Human Evaluation: Online Surveys

The previous two chapters were dedicated to the design and implementations of algorithmic composition systems. This chapter provides the opportunity for studying the human evaluation of the audio generated by the systems in the conducted experiments. This will give an insight to the quality of the performance of the designed systems and reveal further steps for refining the architecture of the systems and suggesting the future research directions. Evaluation provides the chance for determining the progression of the research against the pre-specified goals of the thesis.

In this chapter, three surveys are designed and published. The criteria for the evaluation of each of the surveys are different from each other; however, there is some overlap between the evaluation criteria of the three surveys. The first survey is based on the evaluation of the audio output from the architecture presented in chapter five (The first survey is older than the consecutive surveys (2 and 3), and was published before performing the experiments in chapter 6). The second survey targets the audio generated from the models presented in chapters 5, and 6. The second survey makes use of extended set criteria for evaluation. Both of these surveys (1 and 2) target the same group of audience. The third survey is the most comprehensive in its coverage with criteria specifically selected for professional audience. The audio samples for the third survey cover the audio pieces generated from both architectures in the two previous chapters 5, and 6 (No separate surveys have been designed for chapter six, Although, the sets of the algorithms are different from the ones used in chapter

5. The genetic algorithm in chapter 5 was responsible to experiment with exploratory and possibly transformational type of creativity, while the Boltzmann machine families employed in chapter six experimented with combinational type of creativity. Both the compositional algorithms build the project for creating Dastgāh-like music. The outputs of both of the systems are embedded in the survey. In this chapter by referring to the evaluation of the system, we mean the evaluation of both of the systems in chapters 5, and 6 as a whole).

7.1 Introduction on the Evaluation of Creativity in Creative Systems

The beginning of this section presents an overview on the necessity of evaluation of the results of the systems generating creative materials. Some of the challenges in the evaluation pathway are described. General standard approaches for evaluating the creative systems are briefly discussed and the criteria for assessing the results of the computational composition models in the previous chapters are specified.

The establishment of evaluation benchmarks have been identified as a difficult task throughout the history of computational creativity (Cardoso, Veale, & Wiggins, 2009). This is partially due to the lack of a comprehensive definition for creativity. Evaluation of many artefact generator systems based on standard criteria has not always been the focus of attention of the founders of those systems. A part of this problem goes back to the nonexistence of established standardized evaluation methods at the time. Nonetheless, scientists and philosophers in the area of computational creativity have not yet claimed a full-scale standard evaluation method. However, this does not imply that one is immune to not following the evaluation procedures at hand.

The evaluations, which are designed based on standard milestones, provide the opportunity for comparing various systems according to the specified criteria. This comparison may identify the potentials of the initial tools and computational models involved. The evaluation would provide guidelines for the development of systems. On this account the design of standard evaluation methods, become important.

A note to make here is that there might be some existing bias in participants' evaluations. People may have bias towards the design of the survey, as well as the length of the survey. Other influencing factors that may contribute to bias are the judge's moods, the time in the year when the survey is published, the critic's background knowledge in general. What people expect from the system may also affect the way people rate their answers. More importantly, the bias might be associated with the fact that the presented artefacts are machine improvisations. It might be a good idea to raise the consciousness of the audience about their possible existing bias towards machines generating music. In (Moffat & Kelly, 2006) an investigation was performed towards people's bias against computational creativity and music composition. Moffat and Kelly found that people's ratings on the creativity of machine-generated music is influenced by their conscious and subconscious prejudice towards computational creativity. Identification of people's bias has been practiced by Jordanous in (Jordanous, 2012b). The audience were asked to answer some questions related to their mind-set towards computers being creative. This factor was taken into account in the design of the second and third surveys where the audience were asked to answer to some questions related to computational creativity and Persian music.

7.1.1 Some Evaluation Methodologies

During the past decades, the attempt has been towards standardizing the evaluation methodologies. In the following, some of the most significant standard methodologies for evaluation are demonstrated. The remainder of the section considers some main standard evaluation techniques. The suitability of each of these evaluation techniques are justified in the surveys design sections. The surveys designed in this chapter cover some domain dependent and domain independent aspects for evaluating creativity and in regards to the targets of the thesis. The first survey is based on human evaluation of creativity. The second and third surveys will cover human opinions on creativity and Ritchie's evaluation methods.

- Turing-style Test

One of the famous directions often followed by researchers for evaluating the results of their systems is the conducting of a Turing test. On this basis, the measurement of the success of a system is according to the opinions of human subjects on the origins of the productions. For instance, the agenda of a Turing test for evaluating a system whose artefacts are musical pieces is as follows: The human subjects distinguish the machine-composed pieces from the human composed ones. Higher scores are assigned to undistinguishable pieces. The Turing tests usually suffer from biases caused by selection of testing samples. However, Turing test was widely employed and has been very popular since its introduction. Turing test attain general measurements based on human evaluation. More scrutinized assessment criteria are often required for determining the creativity of a system. There are methodologies applied for evaluating the output of different systems in terms of creativity.

Unlike Turing test which can be considered as a *blind* test, the following methodologies usually consist of informing the subjects that the artefacts are machine generated:

- Human Outlooks and Judgements On Creativity

In this methodology, the human judgements on creativity are taken as ground truth by assuming that people by nature recognize the difference between creative and non-creative items (Zhu, Xu, & Knot, 2009). Many researchers have benefited from the human viewpoint on creativity in the evaluation of their works. A definition on creativity can be presented to the people taking the survey and they are later asked to rate the system generated artefacts (Jordanous, 2012b). One of the possible definitions of computational creativity is as follows: "A behaviour or action generated by computer is said to be creative if the same action enacted by human would appear to be creative."

- Ritchie's Creativity Evaluation Criteria

Ritchie provided a framework for evaluating the creativity of a system based on its produced artefacts (Ritchie, 2001, 2007). Ritchie's methodology evaluates aspects of the *typicality* and *quality* of the outputs of the system:

- To what extent is the produced item an example of the artefact class in question? (Ritchie, 2007)
- To what extent is the produced item a high quality example of its genre? (Ritchie, 2007)

Ritchie determined 18 criteria for creativity (Ritchie, 2007) which are formally represented in set-theory form. The suggested criteria can be customized for different applications. The parameterization and weightings of the criteria may be different from each other for evaluating outputs from different systems. Likewise, some of the 18 criteria may be discarded during the evaluation if they are found to be impractical or redundant. The Ritchie's framework for evaluation is used in the design of the second and third surveys in this chapter.

Ritchie's evaluation criteria for creativity are presented in sections 7.3.6 (in table 7-3), and 7.4.5 (in table 7-5) where they are analysed for the second and third surveys.

One of the drawbacks from Ritchie's framework is that it lacks guidelines for setting up the parameterizations of the different criteria. This has been left to the choice of the evaluators. Ritchie suggests using the data gathered from human assessments to deduce the parameters for implementing the evaluation criteria (Jordanous, 2012b). There are application examples in (Jordanous, 2012b; Pereira, Mendes, & Gervas, 2001; Ritchie et al., 2007) available, which implement Ritchie's evaluation criteria.

- Pease's Evaluation Tests for Creativity

Pease et al. in (Pease, Winterstein, & Colton, 2001) suggests a wide area for making judgements on the creativity levels of a system. This criterion not only investigates on the outputs of the system, it takes a scrutinized look over the input of the system and the process involved in the system as well. Pease suggested a collection of evaluation criteria to be measured. Some of the evaluation criteria include perceived novelty, surprisingness, and emotional response.

- Colton's Tripod for Evaluating Creativity

Colton's (Colton, 2008) evaluation methodology highlights the evaluation of the creativity process. Colton suggested that a creative process should possess three behaviours *Skill*, *Imagination*, and *Appreciation*. Skill demonstrates the extent the system is skilful in a certain domain, Imagination demonstrates how the system can come up with variety in their creation, and Appreciation governs on how the system can think and evaluate during the process.

- SPECS (Standardised Procedure for Evaluating Creative Systems)

Jordanous in (Jordanous, 2011, 2012a, 2012b) proposed SPECS evaluation methodology. It demands the evaluators to specify the creativity requirements in the domain they are working with. As a standardized methodology, SPECS consists of three main steps:

Step 1: Determine domain dependent and domain independent aspects of creativity

Step 2: Standardize the determined aspects of creativity for evaluation

Step 3: Perform evaluation tests

Jordanous found 14 criteria for measuring the creativity, obtained by mining through previous publications about creativity in a wide discipline. These fourteen criteria can be used in SPECS evaluation method and are available to other researchers. Some of these criteria include social interaction and communication, intention and emotional involvement, domain competence, active involvement and persistence, variety divergence and experimentation, dealing with uncertainty, originality, and independence and freedom.

7.2 The First Public Survey

7.2.1 The Criteria and Design for the First Survey

The first survey was designed for assessing the progression of the project. The evaluation criteria in the first survey addressed musicality, Dastgāh-likeness, and pleasantness of the evolved sounds. At the time of the first survey, the author was not intending to evaluate the creativity of LPM system itself. She performed the test for getting feedback on the produced audio for progressing her research as described and continued in the sixth chapter. The results from the first survey contributed in the progression of the thesis.

Although the first survey is about two years older than the consequent ones (second, and third survey), the author decided to keep it in the thesis.

Designing the process of evaluation requires determining the evidence for the chosen evaluation criteria. The evaluation can take place as quantitative or qualitative judgements. The quantitative methodology involves gathering the evaluation values and performing statistical computations on the test outcomes. The qualitative assessments are usually performed by means of feedbacks from surveys. The qualitative survey accompanies our evaluation methodology, which provides us with some feedback on the evolved sounds.

The first survey was designed in survey monkey and has 10 questions, 9 of which consist of audio evaluations and the last question asks the participants to provide their opinions and feedbacks on the presented material. The audio files were made available in sound cloud where they were hosted and later embedded in the survey monkey questionnaire design. The questions try to gather evaluations on the produced audio based on their musicality, whether the audience liked them and if they were Persian-like. In the beginning of the survey the audience were presented with some Persian music (of about one minute length) to give the participants some clue about the Persian likeness of the produced audio. Appendix G provides further details on the first survey design.

7.2.2 The Results of the First Survey

The survey invitation was circulated to a list of postgraduate students and staff at the University of Hull. 35 participants took part in the survey. Only a few questions were skipped during the survey. The feedback areas were filled out by 20 participants. The participants were not presented with any statement such as “computer generated audio”. This was done in order to avoid possible biased responses.

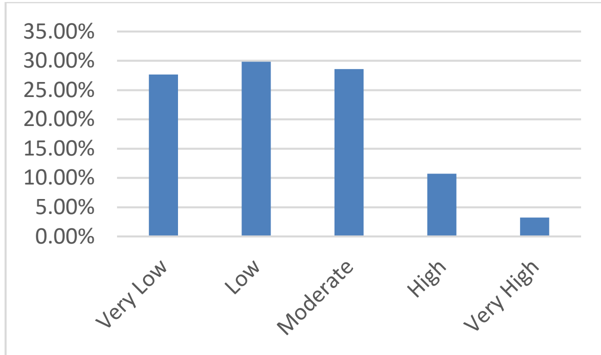


Figure 7-1. The average response value for different pieces according to their musicality criteria.

A few number of participants might knew about the fact that the audio materials were machine generations due to previous contacts in academia and poster presentation gatherings held at the University of Hull. The rates of the responses for the evaluation criteria are presented in the following. (A complete table of the data gathered from the first survey are provided in table G-1 in appendix G):

- Musicality: The total proportion of agreed responses on the musicality of the generated audio is around 15%. Moderately agreed responses are around 30%. Around 56% of the total responses were rated as low or very low. Figure 7-1 demonstrates the result of the assessment on the musicality of the audio samples.

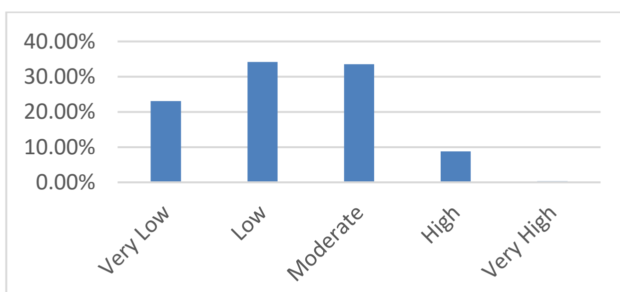


Figure 7-2. The average response value for different pieces according to their Persian-likeness criteria.

- Dastgāh-likeness: Approximately 43 percent of the responses were positively rated upon Dastgāh-likeness of the audio pieces. The rest of the respondents (around 57 percent) did not agree about the Dastgāh-likeness of the audio pieces in total (Figure 7-2).

- Respondents' musical preferences: On average, an approximate proportion of 30% of participants liked the audio pieces moderately or highly. Around 70 percent of the audience rated their preferences rather low or very low among the audio pieces. Figure 7-3 shows the preference ratings of the audio samples.

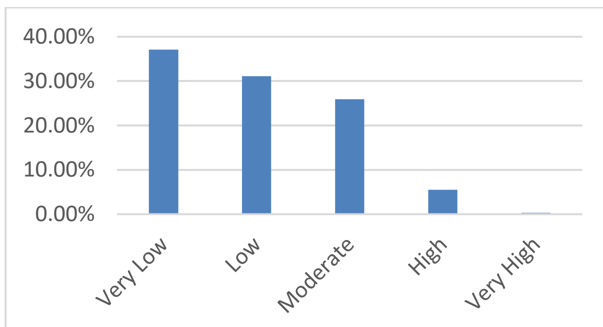


Figure 7-3. The average response value for different pieces according to the participants' preferences criteria.

Some notable feedbacks in the first survey include those that associated the produced audio to be as if a child was trying to learn a new musical instrument or was trying to compose a musical piece. The audio samples were attributed with randomness. Some comments suggested over defining better constraints over the duration of the musical notes since they appeared to be random. Most of the feedbacks attribute the sounds to be sourced from a musical instrument such as Piano, Santur, and guitar. Feedback of this kind shows the success of the system on mimicking the timbre of musical instrument. Comparing the audio to Santur (the Persian musical instrument) and the Piano as a hammered musical instrument (which follows the same excitation mechanism as Santur) are both warming comments. Some of the comments suggested that some audio pieces were dissonant and harsh to the ear: "*As if someone is randomly exciting the broken strings of a guitar*." Is one of the comments that states this situation.

7.3 The Second Public Survey

7.3.1 The Design and Evaluation Criteria of the Second Survey

In this section the design and evaluation for the second survey is discussed. This test was published throughout the mailing list of postgraduate students at the University of Hull. The introductory part of the survey is similar to the welcoming part of the first survey. A brief introduction of the research at hand and its purpose were presented. People were invited to familiarize themselves with Persian Dastgāh music through some links for those people who were not familiar with traditional Persian music. The audience were suggested to contact the creators of the survey for further questions, assistance and requests for the results.

The evaluation tests performed in this thesis were kept as compact as possible yet designed in such a way for the evaluations to be robust. The design of the surveys should not have needed a significant amount of time from the audience to complete. More exhaustive and time-consuming evaluations would have possibly required the assignment of rewards for the judges. However, this option was negated due to constraints on the budgets and fundings of the research project. Nevertheless, for avoiding possible fatigue in the critics, the tests were kept short and simple.

The first question in the survey evaluates people's bias towards computational creativity and generated Persian music. Questions about people's mind-set towards computational creativity make them aware of their subconscious biases before entering the main part of the survey. The questions associated to bias issues are presented in table 7-1 (some of which are inspired from (Jordanous, 2012b)). The participants were made available with a likert for each component. The likert choices are Strongly Disagree/ Disagree / Neutral /Agree / Strongly Agree.

Table 7-1 The first question components in the second survey is for identifying the particiants' bias towards computational creativity, and Dastgāh music.

a	Computers can produce creative outputs.
b	Computers can occasionally or randomly be creative.
c	Computers cannot be creative because they merely reflect the creativity of programmer.
d	The idea of computers being creative disturbs me.
e	Computers might be or can be creative in the future but currently are not creative.
f	Computers will never be creative.
g	Computers cannot generate Persian music.
h	Dastgāh Persian music should not be a subject for computational creativity.
i	I like the idea of computers being creative.
j	I do not like the idea of computers generating Dastgāh-like music.

In the auditory section of the second survey, a total number of seven audio samples are presented. Three of the samples were selected from the audio pieces generated from the compositional algorithm from chapter 5. Four of these audio samples were selected from the pieces generated from the models in chapter 6.The participants were invited to listen to each of the audio samples and answer the related queries. The audience were informed that they were going to listen to machine generated audio. Once the assumption was established that the audio pieces were machine generations, they could be judged against the predetermined criteria.

The evaluation criteria in the second survey are a combination of human opinion on creativity, Ritchie's criteria, and some other criteria that are related to the targets of this thesis. The combination of these methodologies were taken into account both in the hope of finding out more information that can help progress the work and both for establishing stronger grounds for reasoning about the achieved targets of this PhD thesis. The possibility of using Pease's criteria for evaluation, Colton's creativity tripod, and Jourdanous evaluation criteria were considered. The evaluation processes in these methodologies rely on the input, output, and the process itself. However, in this thesis the attention was kept towards evaluating the system by its outputs.

Applying Ritchie's criteria as a computational creativity evaluation method enables the assessment of the creativity of the system by its output. Ritchie's criteria evaluate the typicality, and value of the generated audio pieces. Ritchie's criteria were customized for evaluating the audio samples and are rephrased accordingly to be suitable for the current application. Therefore, the Ritchie's questions in this survey are slightly different from Ritchie's original questions and are taken from (Jordanous, 2012b) to fit the evaluation of a musical system:

- Is the audio an example of a musical improvisation? (Typicality)
- Is the audio a good musical improvisation? (Value)

Evaluation of the typicality and value of the generated audio pieces as musical improvisations are important. This is due to the improvised nature of performance of Dastgāh. The above Ritchie's criteria evaluate how the generated audio resemble Dastgāh music in this respect. Asking people about the creativity of the audio pieces and relying on their perspectives on creativity can be considered as a straightforward way of evaluation. The designed question for evaluating people's opinions on the creativity of the system is:

- Is the piece the result of a creative process?

The rest of the queries are domain specific questions and are designed to see how the system is successful in achieving the targets of the thesis.

- Is the audio music-like?
- Is the audio Dastgāh-like?
- Is a Persian musical instrument being played?

The participants were asked whether the audio are music-like. This question is different from the first query about typicality of audio pieces. If an audio piece is considered as a musical improvisation then it would be music-like.

A Persian musical instrument (Santur) model was employed as the synthesizer. We wanted to investigate how the changing of the parameters would keep the generated audio to be perceived as being played by Santur. Moreover, we wanted to take a look at how the acoustical aspects of the audio would effect the way they were perceived as Persian music.

The additional queries were:

- Did you like this audio piece?
- How much confident were you in answering these questions?

Liking an audio piece and rating it as creative or perceiving it as a good music, are different matters. It was assumed that separating these queries from each other might help the participants to differentiate these criteria from each other and rate them more independent of one another. The audience were asked about the level of their confidence in answering the queries (Jordanous, 2012b). The evaluation queries were repeated for all the seven audio samples presented in the survey from second to eighth question.

Table 7-2.The subqueries presented in the ninth question of the second public survey.

a	I spend as much time as I can listening to music.
b	I consider myself a Musician.
c	I consider myself as a Computer Scientist/ Computer Programmer.
d	I play at least one musical instrument.
e	I am familiar with Persian music.
f	I am familiar with Dastgāh Persian music.
g	I can identify the genres of music relatively easy.
h	I have/had formal training on music theory.

The ninth question in the survey (table 7-2) asks people about their familiarity with music, and computer science and their levels of proficiency (whether they are amateur or professional). People were asked if they are computer scientists or musicians. The queries also ask the participants whether they are familiar with Persian music. Responses to this question set should reflect the nature of the society of the participants. The last question in the survey asks the participants to leave their comments. Further details about the design of the second survey are presented in Appendix H. In the next sections, the survey evaluation results are presented.

7.3.2 Analysis according to Musical Backgrounds and Computer Science Knowledge

The second survey was circulated around the postgraduate students of the University of Hull. Overall, 53 people took part in the survey. The survey results show that most of the respondents spend as much time as they can listening to music. Likewise, most of these people can identify the genres of music relatively easy. 18 people out of 53 considered themselves moderate to very good computer scientists. 21 people present in the survey consider themselves as moderate to professional musicians. More than 50% of the survey takers play at least one musical instrument from modest to professional levels.

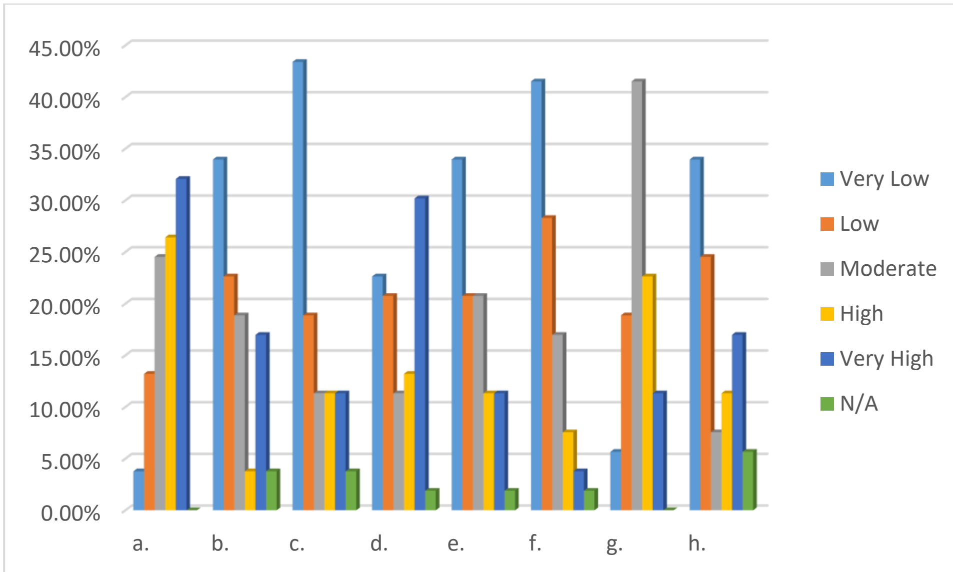


Figure 7-4 Collected responses about musical background, and computer science knowledge of respondents.

Each of the bars from left to right are associated with the queries a-h (in question 9) respectively (Please refer to table 7-2 for criteria a-h).

Around 59% of the respondents did not receive formal music theory to a moderate level. The familiarity of the respondents with Persian music and Dastgāh were limited. This is inferred from the fact that higher numbers of people selected the negative side of the likert (Figure 7-4).

7.3.3 Analysis of the Biases towards Computational Creativity and Dastgāh Music

In this subsection, an analysis is given to the set of queries designed for identifying the possible respondent's biases towards Persian music and computational creativity. Figure 7-5 illustrates the responses collected about queries regarding people's bias towards computational creativity, and Dastgāh music. The quantitative analysis of each of the under studied bias criteria (a-j) are presented in the following:

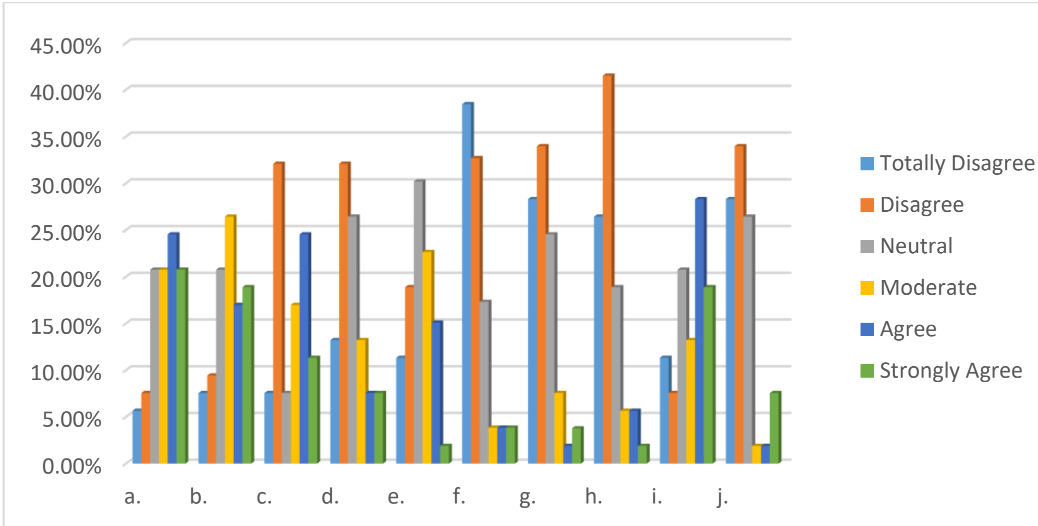


Figure 7-5 The responses collected about queries regarding people's bias towards computational creativity, and Dastgāh music.

Each of the bars from left to right are associated with the queries a-j (in question 1) respectively.

a. "Computers can produce creative outputs." 35 out of 53 people had moderately to strongly agreed that computers can produce creative material. 11 people stayed neutral, while the rest (7 people) had negative tendencies towards the matter.

b. "Computers can occasionally or randomly be creative." The second statement were rated with a similar trend to the previous statement in (a). 34 out of 53 people agreed with the statement (including all the moderately to strongly levels of agreement). 9 people disapproved the verdict. While the rest of the respondents (10 people) stayed neutral.

c. "Computers cannot be creative because they merely reflect the creativity of the programmer." The responses for the third statement are quite varying. However, a symmetrical pattern in the responses can be detected. The number of people who expressed strong approval (6) and strong disagreement (4) do not differ to a high extent. The rate of people who agreed (13) or disagreed (17) with the statement are near to each other and most of the ratings for this statements are allocated to the agree/disagree groups. The rest of people moderately agreed (9) or selected the neutral ground (4). The figure for this statement shows that 28 respondents in total agreed with the statement, and 21 people objected.

d. “The idea of computers being creative disturbs me.” Only 8 out of 53 people agreed or strongly agreed that computers being creative are disturbing to them. 7 respondents agreed moderately. 24 people disagreed with the statement. 14 people stayed impartial.

e. “Computers might be or can be creative in the future but currently are not creative.” This claim attracted various ratings while about 30 percent (16/53) of the respondents stayed indifferent. The percentage of people who had positive tendencies towards this statement were 40%. Thirty percent of the respondents disagreed with this statement.

f.“Computers will never be creative.” Around 71 percent (37/52) of the respondents disagreed that computers will never be creative. This figure is in contrast to the low number of people (8/52) who positively rated their agreement. The rest of the survey attendees expressed their neutral position towards the comment. One person did not leave their ratings for this comment.

g. “Computers can not generate Persian music.” Most of the survey takers (62%) opposed to the statement that computers cannot generate Persian music. About 25 percent of the respondents stayed neutral. While approximately 13 percent of the respondents (moderately to strongly) believe that computers cannot generate Persian music.

h. “Dastgāh Persian music should not be a subject for computational creativity.” This comment was opposed by 68 percent of the survey attendees. While only around 13 percent of people agreed with the statement, and around 19 percent of the respondents stayed impartial. The figure of the ratings for this statement is similar to that of the previous comment.

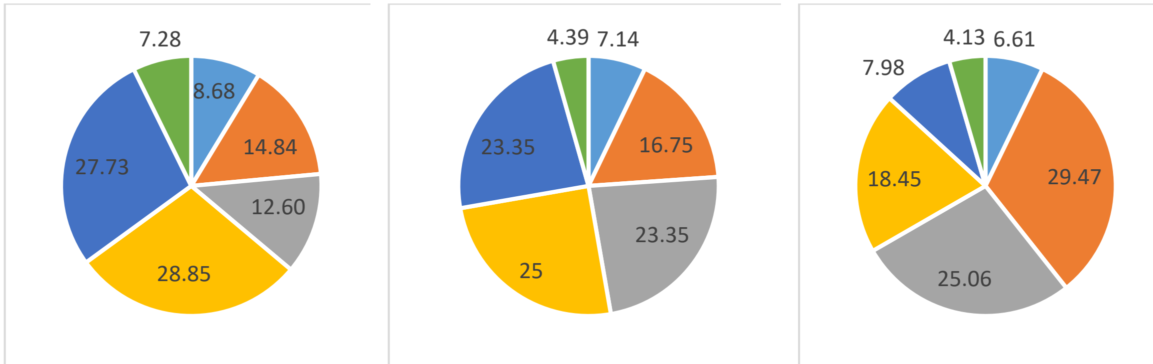
i. "I like the idea of computers being creative." Most of the people who took part in the survey (32/53) do like the idea of computers being creative. This liking is rated from moderate to strong level. Eleven people did not particularly had preferences about this statement. 10 people (about 18%) did not like the idea of computers being creative.

j. "I do not like the idea of computers generating Dastgāh-like music." There were only about 11 percent of the respondents who supported this comment. About 62 percent of the respondents disapproved the statement. Around a quarter of people who participated in the survey stayed neutral about this statement.

7.3.4 Quantitative Analysis of the Responses about the Machine Generated Audio

In this subsection the quantitative analysis of the responses about the machine generated audio are presented. The represented numbers and figures are calculated by taking the average responses over all respondents on the likert for all the seven audio pieces. Figures 7-6, 7-7, and 7-8 demonstrate pie charts related to the analysis of the responses to queries a-h about the machine generated audio in the second survey.

a. "Is the audio music-like?" 29 percent of the respondents on average moderately agreed that the audio pieces are music like. On average, the rate of people who agreed with the music-likeness of the audio pieces are slightly more than 27 percent. This figure shows that around 56 percent of the respondents go for the idea of music-likeness of the generated audio pieces. A mean of 12 percent of the respondents stayed impartial. While less than 15 percent and 9 percent of the participants disagreed and totally disagreed with the statement.



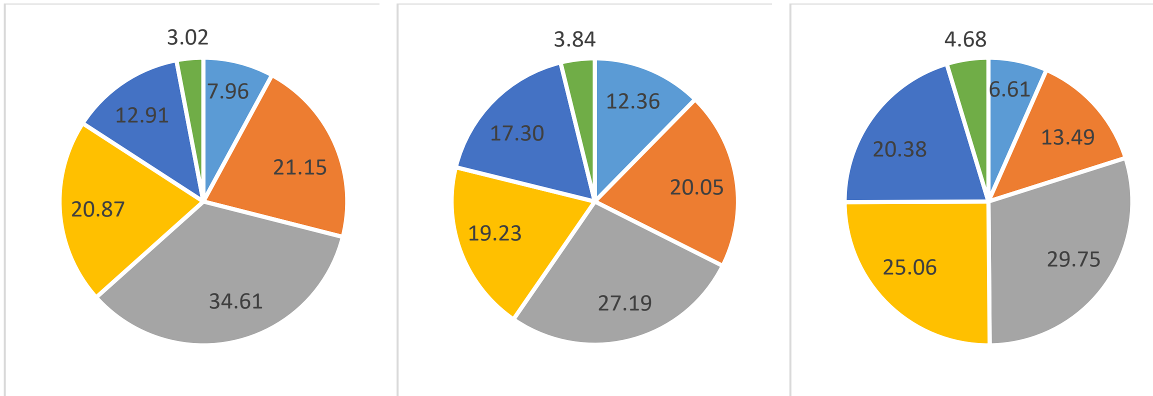
(a) Music-Like (b) Improvisation (c) Good improvisation

Figure 7-6 The pie-charts related to the analysis of the average responses to queries a,b, and c, about the machine generated audio in the second survey.

The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

b. “Is the audio an example of a musical improvisation?” On average, the highest rate of the likert is dedicated to moderate rating for this statement. This rating is around 25 percent in the moderate group which is followed by the ratings in the agree and neutral groups with 23.35% each. Slightly less than 17 percent of the participants disagreed with the statement while 7 percent of the survey attendees totally disagreed that the audio samples are examples of musical improvisations.

c. “Is the audio a good musical improvisation?” The mean proportion of the respondents who disagreed or totally objected this statement was around 26 percent in total. An average of 23 percent of the participants stayed impartial in rating the various audio pieces. A mean proportion of 18.45% of the respondents moderately agreed that the audio samples are good musical improvisations. The mean percentages of the participants who agreed or strongly agreed with the statement are 7.98%, and 4.13% respectively.



d) Dastgāh-like (e) Persian Instrument (f) Creative

Figure 7-7 The pie-charts related to the analysis of the responses to queries d,e, and f, about the machine generated audio in the second survey.

The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

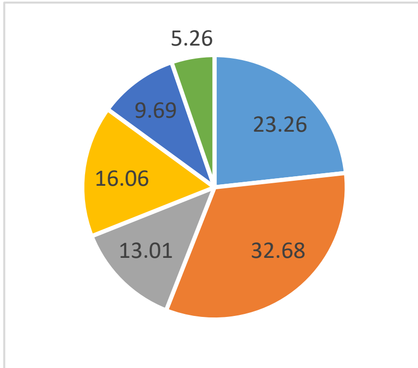
d. "Is the audio Dastgāh-like?" 34.61% of the participants stayed neutral in rating this comment while going through the audio samples. The proportion of the respondents who moderately agreed and those who disagreed with this statement are very similar and they are both approximating a mean of 21 percent. Around sixteen percent of the respondents stayed in the agree, and strongly agree groups together. About 8 percent of the respondents strongly disapproved of the audio samples to be Dastgāh-like. The quantitative analysis of this comment shows that an average of around 36 percent of the respondents approved the comment with different levels of positive agreements. This figure is around 30 percent for those who disapproved the Dastgāh-likeness of the audio samples.

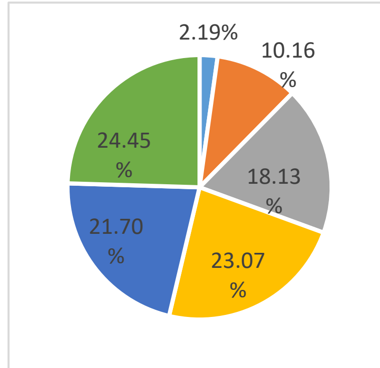
e. "Is a Persian musical instrument being played?" On average, the highest proportion is allocated to the group of neutral orientation as in the case of previous statement about Dastgāh-likeness of audio samples. 27.19% of the participants on average chose to be neutral in rating this statement. The approximate proportion of the respondents who agreed with this statement (from moderate to strongly agree level) over all the audio clips is 40 percent. The average percentage of the respondents who chose the negative side of the likert for this statement is slightly more than 32 percent in total.

f. “Is the piece a result of a creative process?” On average about 30 percent of the respondents stayed neutral in giving rates to this comment. The gravity of the responses stays in the positive side of the likert. A mean of 25.06% of the respondents moderately agreed that there had been a creative process behind the scenes of the production of audio samples. An approximate average proportion of 25 percent of the participants agreed or strongly agreed with the statement. The average proportions of the disagreements are around 20 percent in total.

g. “Did you like this audio piece?” The average gravity of the ratings for this statement is in the negative side of the likert for different audio pieces. These ratings are 32.68%, and 2.19% for the disagree and totally disagree ratings for the different presented audio clips in the survey. On average, the preferences of the respondents for the audio pieces decreases from 16.06% in the moderate group to 5.26% in the strongly agree group. There were about 13 percent of the respondents who stayed impartial about their preference of the audio samples presented.

h. “How much confident were you in answering these questions?” An average approximation of 70 percent of the ratings fall in the positive categories for this statement; meaning that most of the respondents rated to the queries with confidence. Only a mean of 12 percent of the participants reported negatively to this statement. A mean of 18.13% of the people taking the survey stated neutral position while rating this comment.





(g) Like (h) Confidence

Figure 7-8 The pie-charts related to the analysis of the responses to queries g, and h about the machine generated audio in the second survey.

The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

7.3.5 Summary Analysis of the Second Public Survey

Fifty-three people from the list of postgraduate students of the University of Hull participated in the survey. The analysis of their musical background shows that they generally like to listen to music, and they are capable to identify genres of music. Most of the participants have played at least one musical instruments in their life. Generally, they have limited familiarity with Dastgāh music. Most of them are not musicians or computer scientists.

The result of the analysis for identifying the possible biases of the respondents towards computational creativity, and Dastgāh music are briefly discussed here: Most of the respondents agreed that computers can produce creative materials. In addition, most of the survey participants agreed that computers can be creative or can occasionally or randomly be creative. However, the dominating thought is that computers merely reflect the creativity of the programmer. Almost half of the participants are not disturbed by the idea of computers being creative. The rest of the respondents whether belonged to the neutral group or to the group of people who are actually disturbed by the matter. The belief of the possibility of computers being creative in the future and not as a current fact, raised three major orientations among people: the three groups of orientations of neutral, and positive and negative tendencies are almost of the same sizes. The number of people who agreed with the statement were slightly more than those who disagreed. Most of the people who participated in the survey did not go for the idea that computers will never be creative. Likewise, more than half of the respondents did not agree that computers cannot generate Persian music. A quarter of the respondents did not oppose or agree with this statement (stayed neutral). Around 70 percent of the respondents approved that Dastgāh Persian music should be a subject for computational creativity. 62 percent of the participants approved that they like the idea of computers generating Dastgāh-like music. Generally, most of the people who took part in the survey do like the idea of computers being creative.

More than half of the respondents on average agreed with the idea of music-likeness of the audio samples. Slightly less than 50 percent of the participants, on average agreed that the audio clips are examples of musical improvisation. This is while around a quarter of them stayed neutral. The average proportion of the respondents who disagreed that the audio samples are good musical improvisation are 6 percent more than the rate of people who agreed about the matter. The mean rate of the people who agreed that the audio clips are Dastgāh-like are around six percent more than those who disapproved the Dastgāh-likeness of the productions. On average one-third of the respondents stayed impartial in rating the Dastgāh-likeness of the audio samples. The statement about the usage of a Persian musical instrument in the samples attracted a mean of 40 percent of various degrees of approvals from the audience. The disapprovement figure is about 9 percent less than the agreement ratio. A mean rate of around 27 percent of the respondents stayed neutral in rating this comment. The statement about whether a creative process is being involved in the audio generations gravitates towards positive responses rather than negative ones. The figure of about 30 percent neutrality in the rating of this statement persists (as in the ratings of the previous cases about Dastgāh-likeness and the statements regarding the typicality, and value of the audio pieces). More than 50 percent of the people who took part in the survey did not like the audio samples in general. A large proportion of the respondents stated that they had confidence in responding to the questionnaire.

7.3.6 Summary Results from Ritchie's Criteria for Evaluation

The set of questions for evaluating the Ritchie's criteria in the second survey are:

- Is the audio an example of a musical improvisation (Typicality)?
- Is the audio a good musical improvisation (Value)?

The full procedure for obtaining the Ritchie's evaluation criteria are presented in the appendix J. The respondents were provided with a likert to rate their answers. The ratings were weighted according to Ritchie's 0-1 standard range. The obtained values were applied for extrapolating the parameters for Ritchie's criteria formulations.

The mean typicality and mean value for each of the seven audio pieces in the second survey were calculated. These measurements were then compared with the *typicality*, *atypicality*, and *value* threshold parameters from Ritchie's formulations. In this stage, the audio pieces might be tagged as typical/atypical or valuable items. All the seven audio pieces were tagged as typical, and only two of them were considered to be valuable. The 18 Ritchie's criteria were then implemented by counting the members of different sets (for example members of typical and valuable groups) and populating the parameters of the set theory formulations. Table 7-3 demonstrates how the different Ritchie's criteria are satisfied or dissatisfied in the second survey.

Table 7-3 Results from applying Ritchie's criteria in the second survey.

	Evaluation Criteria	**RESULTS**
1	At least an average amount of the outputs should be competently typical.	TRUE
2	An acceptable proportion of the outputs should be competently typical.	TRUE
3	At least an average amount of the outputs should be adequately valuable.	TRUE
4	An acceptable proportion of the outputs should be adequately valuable.	FALSE
5	An acceptable proportion of the results should be both competently typical, and adequately valuable.	FALSE
6	An acceptable proportion of the outputs should be competently atypical and adequately valuable.	FALSE
7	An acceptable proportion of the atypical results should be worthy.	NOT APPLICABLE
8	An acceptable proportion of the qualified valued outputs should be competently atypical.	FALSE
9	The system should be capable of reproducing an acceptable proportion of the artefacts originally presented to the system as inspiring set.	NOT APPLICABLE
10	A decent proportion of the outputs of the system should be novel (should not be replications of the items in the inspiring set).	TRUE
11	On average, the outputs of the system which are novel should also be competently typical.	TRUE
12	On average, the outputs of the system which are novel should also be worthy.	TRUE
13	An acceptable proportion of the results of the system should be competently typical novel outputs.	TRUE
14	An acceptable proportion of the results of the system should be worthy novel outputs.	FALSE
15	An acceptable proportion of the novel outputs should be competently typical.	TRUE
16	An acceptable proportion of the novel outputs should be worthy.	TRUE
17	An acceptable proportion of the novel results should be both competently typical and adequately valuable.	FALSE
18	An acceptable proportion of the novel results should be both competently atypical and adequately valuable.	FALSE

The summary of applying Ritchie's criteria is demonstrated in table 7-4. In the appendix J it is demonstrated how some of the criteria are equivalent for this application

(This is due to the fact that none of the generated audio pieces were replicating members of the training set also known as inspiring set). Neglecting the equivalent criteria would leave 11 distinct criteria. Allowing the equivalent criteria, 50% of the 18 criteria are TRUE, 2 of them are not applicable and the remaining 7 are FALSE. This can be interpreted that the system fulfils half of the Ritchie's criteria to be considered as a creative system. Four of the distinct criteria out of 11 were true, while five were false and two of the criteria were not applicable. This suggests that the interpretation of the system as creative is compromised through using stricter criteria.

Table 7-4 The summary of the results of applying Ritchie's criteria in the second survey. Top three rows show all 18 criteria; last three rows show just the 11 distict criteria

9/18	Criteria TRUE	1,2,3,10,11,12,13,15,16
7/18	Criteria FALSE	4,5,6,8,14,17,18
2/18	Not Applicable	7,9
4/11	Distinct Criteria TRUE	1,2,3,10
5/11	Distinct Criteria FALSE	4,5,6,8,17
2/11	Not Applicable	7,9

7.4 The Survey for Professional Persian Musicians

7.4.1 The Design and Criteria for the Survey aimed for Professionals

The design of the third survey is similar to the second survey with some additional evaluation criteria. The question regarding participants' bias towards Persian music, and computational creativity are the same as in the second survey. There were 8 audio samples evaluated in the third survey and the audio samples were selected from the output results from systems described in chapters 5, and 6. Half of the samples belong to the composition set from algorithm in chapter 5, and the other half were associated to the outputs from the algorithm in chapter 6. The people who were invited to take part in the third survey belong to one or more of these categories:

- Have sufficient/great domain knowledge from Persian Dastgāh music.
- They are improvisers and music producers in the style of Dastgāh Persian music.
- They are professionals in playing at least one Persian musical instrument.
- They are singers who sing vocals and lyrics for Dastgāh music.

The additional evaluation criteria target the typicality and value of the audio pieces as Dastgāh music. These queries are:

- To what extent is the generated audio an example of the Dastgāh Persian music genre? (typicality)
- To what extent is the produced audio a high quality example of Dastgāh?(value)

While this survey gathered only 7 responses, it is presented as it offers a different audience set to the previous two surveys. Further details about the design of the third survey can be found in appendix I. In the following subsection, the results of the professional people evaluations on the outputs of the systems are presented.

7.4.2 Analysis of the Biases towards Computational Creativity and Dastgāh music

An analysis is provided according to the set of questions designed for identifying the possible respondent's biases towards Persian music and computational creativity. Figure 7-9 illustrates the responses collected about queries regarding people's bias towards computational creativity, and Dastgāh music. The quantitative analysis of each of the under studied bias criteria (a-j) are presented in the following:

a. “Computers can produce creative outputs.” Around 43 percent of the professional respondents were moderately fond of the idea that computers can produce creative materials. The rest of the respondents had disagreement about computers producing creative materials. The proportions of the respondents who went for each of the disagreement and totally disagreement choices were 28.57%. This contributes to the fact that more than half of the participants believe that computers cannot produce creative outputs.

b. “Computers can occasionally or randomly be creative.” 5 out of 7 respondents had negative tendencies about the statement that the computers can randomly or occasionally be creative. 2 people moderately supported this statement.

c. “Computers cannot be creative because they merely reflect the creativity of the programmer.” A total number of 4 people out of 7 agreed that computers cannot be creative because they merely reflect the creativity of the programmer, however 3 respondents disagreed with the statement.

d. “The idea of computers being creative disturbs me.” Around 81 percent of the survey participants were not disturbed by the idea of computers being creative. 14.29 percent of the respondents did strongly agree with the statement. The same proportion of the respondents (14.29%) stayed neutral. Most of the people taking the survey were not disturbed by the idea of computers being creative.

e. “Computers might be or can be creative in the future but currently are not creative.” 42.86% of the participants moderately agreed with the statement and a proportion of 28.57% of the respondents agreed that the computers can be creative in the future but are not currently creative. 28.57% of the survey takers totally disagreed with the statement.

f.“Computers will never be creative.” A proportion of 42.86% of the respondents disagreed with the statement and 14.29% of the people totally objected with the remark that computers will never be creative. This means that the gravity of the responses for this comment lies in the negative side of the likert. The percentages of the people who selected the neutral choice or those who agreed, or strongly agreed with the statement were all the same: 14.29%.

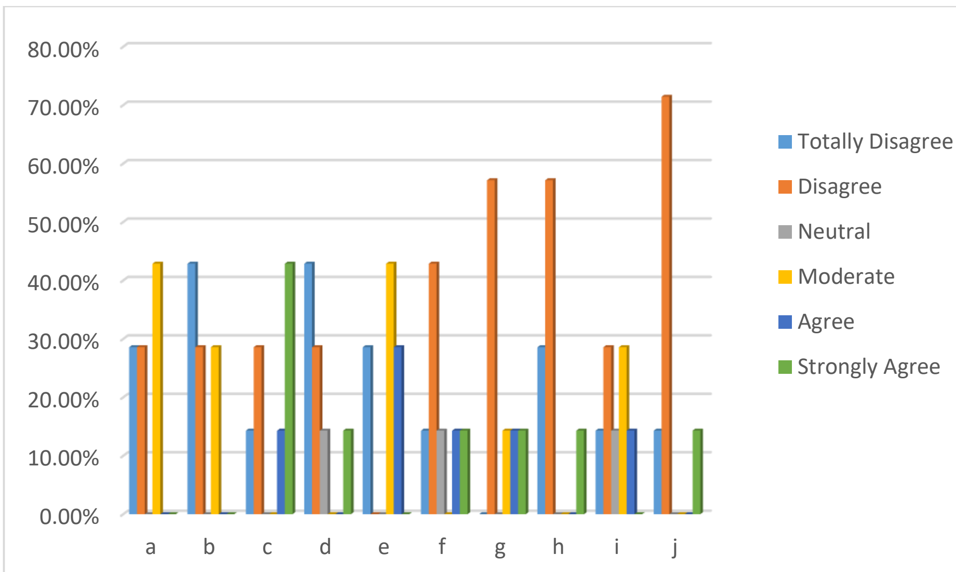


Figure 7-9 The responses collected about queries regarding people’s bias towards computational creativity, and Dastgāh music in the third survey.

Each of the bars from left to right are associated with the queries a-j (in question 1) respectively.

g. “Computers cannot generate Persian music.” A proportion of 57.14% of the people who took part in the survey disagreed with the comment that Computers cannot generate Persian music. The rest of the ratings for this comment is equally distributed between the choices of moderately agree, agree, and strongly agree. The proportions of the responses in each of these three positive groups are 14.29%.

h. “Dastgāh Persian music should not be a subject for computational creativity.” Most of the respondents (6 out of 7) opposed the comment that Dastgāh Persian music should not be a subject for computational creativity (It can implicitly infer that 6 out of 7 people thought Dastgāh music can be a subject for computational creativity). The figure in the negative side of the likert, suggests that there are 57.14% of the respondents who disagreed with the comment, and a proportion of 28.57% who totally disagreed. The proportion of people who went for strongly agree choice was 14.29%.

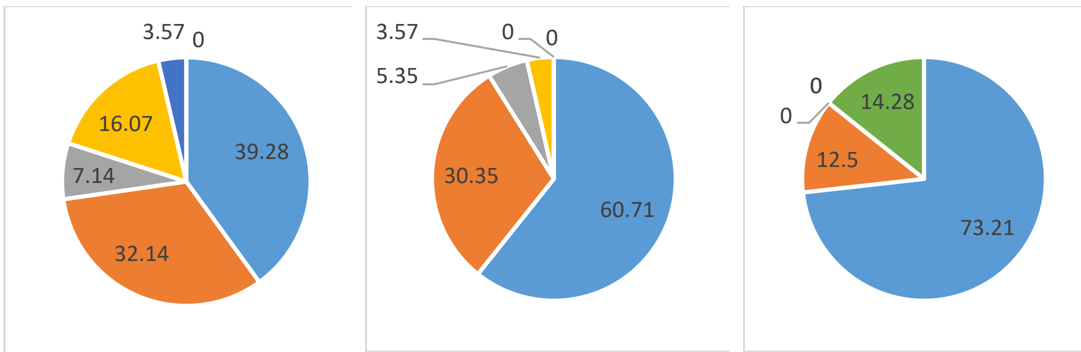
i. “I like the idea of computers being creative.” 2 of the respondents moderately supported the statement and only 1person agreed that they like the idea of computers being creative. On the negative side one of the respondents totally disagreed with the statement and 2 people disagreed. 1 out of 7 people stayed neutral about this remark.

j. “I do not like the idea of computers generating Dastgāh-like music.” A total number of 6 respondents out of 7 stated their opposition towards this statement. Only 1 person strongly agreed with the statement.

7.4.3 Analysis of the Responses about the Machine Generated Audio Clips

In this subsection the quantitative analysis of the responses about the machine generated audio are presented. The represented numbers and figures are calculated by taking the average responses over all the respondents on the likert for all the eight audio pieces. Figures 7-10, 7-11, and 7-12 demonstrate pie charts related to the analysis of the responses to queries a-i about the machine generated audio in the third survey.

a. “Is the audio music-like?” On average, there are only 16.07% of the respondents who moderately agreed and 3.57% who agreed with the statement. A considerable proportion of the participants counter voted for the musicality of the audio pieces. The average rate of people who chose the negative side of the likert for the different audio pieces is slightly more than 71 percent. A mean proportion of 7.14% of the participants stayed neutral while rating different audio pieces.



(a) Music-like (b) Improvisation (c) Good improvisation

Figure 7-10 The pie-charts related to the analysis of the responses to queries a,b, and c, about the machine generated audio in the third survey.

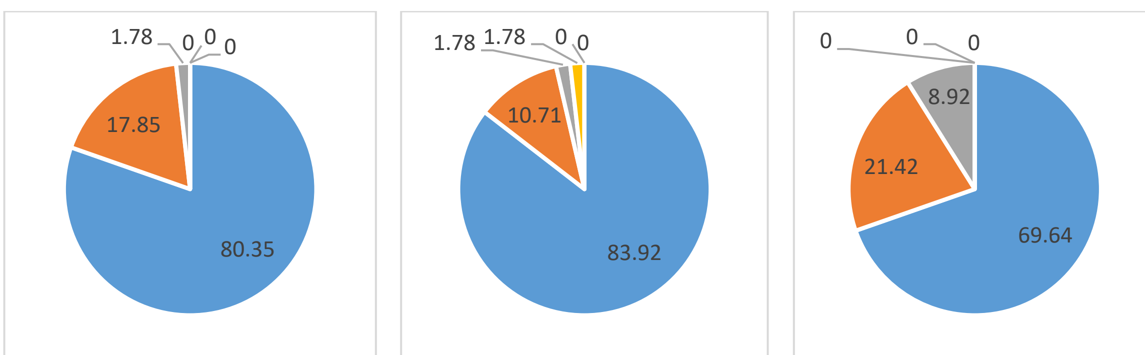
The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

b. “Would you agree that this is an example of musical improvisation?” The figures show that a large average proportion of around 91 percent of the respondents disagreed or strongly disagreed that the audio-clips represent samples of musical improvisation. Only 3.57% of the respondents on average moderately agreed that the audio pieces represent examples of musical improvisations. A mean proportion of 5.35% of the participants chose the neutral option while rating this statement over different audio pieces.

c. “Would you agree that this is a good musical improvisation?” A mean proportion of 73.21% of the respondents totally disagreed and an average rate of 12.5% disagreed that the audio samples are good musical improvisations. Only 14.28% of the respondents on average strongly agreed that the audio pieces are good musical improvisations.

d. “To what extent is the generated audio an example of the Dastgāh Persian music genre?” A large proportion of participants did not agree that the audio samples are in the genre of Dastgāh Persian music. An average proportion of 98 percent of the survey attendees disagreed and totally disagreed with this comment. A mean of 1.78% of the respondents stayed neutral while rating the audio pieces.

e. “To what extent is the produced audio a high quality example of Dastgāh?” A notable mean proportion of 83.92% of the respondents, totally disagreed, and 10.71% of the respondents disagreed with this statement. This implies that most of the concentrations of the ratings for this comment are in the negative side. An average proportion of 1.78 % of the respondents moderately agreed with this remark and a mean proportion of 1.78% of the respondents stayed neutral.



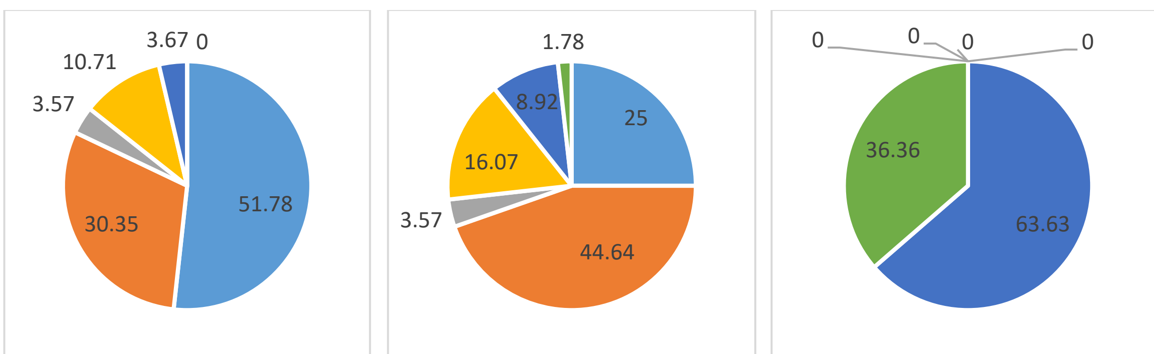
(d) Dastgāh-like (e) Quality Dastgāh (f) Liked

Figure 7-11 The pie-charts related to the analysis of the responses to queries d,e, and f, about the machine generated audio in the third survey.

The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

f. “Do you like this audio piece?” Most of the respondents did not like the audio samples. A mean proportion of 69.64% totally disliked the audio clips and 21.42% did not like them. A mean percentage of 8.92 of the respondents stayed neutral.

g. “To what extent would you agree the piece was a result of a creative process?” Only 10.71% of the respondents on average moderately agreed the productions as a result of a creative process. A mean rate of 3.67% of the survey participants had slightly better agreements with the matter. On average around 82 percent of the professionals did not agree the pieces to be the result of a creative process. A proportion of 3.57% of the respondents on average chose the neutral grounds in rating this statement.



(g) Creative process result (h) Persian instrument (i) Confidence

Figure 7-12 The pie-charts related to the analysis of the responses to queries g,h, and i, about the machine generated audio in the third survey.

The colours in the pie-charts are associated to the likert ranges: Totally Disagree, Disagree, Neutral, Moderate, Strongly agree, Agree

h. “Does the audio sound as if a Persian musical instrument is being played?” Most of the responses belong to the negative side of the likert. A mean proportion of around 70 percent of the respondents disagreed that a Persian musical instrument was being played throughout the audio clips. An average rate of 16.07% of the respondents moderately agreed with the statement. The mean proportion of respondents who agreed and strongly agreed with the matter is around 11 percent. A mean percentage of 3.57 of the participants stayed impartial.

i. “How much confident were you in answering these questions?” On average, the participants had high levels of confidence in rating the audio samples. All the concentration of the ratings lies in the high end of the likert.

7.4.4 Summary of the Result of the Professional Survey

More than half of the professional participants disagreed that computers can produce creative outputs. Most of them believe that computers cannot occasionally or randomly be creative. The number of people whom agreed that computers merely reflect the creativity of the programmer is almost the same as the number of respondents whom disagreed with the statement. A large proportion of the people taking the survey stated that they were not disturbed by the idea of computers being creative. Most of the respondents (approximately seventy-one percent) agreed with the comment that computers might be or can be creative in the future but are not currently creative. Around 30 percent of the respondents totally disagreed with this remark. Most of the respondents disagreed with the claim that the computers will never be creative, meaning that they believe at some point or in some ways that computers will be creative in the future. The survey takers disagreed with the negative statement that the computers cannot generate Persian music. All the survey participants except one of them agreed that Dastgāh Persian music should be a subject for computational creativity. Equal numbers of respondent like and dislike the idea of computers being creative. Almost none of the respondents had oppositions about computers generating Dastgāh-like music.

The results of the analysis of the ratings of the audio samples do not seem to be very satisfying in the survey designed for professionals. On average, there are slightly less than a fifth of the respondents who agreed on the musicality of the generated audio pieces. Moreover, a large proportion of the respondents on average, did not agree on audio clips to be valuable examples of musical improvisations nor even typical examples of musical improvisations. Almost all of the participants disagreed that the audio samples are examples of Dastgāh music genre. Likewise, the audio clips were not considered as high-quality examples of Dastgāh. Most of the participants did not like the audio pieces. Only a mean of 13 percent of the respondents agreed that the audio clips are the results of creative process. A large mean proportion of around 70 percent of the respondents disagreed that a Persian musical instrument is being played throughout the audio clips. The respondents stated that they were confident while rating the audio samples according to the evaluation criteria.

7.4.5 The Results of Evaluation from Ritchie's Criteria

In the third survey analysis, the Ritchie's evaluation criteria were implemented for evaluating the systems based on the audio samples being typical and valuable musical improvisations as well as being typical and valuable Dastgāh musical pieces. The first set of typicality and value questions in the third survey are:

- Is the audio an example of a musical improvisation? (Typicality)
- Is the audio a good musical improvisation? (Value)

The second set of typicality, and value questions in the third survey are:

- To what extent is the generated audio an example of the Dastgāh Persian music genre? (Typicality)
- To what extent is the produced audio a high quality example of Dastgāh? (Value)

The full procedure for obtaining the Ritchie's criteria evaluation for the first set and second set of questions are presented in the appendix K. In this paragraph the summary of the procedure for evaluating Ritchie's criteria for the question sets are presented. These question sets were evaluated independently in appendix K-1, and K-2, yet the procedure is similar. In the main survey the respondents were provided with a likert to rate their answers. The ratings were weighted according to Ritchie's 0-1 standard range. The obtained values were applied for extrapolating the parameters for Ritchie's criteria formulations. The mean typicality and mean value for each of the eight audio pieces in the third survey were calculated. These measurements were then compared with the typicality, atypicality, and value threshold parameters from Ritchie's formulations. In this stage, the audio pieces were investigated and they were tagged as typical/atypical or valuable items where suitable.

The audio pieces were all tagged as atypical samples for both sets of questions. None of the audio pieces was tagged as valuable for the first set of questions (related to musical improvisation). Only one of the pieces was tagged as valuable for the second set of question (related to Dastgāh music). The 18 Ritchie's criteria were then implemented by counting the members of different sets and populating the parameters of the set theory formulations. The results from applying Ritchie's criteria for the first set of questions are demonstrated in table 7-5. The results indicate that only 1 out of 18 criteria is fulfilled. The results from applying Ritchie's criteria for the second set of questions are shown in the last column of table 7-5.

The results show that only 2 criteria out of 18 criteria are met.

Table 7-5 Evaluating the Ritchie's 18 criteria for the third survey.

The evaluation results are given for the two sets of evaluation criteria: The first set of results refers to the typicality, and value evaluations of the audio pieces as musical improvisations. The second set of results refers to the typicality, and value evaluations of the audio pieces as Dastgāh music.

	Evaluation Criteria	**First set**	**Second set**
1	At least an average amount of the outputs should be competently typical.	FALSE	FALSE
2	An acceptable proportion of the outputs should be competently typical.	FALSE	FALSE
3	At least an average amount of the outputs should be adequately valuable.	FALSE	FALSE
4	An acceptable proportion of the outputs should be adequately valuable.	FALSE	FALSE
5	An acceptable proportion of the results should be both competently typical, and adequately valuable.	FALSE	NOT APPLICABLE
6	An acceptable proportion of the outputs should be competently atypical and adequately valuable.	FALSE	FALSE
7	An acceptable proportion of the atypical results should be worthy.	FALSE	FALSE
8	An acceptable proportion of the qualified valued outputs should be competently atypical.	NOT APPLICABLE	TRUE
9	The system should be capable of reproducing an acceptable proportion of the artefacts originally presented to the system as inspiring set.	NOT APPLICABLE	NOT APPLICABLE
10	A decent proportion of the outputs of the system should be novel (should not be replications of the items in the inspiring set).	TRUE	TRUE
11	On average, the outputs of the system which are novel should also be competently typical.	FALSE	FALSE
12	On average, the outputs of the system which are novel should also be worthy.	FALSE	FALSE
13	An acceptable proportion of the results of the system should be competently typical novel outputs.	FALSE	FALSE
14	An acceptable proportion of the results of the system should be worthy novel outputs.	FALSE	FALSE
15	An acceptable proportion of the novel outputs should be competently typical.	FALSE	FALSE
16	An acceptable proportion of the novel outputs should be worthy.	FALSE	FALSE
17	An acceptable proportion of the novel results should be both competently typical and adequately valuable.	FALSE	FALSE
18	An acceptable proportion of the novel results should be both competently atypical and adequately valuable.	FALSE	FALSE

In the appendix K, it is demonstrated how some of the criteria are equivalent for this application. This is due to the fact that none of the generated audio pieces were replicating members of the training set (known as inspiring set). Neglecting the equivalent criteria would leave 11 distinct criteria.

Tables 7.6 and 7.7 show the summary results for first and second set of questions on Ritchie's Criteria, respectively. Allowing the equivalence of criteria, 1 out of the 18 criteria are TRUE for question set 1, 2 of them were not applicable and the remaining 15 were FALSE. The second question set shows similar results with 2 of the 18 criteria as TRUE, 2 not applicable and the remaining 14 FALSE. This can be interpreted as stating the responses of the professional Persian musicians suggest the system is NOT being creative.

Table 7-6 The summary of Ritchie's criteria for the first set results in the third survey.

The first set of results refers to the typicality, and value evaluations of the audio pieces as *musical improvisations*. Top three rows show all 18 criteria; last three rows show just the 11 distict criteria

1/18	Criteria TRUE	10
15/18	Criteria FALSE	1,2,3,4,5,6,7,11,12,13,14,15,16,17,18
2/18	Not Applicable	8,9
1/11	Distinct Criteria TRUE	10
8/11	Distinct Criteria FALSE	1,2,3,4,5,6,7,17
2/11	Not Applicable	8,9

Table 7-7 The summary of Ritchie's criteria for the second set results in the third survey

The second set of results refers to the typicality, and value evaluations of the audio pieces as *Dastgāh music.* Top three rows show all 18 criteria; last three rows show just the 11 distict criteria

2/18	Criteria TRUE	8,10
14/18	Criteria FALSE	1,2,3,4, 6,7,11,12,13,14,15,16,17,18
2/18	Not Applicable	5,9
2/11	Distinct Criteria TRUE	8,10
7/11	Distinct Criteria FALSE	1,2,3,4,6,7,17
2/11	Not Applicable	5,9

A similar set of results are shown for the distinct criteria. 1 of the 11 were TRUE for question set 1, 2 not applicable and the remaining 8 FALSE. The second question set shows similar results with 2 of the 11 TRUE, 2 not applicable and the remaining 7 FALSE. This suggests that the interpretation of the system as creative is further compromised through using stricter criteria.

7.5 Chapter Summary

After the completion of modelling and implementation of the suggested architectures for algorithmic composition in chapters 5, and 6, the next important step in the progression of this research was to provide means for evaluating the results. This would attain an insight to the quality of the performance of the designed system. The assessments of the results of the systems identify their potentials. The tests would achieve a sense of the nature of the system's output.

Three surveys were presented in this chapter. The first survey was published right after the release of the results of the system according to the proposed architecture for genetic LPM system in chapter five. This evaluation method was based on human assessment regarding the musicality and Persian-likeness of the audio, and people's preferences.

The second survey was accompanied by further questions and was targeted for evaluating LPM system according to the architectures proposed in chapters five and six. This survey was designed for evaluating the creativity of the system based on two standard evaluation criteria: human opinion towards creativity and Ritchie's criteria for evaluation.

The third survey was aimed for professional groups of people and benefits from their perspectives towards Dastgāh Persian music and computational creativity. The third survey was also according to the architectures proposed in chapters five and six. The survey embeds a few more questions regarding the level of knowledge of the participants towards Dastgāh music.

Nonetheless, we cannot escape the reality that people would critic the generated sequences based on their own tastes of music and perception of creativity. Despite this fact, the surveys of this kind possibly reveal some hidden aspects about human aesthetical perceptions beyond current knowledge for measuring such, especially for the case of Persian music. Further collected information through the survey comments are presented in the discussion chapter.

Chapter 8. Discussion

This chapter discusses the theoretical aspects of creativity of the composition systems proposed in this project. It determines how the system fulfils the aesthetical goals and to what extent. It discusses whether the system was successful in achieving the goals mentioned in the hypothesis section from the first chapter.

There are some questions left here which we would want to find an answer for:

- How successful was the system in creating Dastgāh-like music?
- Can the productions of this project be considered as H-creative materials or do they contribute to P-creativity?
- Are the systems presented in the thesis creative?
- How Boden's three types of creativity have been targeted by the thesis work, and to what extent they have been achieved by the thesis work?

In order to answer to these questions, there would be a more scrutinized look over the comments provided by the participants in the survey to grasp a better perspective towards the extent of success of this project.

8.1 Evaluation of the Systems through the Survey Comments

In this section, some of the comments left by the survey participants are highlighted. They are used to achieve further insight towards the potentials and weaknesses of the generated audio.

- In the beginning of the second and third survey, it was emphasized that the audio were computer generated. There were comments, which suggested that some of the audio pieces were indistinguishable from human audio generations. The lack of human performer expressions made some of the audio to appear as computer generated; for others lack of consonance had made the audio pieces to seem random. There is an audio piece in which the parameterization of the model of the synthesizer produced piano like sounds; indeed, in a number of generated pieces the audio notes have the timbres of Piano or guitar. The space of familiar musical instruments might have had impact on the way an audio piece was perceived as human production. Here are some of the comments of this kind:

‘Much of the material very much felt computer generated. What I mean by that is it lacked the microscopic blemishes you'd expect from any human performer. With the one exception being the track on prepared piano, I was unable to determine whether this was computer generated or actually performed.’

‘The audios do sound music-like but in a very computerised way. If I heard the audios with no context then I would still come to the conclusion that they were computer generated.’

‘I'm not sure if some of these are computer generated and some done on real instruments.’

- The synthesizer produces the sound of a Persian musical instrument (Santur). What is normally expected is that the timbre of Santur should be perceived through all the audio samples, however, this has not been the case. The obtained parameters from the pattern-matching rules (the 20 original LBM pattern-matching rules and those introduced in chapter 6) had sometimes caused the pitch frequencies to fall in the ranges which resulted in the sounds to be normally excluded as Santur tones as perceived by audience. For instance since one of the audiences subjects commented the sound to be like the sound of Piano or guitar. This effect is not considered as a weakness. The author wished the algorithms to explore tones and frequencies beyond the limitations of the physical instrument itself. Therefore, the produced audio does not contribute to realistic pitch ranges. There are also audio samples which are similar to bursts of energy or noisy sounds. In fact, the parameters make the synthesizer unstable causing to those auditory effects. Some of the mentionable comments in this range are:

'After listening to the first Youtube clip provided, which I really enjoyed listening to, moving to these clips seemed coming from a computer generated sound (an electronic version of the instrument). Being a musician I was very aware of the difference in sound, which is why I was a little hesitant at the question on whether a Persian instrument was being played in the questionnaire clips. At times this was better than others though.'

'In some of the samples, I have also had a feeling of hearing fast and varied Jazz rhythms.'

'Not very experienced with Persian music or Dastgāh music but most of the samples sounded plausibly realistic.'

'It is clearly produced using some form of digitised production technique. If we are studying the compositional process of intelligence it may be worthwhile to transfer these across to the real instruments. This would increase our ability to determine an opinion based solely on the composition. The production of the audio is therefore too unreflective of the acoustic sound of such zither instrumentation that it automatically adds connotations to our

listening perceptions. The music also is not reflective of the realistic pitched range of the 'instrument' used. This could be a point of development for random computer generated music to be specified and limited by human intuition and/or musical expertise.'

- Some people left comments on the creativity of the system. The ideas in this range of responses were diverse. Some people believe that creativity is a human possession and is not transferrable to computers. Some comments implied that more clarification were needed about what we meant by creativity and about the description process involved in the genration of audio samples. Some notable comments towards creativity are:

'I didn't feel I could comment on whether or not the musical process was creative with little background on what process was being used. If it was totally random code then I would have said no, but if the computer was programmed to make a choice based on a preference then I would have agreed it was creative.'

'In my opinion 'creativity' is subjective; therefore, what I may deem to be creative for a computer may not be what I would consider to be creative for a person.'

'It seems that there is a far way to go for computers being creative. I'm happy to hear from this type of research.'

'Creativity in the music is nascent from human brain. Creativity is not acquired. What is acquired is experience that becomes a tool for manifesting creativity and flourishes the creativity itself. Similarly, it can be concluded that the transfer of creativity to the machine is not possible. At least not at the moment.'

- The preferences of people towards the audio pieces was quite varied and subjective.

'I was very surprised with the rhythms because I find them surprisingly credible and nice. I immediately imagined Bossa Nova songs with such rhythms. There is something strange to my ears, though... there are very rich rhythms and the notes variations generally as well... I asked myself is there any way of extracting the rhythms and some notes and building on top of them with other genres?'

'In my opinion all of the audios are very creative and unique.'

'Some of them were very nice to listen to and although they weren't very familiar to me, but I think they were very creative and I even played them twice.'

'This was quite painful to listen to. There wasn't even one piece that I liked.'

- The timings of the audio pieces were determined by the patterns generated by the algorithm. No attempts took place for defing tempo for the audio pieces. Not setting tempo for the audio pieces caused them to appear random in the auditory tests. It can be inferred that the settings of the onset time in the mapping process needs a more intelligent design as a future work. Nonetheless it would have been more fair to have audio samples with set time signatures in the samples collection. Moreover, some part of the randomness results from the jumping effect caused by sequencing random musical motives in the initial population of the genetic algorithms.

'The timing of the zither parts seems random, almost like it is electrical signals literally transferred into zither sound..'

'All the audio had the same thing in common- it didn't have a set time-signature, which made it uncomfortable to listen to and very different from almost all other types of music.'

'Many of these pieces seem to be the result of a random process, rather than innovative improvisation. (I wish you would have explained more about the creative process and its difference with a random process especially for the case of improvisation of traditional music).'

'I focused too much on what I perceived as discontinuities in the music.'

- The feedback provided by the professionals mention the lack of the formation of modes in the generated audio. This is one of the areas that needs further investigations.

'Very interesting experience to listen to. I think the timbres were often quite convincing but it definitely lacked a sense of mode in both repetitive shaping (formulas) and chromatic shading (microtones).'

'For me the thing that imposed gaps between the generated pieces and Persian music was the lack of modes and melodies based on the Dastgāh and vocal models of Persian music.'

'The generated audio enlivened more of Persian Santur's tuning space. Of course in some of the cases the tonality quality represented the sound of Persian Santur. However, the lack of the formation of musical modes did not guide me through the space of improvisation of traditional music.'

'In the discussion of improvisation in Persian music, the mere production of a random melody cannot be taken into account as an improvisation. Improvisation has its own coordinates and is based on a very strong infrastructure,'

- The execution of the systems proposed in the thesis requires high demands of time and memory capacity. Therefore the sizes of the audio pieces that the systems were supposed to generate were often selected to be short. Furthermore, the audio samples which were used in the surveys were required to be short, in order to make the surveys brief. However, this has harmed the sense of narrative nature which is often found in longer pieces of audio.The following comment states this problem clearly:

'I have listened to the beautiful piece given in the link prior to the survey and I was captivated by the sense of musical narrative I felt whilst listening. About the samples, I have felt that it sounded like it, but because the samples where short, I didn't feel the sensorial story/narrative that I have felt with the Dastgāh player. Maybe I would need a full piece to get it.'

- The auditory results suggest that the pieces were distant from Dastgāh music. Introducing further constraints to the algorithm may cause the generated audio to stay in the space of Persian music. This is opposed to discovering spaces alien to traditional notions of Persian music. Producing historical creative artefacts and yet sustaining a specific musical space or style is a narrow pathway.

8.2 How successful were the Systems in Creating Dastgāh-like Music?

Some principal questions addressed in this section are:

- How musical are the audio pieces?
- Can the generated audio be considered as Persian Dastgāh pieces?

In order to find a response for these questions we first consider the evaluations provided in the (second and third) surveys.

There were fifty-three respondents in the second survey. More than half of the public respondents in the second survey agreed about music-likeness of the audio pieces on average. Similarly, a mean of about 50 percent of the respondents agreed that the audio pieces are examples of musical improvisation. On average, about a third of the public respondents agreed about audio pieces to be good musical improvisations. Slightly less than forty percent of the public respondents agreed that the audio pieces were Dastgāh-like. The responses in the second audio survey imply average success in the generation of Dastgāh-like music.

The responses in the third survey are somehow despairing. It should be noted that there were only seven participants in the third survey. The bias analysis showed the existence of some levels of bias towards computational creativity and Dastgāh music among the professional participants. The average proportion of the professional participants who agreed about music-likeness of the audio pieces was around 20 percent. An average rate of about 4 percent of the respondents agreed that the audio pieces were examples of musical improvisation. While the people who agreed on the quality musical improvisations were around fourteen percent, on average. None of the participants agreed that the audio pieces were examples of Dastgāh music, while a mean rate of 2 percent agreed that they were good examples of Dastgāh music. According to the third survey and perspective of professionals, the system was not successful in creating Dastgāh-like music.

8.3 Evaluation of the Creativity of the System

Some of the questions revolving around the creativity of the system are:

- Are the systems presented in the thesis creative?

- Can the productions of this project be considered as historical creative materials or do they contribute to psychological creativity?

Around half of the public participants agreed that the audio pieces were resulted by a creative process. In order to find more robust responses to these questions we consider the Ritchie's evaluation criteria. The analysis of the second survey shows that nine out of eighteen of the Ritchie's criteria were met. An average amount (and an acceptable proportion) of the audio samples are competently considered as examples of music improvisation. At least an average amount of the outputs are adequately valuable music improvisation. A decent proportion of the outputs of the system are novel (they are not replications of the items in the inspiring set).

Some of the comments in the second survey suggested that the audio pieces look like machine generations, or they assimilate the productions of a child starting to learn a new musical instrument. Some of the other comments suggested mixing the audio pieces with Persian music or other genres of music in the future. The comments suggest that the audio pieces are not generally historical type of creativity. However, they can be considered as psychological type of creativity. This is due to the fact that the audio pieces remind people of generated audio or of a child trying to learn a new musical instrument.

In the third survey the proportion of people who agreed that the audio productions are the results of a creative process were about 14 percent only. The study on the Ritchie's evaluation criteria about audio samples being examples of musical improvisation, gathered one out of eighteen TRUE results. The fulfilled criteria imply that the audio samples are novel and have not been previously existed in the training set. The Ritchie's criteria about audio samples being typical examples of Dastgāh music fulfilled two out of eighteen criteria. One of the criteria was that the items were novel, the other criteria shows that an acceptable proportion of the qualified valued outputs are competently atypical. This means that the professional people believe that there are valuable audio samples, yet they cannot be considered as Dastgāh pieces. This may imply that the valued outputs belong to a different genre of music (unknown) other than Dastgāh.

Producing creative audio pieces while maintaining the style of Persian music has been the attention of the designed systems. However, the auditory survey from professionals suggests that this navigation were distant from Persian music. On this basis, criticizing the audio pieces to be samples of psychological or historical types of creativity becomes difficult or non-applicable. Never the less, this direction of research which is established in this thesis is in its infancy. More progression in this direction requires better scrutiny in the area of computational creativity itself, and better grasping of Dastgāh music.

8.4 Discussion on Boden’s Three Types of Creativity

One fundamental question arising out the reported research is: How Boden’s three types of creativity have been targeted by the thesis work, and to what extent they have been achieved by the thesis work?

Boden defined three types of creativity: combinational, exploratory, and transformational types of creativity. The systems designed and described in this thesis try to navigate the spectrum of various types of creativity by the application of computational intelligence tools. In chapter 5, an evolutionary system was designed which was responsible for experimenting with exploratory and possibly transformational types of creativity. Chapter

6 provided a computational structure to experiment with combinational type of creativity.

In these experiments, a population of random sequences of LPM were generated. The individuals in the population were considered as potential artefacts that can exist in the space of all possible LPM sequences. The system in chapter 5 explored this space of creativity by the help of genetic algorithm operators. By performing crossover, and mutation operators it is hoped that the system visits possible new artefacts. There is also a chance that the exploration ends up in transformational creativity; meaning that one or more of the governing dimensions of a sample are transformed with admissible results. The designed fitness function was a support vector machine regression model trained with Zipfian features extracted from Dastgāh music, and LPM sequences. The fitness function guides the search by giving higher scores to individuals with Zipfian metrics more similar to Persian music Zipfian metrics.

The Multimodal DBM employed in chapter six was responsible for experimenting with combinational creativity by discovering new links between patterns that were not directly associated with each other. Patterns were extracted from Persian music pieces and cellular automata progressions and were stored in consequent layers in the multimodal architecture. The resonation occurring between similar patterns in the Zipfian channel activates associated patterns in the other channels. The idea behind this architecture is inspired from some of the possible underlying mechanisms related to combinational creativity. The constituent elements or materials in an artefact have features (or patterns). The similarity or associations between those features are assumed to be the hidden links between the constituent elements in the artefact.

8.5 Chapter Summary

In the discussion chapter, we attempted to find answer to some questions regarding the success of the project and the level of creativity of the designed systems. The results from the surveys in the previous chapter were taken into account for finding answers for the questions presented in the beginning of this chapter. The comments left by survey participants were helpful in identifying the strengths and potentials of the systems. According to the analysis of the public survey, on average the system was successful in creating Dastgāh-like music. The survey for professionals shows the contrary: that the system was not successful in creating Dastgāh-like music. Instead, the analysis of the third survey suggested that an acceptable proportion of the audio pieces were not in Dastgāh category, yet they were valuable pieces of audio. It can be inferred that the audio pieces did not meet the expectations of the professionals about being in the Dastgāh genre. However, they can still be categorized in another genre (presumably unknown). Moreover, the second public survey shows the creativity scores of the system being in the midrange. The professionals' scores given to the creativity measures were low.

The comments left by the participants revealed so many facts. It can be indirectly inferred that the Zipfian metrics were not sufficient aesthetical measures for guiding the systems in creating Dastgāh-like music. The fitness function designed based on the Zipfian metrics do not seem to be able to adequately reflect all the possible aesthetical aspects related to Persian music. There are possible other areas which needs to be identified for measuring the aesthetical aspects of Persian music. Moreover, better possible ways should be taken into account for generating the initial population in genetic algorithms.

Chapter 9. Conclusion

In this last chapter an overview on the goals, methodologies taken, and the contributions of the PhD project are provided. Other possible directions and future works are discussed. Firstly a summary of the research underpinning this thesis is given.

The first chapter introduced the research motivations and hypothesis. The second and third chapters were dedicated to background knowledge required for performing the research experiments. Various machine learning tools, computational intelligence tools, and signal processing techniques were described in the background chapters. Cellular automata as computational intelligence tools was the main core responsible for generating creative materials in the designed systems. Liquid Persian Music (LPM) software was introduced in chapter 4. LPM is an audio generator that works by employing pattern-matching rules for extracting features from CA progression and feeding the parameters of a synthesizer. Later in chapter 4, sequences of pattern-matching outputs were studied according to their Zipfian characteristics. The fourth chapter determined the research direction by designing an evolutionary framework for evolving LPM sequences according to their Zipfian characteristics.

In chapter 5, Zipfian metrics were embedded in the presented evolutionary algorithm as aesthetical critics. A support vector regression (SVR) model was employed as fitness function in the evolutionary algorithm framework for evolving LPM sequences. The SVR model was trained with the Zipfian metrics extracted from Dastgāh music and LPM sequences. The evolutionary algorithm was used to navigate the space of possible LPM sequences. The Zipfian metrics were used to guide the search. In chapter 6, Boltzmann machine families were the focus of attention, for their various different capabilities. Restricted Boltzmann machines and conditional restricted Boltzmann machines were used as pattern-matching tools for extracting features from CA progression. Two multimodal deep Boltzmann machines (Multimodal DBM) were trained with Persian music and CA progression. The Multimodal DBMs were mainly used as resonators for experimenting possibilities for assimilating some of the mechanisms related to combinational creativity.

Chapter 7 dealt with the design and execution of three auditory surveys. The surveys criteria were in alignment with the targets of the thesis about creating new types of Dastgāh music. Ritchie's evaluation criteria were taken into account. People's opinions about the creativity of the generated materials from the experiments in chapters 5, and 6 were measured and analysed. Chapter 8 is the discussion chapter and provides further analyses on the survey results. The extent of success of the project in creating Dastgāh-like music was discussed. It was shown that the system achieved mid-range success in creating Dastgāh-like music according to public survey participants. The levels of creativity of the systems were rated in average levels. The third survey analysis shows that the systems were not successful in creating Dastgāh-like music and the systems were not creative.

9.1 Research Questions Investigated

It is worthwhile mentioning about how the research questions posed in chapters 1 and 4 were fulfilled implicitly throughout this project. The research questions cited in sections 1.5 and 4.7 are presented in the following:

- How to use computational intelligence methods to produce creative (audio) artefacts?

Cellular automata as a computational intelligence tool was employed as the main core for providing raw creative materials for the algorithmic composition models in the thesis. LBM software concepts were established as a basis for this research. Pattern matching rules were responsible for extracting features from CA progressions. Moreover, Boltzmann machines extracted patterns from CA progressions. The stochastic and generative natures of Boltzmann machines have also provided us with creative materials.

- How to guide this process to produce Persian Dastgāh-like music?

I investigated how the Liquid Brain Music system can be applied to the task of generating Persian Dastgāh-like music. LBM software was developed to LPM with a new synthesizer based on a Persian musical instrument. Later the voices produced by this system were studied according to their aesthetical aspects. Genetic algorithms were used for evolving LPM voices sequences. Boltzmann machine families were also used in a separate experiment for generating Persian-like music.

- How to design an efficient search space traversal, which resolves the sequencing problem within the constraints of given hardware resources?

A search space was designed at the end of chapter 4, and some of the dimensions of this space were eliminated for addressing the hardware constraints issues and for simplifying the search scope. The remaining dimensions determine the melody of the compositions. No attempts have been taken for harmonizing the audio pieces. Meanwhile the ADSR envelopes were not allowed to vary throughout the experiments.

- What are the possible approaches for sequencing voices in an aesthetically pleasing manner?

An evolutionary algorithm agenda was used to address the voices sequencing problem. The design and implementation of the related model is the subject of chapter 5.

- What are the possible designs for the genotypes and phenotypes of a musical sequencer based on LPM?

The beginning of chapter 5 is dedicated to the design of the genotypes as the members of the genetic algorithms populations, It is specified in the same chapter that the phenotypes are the auditory versions of the genotypes encodings.

- How to assess our musical productions in terms of aesthetics?

Zipfian metrics were identified as suitable aesthetical criteria in our applications. One of the main reasons behind this choice was the measurability of the Zipfian metrics. This enables the automation of the fitness evaluation of the GA population members.

- How to define musical critiques in order to criticize the musicality of LPM sequences?

The musical critiques employed in this project were defined based on the Zipfian metrics of Persian music pieces. We wanted the LPM voice sequences to have Zipfian metrics similar to those of Persian music. A fitness function based on support vector regression model was trained to credit the LPM voice sequences according to their Zipfian metrics. The more competent individuals in the population have the chance to survive and pass their genes to the future generations.

- Is there a measurement for the creativity of the generated materials?

Chapter 7 was dedicated to the design of evaluation criteria for determining the creativity associated with the generated materials. Three research surveys were published accordingly, and the results are reported in chapters 7, and 8.

9.2 Achievements

The research hypothesis explored in this PhD thesis was that it is possible to create Dastgāh-like music using appropriate computational intelligence methods. We hoped that by applying the main concepts of LPM software and establishing advanced systems using computational intelligence tools we would become able to create and experience new dimensions of composition for generating Dastgāh-like music.

This thesis is the first research project that considers the creation of Dastgāh music by the help of computational intelligence tools. Research of this kind has existed for a longer time for the case of Western music. The results presented in Chapter 7 and further discussed in Chapter 8 demonstrate that there is a journey ahead to get results that satisfy professional Persian musicians. In this thesis, only one aesthetical aspect has been considered for directing the compositions to follow Persian music: Zipfian metrics. The knowledge about aesthetics of music is limited and, in addition, it is a very subjective matter. However, the results from the surveys suggest that Zipfian metrics may be able to distinguish between LPM sequences but provide poor guidelines for generating Dastgāh-like music.

In this thesis, less discussion was directed towards different Dastgāh and their melodic characteristics. Most of the efforts were concentrated on tools and techniques for extracting features from Persian music and employing those features for algorithmic composition. Zipfian aesthetical features extracted from different Persian musical pieces became the guidelines for the composition task. There are still various other possibilities that can be used as aesthetical measures to be applied in this research, such as explicit representations that model domain knowledge about Dastgāh. Other possible directions are discussed in the section related to future work.

The list of contributions and achievements in this thesis are given here:

- This project is the first research that has been performed for creating Persian Dastgāh-like music by the usage of computational tools. (Previous and ongoing works on Persian music are related to music information retrieval and classification tasks. The resources in (Abdoli, 2011; Beigzadeh & Koochesfahani, 2016; Heydarian, 2016; Lāyegh et al., 2013) have investigated information retrieval from Persian music databases and performed classification of different Dastgāh)

- Developing LBM software to LPM software. Updating the additive synthesizer that worked based on adding sinusoidal waves and transforming the LBM synthesizer to a musical instrument synthesizer (guitar, Sitar, and Santur model) have been explored. It should be emphasized that part of producing Dastgāh-like music was to use appropriate timbres compatible with Persian music. On this account, the LBM synthesizer was updated to produce the sounds of a Persian musical instrument (Santur). Some other developments also took place to achieve further means for extracting features from cellular automata progressions. The twenty pattern-matching rules suggested in the initial LBM software were expanded to the number of hidden units in the Boltzmann machine family architecture trained with CA progression. This number is practically infinite according to the arbitrary architectures and topologies selected for Boltzmann machines.

- Extracting Zipfian metrics from the collection of the Radif music by *Ostad Faramarz Paivar* for Santur. Performing this required the extraction of the information related to musical notes, duration time, and onset times from audio databases using signal processing tools and toolboxes. The extracted features were stored inside MIDI databases where they were further analysed for deriving their Zipfian attributes.

- Design of systems (based on Boltzmann machines and cellular automata) for generating creative materials without human domain knowledge contributions. The obtained patterns were applied as musical motives and for populating the parameters of the synthesizer. The stochastic nature of Boltzmann machines are also a favourable trait that contributed to the variety of generated patterns.

- Design an architecture based on evolutionary algorithms for navigating the space of possible LPM sequences. This system was designed in the hope of experimenting with exploratory and transformation types of creativity. The exploration in the space of creativity was constrained by the application of a fitness function that assigned higher credit to LPM sequences with Zipfian metrics more aligned with those Zipfian attributes extracted from Persian music.

- Design of a system based on multimodal deep Boltzmann machines. This system was designed in the hope of experimenting with combinational type of creativity. The associative nature of Boltzmann machine families were investigated for replicating some of the traits related to combinational creativity. Two Multimodal DBMs were trained with Persian music data, and CA progressions. The patterns with similar Zipfian metrics in Multimodal DBM channels were used to resonate and call on the similar patterns in the other channels. Discovering links between items that have not been previously linked is a naive form of combinational creativity.

- Three surveys were designed and published in order to evaluate the creativity associated with the systems. This evaluation was performed based on the outputs generated by the systems. The criteria for evaluation were determined based on the objectives of the thesis. Ritchie's criteria for evaluation were also considered.

9.3 Future Work

In this section, some of the criteria for further development and future work are discussed:

- Designing user interface for the systems presented in the thesis:

One of the potentials of the current LPM system is the addition of manual controllers. This would enable the users to have direct contributions towards the parameterization of the model and the process of mapping to the musical space. The former Liquid Brain Music model was based on the real time interaction with the users. The user could choose the CA rule and pattern-matching rules for various configurations of the system as parameterization sources for different synthesizer configurations on a real-time basis. However, There were some limitations associated to the type of interactivity in LBM system which the author tried to overcome: The output of LBM system were solely based on the user interaction and choice of CA rule numbers and pattern-matching rules parameterization. By adding the option of saving and loading, the user had the opportunity to resume their previous navigations in the musical space. This provided the opportunity to explore various possible paradigms for the output voices manually. The author of this thesis also increased the number of possible concurrent voices in the LBM system, where each channel accepted different parameterizations.

As the development of LPM progressed throughout the PhD, the tasks of the system's voice sequencing were taken over by machine learning and evolutionary algorithms. The directions of the research were more focused on automating the audio generation in a musical manner rather than being performed manually. Building a user interface in the format of a composition software (or a composition assisting software) are suggested as future work. (These are in harmony with the purpose of this research and similar researches in algorithmic composition.)

In the following, some limitations in this direction and some further suggestions are presented. The limitation of the architectures proposed in chapters 5 and 6 is that the support vector machine and restricted Boltzmann machines models are not generally designed as machines with adaptive learning capabilities. These machines are often employed after performing the training phase. Once these systems are trained, they are applied as components in the general architecture. In this respect, the nature of the designed systems can be considered to be autonomous with minor interactions with users. The intensive training process required is particularly time consuming; this constrains a real-time manual manipulation of the parameters by the users. In spite of the mentioned limitations, there are still some measures that can be manipulated by the users and can be taken into account as parameters for designing the user interface of LPM system. The suggested measures and parameters are discussed separately for the systems described in chapters five and six.

In chapter 5, a SVR was trained to differentiate between LPM random sequences and Persian music pieces. The SVR model was later applied as a fitness function for an evolutionary algorithm. The aim of the navigation was towards the enhancement of the LPM voice sequences in a musical manner. Although the computations of the evolutionary algorithm are labour intensive and time consuming, there are still chances remaining for designing an interactive based system by increasing the computational power of the processor. As been described in chapter five there are various parameters and operators involved for initializing the evolutionary algorithm process and even altering those operators and parameters during the evolution; a list of suggested evolutionary algorithm operators are provided in section 5.1.2 in chapter five. Other evolutionary parameters include the mutation rate, and crossover rate, which can be systematically changed. Moreover, the initial population can be chosen arbitrarily from previous population generations to continue to evolve with different parameterizations.

Some of the controllable measures for building a user interface on top of the architecture suggested in chapter six are discussed in the following. It has been demonstrated that once the Multimodal DBM are trained they are capable of storing the higher representations of the data provided as training set. The user interface can be equipped with a graphical tool that allows the user to visualize the connecting weights to neurons and for selecting the neurons for getting the outputs to be mapped to the musical space.

- Design of a hybrid system based on the evolutionary architecture proposed in chapter five, and the related architectures based on Boltzmann machine families in chapter six: The patterns achieved from Boltzmann machine pattern-matching tools can be used as an important option for initializing the first population in the evolutionary algorithm. This choice would provide a hybrid system based on the architectures suggested in chapters five and six. As been demonstrated previously, the effects of the random initialization of the LPM sequences would remain to some extent even after the evolution process. The effect shows itself as significant jumping between the musical motives in the LPM sequences. However, if this initialization is performed by the outputs generated by system suggested in chapter six it is hoped that the jumping effect will be minimised. Furthermore, the generated output from Multimodal DBM architectures will enter an evolutionary system where they will have the chance for being evolved. In this manner, the design of further fitness functions in accordance with the new hybrid system might enhance the overall musical results.

- Expanding the design of evolutionary system to evolve all the eleven dimensions of the produced audio: In the LBM software, seven synthesizer parameters were introduced. All of these parameters contribute to the quality essence of the generated audio. For instance, the ADSR parameters govern the nature of the notes envelopes. A space of all possible configurations for LPM voices was portrayed in chapter four. An evolutionary framework for navigation through this space was designed in chapter five. The dimension of the evolutionary algorithm search space was limited to only 3 of the 11 original parameters. These three parameters determine the pitch frequency, the duration of the notes and the notes onset times. This implies that the focus was kept on evolving sequences of voices based on their melodic structure. The other governing aspects of the audio were kept constant. For instance, the ADSR envelopes parameters were left untouched. The reasons behind this decision are manifold which were stated in chapter five and are briefly mentioned here to be considered as a future work. The reason for the reduction of the dimensions of the voices relied on the designing complexity of the fitness function. The designed fitness functions were based on the features extracted from Persian music databases. The used database consisted of Dastgāh pieces performed with one musical instrument only. The timbre of the instrument did not change. Therefore, presently, no information exists in our databases to attain measures for effective timbral changes to use as a guideline for evolving voices on that basis. The only changing parameters throughout the musical pieces were the pitches and durations of the musical notes. Therefore varying the ADSR envelopes of LPM pieces would require guidelines that were outside the scope of the explored traditional Persian music database. The design of fitness function for evolving LPM sequences based on other ADSR envelopes and other signal processing parameters can be performed based on their psycho acoustical measurements. Designing fitness functions based on psycho acoustical aesthetics are subjects for further investigations.

- Obtaining further means for interpreting patterns from CA progressions and mapping them to musical domain. CA has been in the attention of people in the computer music community. This is because CA produces new genres of behaviour. There have been various approaches for mapping CA patterns to music domain. The interpretation of patterns of CA progressions plays an important role in the result. For instance in (Kirke & Miranda, 2007) some approaches for performing the mapping from CA domain to music space were discussed. In the earlier LBM software, 20 pattern-matching rules were originally introduced. These twenty pattern-matching rules were the basis of the experiments performed in chapters 4, and 5. Chapter 6 introduced new classes of pattern-matching tools by applying Boltzmann machines families. The proposed approach for extracting features from CA would practically produce vast number of pattern-matching tools depending on the number of RBM units and the employed topologies in RBM structures. There is a question left here which goes back to the interpretation of the patterns extracted from CA sequences. Interpreting the extracted patterns from CA using Boltzmann machines or in another words describing what the model has learnt is very difficult. Let alone be the mapping of those features to musical domain. Projecting the extracted patterns from CA progressions to musical space requires further investigations.

- Obtaining further aesthetical measures other than Zipf's law for guiding the designed systems: The truth is that there are limited knowledge about the universal measurements for associating aesthetical qualities to music. Moreover, people's musical preferences are a subjective matter. People state whether they like or dislike a piece of music and they can often describe what they like/dislike about a piece of music. However, the practice of capturing universals for musical aesthetics especially in a measurable manner is extremely challenging and the subject of controversy. In this project, the only guidelines applied for directing the compositional algorithm were the Zipfian metrics extracted from Persian music. Yet there are other possible potentials which can be applied as aesthetical measurements for guiding the algorithmic composition of Persian music. Improving the model can be performed by choosing further aesthetical measures based on Persian music that would cover other aspects of the produced audio.

- Design of a hybrid model based on the compositional systems suggested in this thesis with the addition of knowledge-based systems: research in the area of artificial intelligence and Persian music are recent. With the development of tools and methodologies, further achievements can be obtained in the area of algorithmic composition and Persian music. A hybrid system benefitting from both knowledge based systems and the compositional systems suggested in this thesis can possibly be one of the future directions for this project. A system with this specification could possibly be able to create in the space of the target musical style. This requires the intelligence and creativity to both conform in the target style while being able to break the norms, the rules, and the constraints of the specified style. The identification of balance between maintaining and breaking the style constraints are challenging tasks.

In fact, the designed system can benefit from a hybrid structure where a knowledge-based system can infuse the model with specialized domain knowledge on Persian music. Achieving a knowledge-based system for traditional Persian music would require the cooperation of Persian musician experts that would provide the knowledge they have gained through years of exposure to traditional Persian music. *Farhat's* book (Farhat, 1990) is one of the most complete books which discusses the Dastgāh concept in Persian music and offers some classifications of the modes and occurring patterns in each Dastgāh. To date, no research exists in how to represent the concept of Dastgāh in a form understandable to knowledge systems. However, preparing a comprehensive knowledge-based representation for Persian music can be considered as a necessary task for future researchers in this area.

References

Abdoli, S. (2011). Iranian Traditional Music Dastgah Classification. *International Society for Music Information Retrieval Conference (ISMIR)*, 275–280.

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A Learning Algorithm for Boltzmann Machines. *Cognitive science*,*Elsevier*, *9*(1), 147–169.

Anders, T. (2007). Composing Music by Composing Rules: Design and Usage of a Generic Music Constraint System. *Social Sciences*, *10*(February), 412. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.96.9497

Arshi, S. (2012). An Implementation of Santur musical instrument and the Synthesis of music pieces using Learning Machines, MSc. Thesis. Science and Research Branch of Islamic Azad University, Tehran,Iran.

Arshi, S., & Davis, D. N. (2015). Towards a Fitness Function for Musicality using LPM. *The 6th York Doctoral Symposium*.

Arshi, S., & Davis, D. N. (2016). A Computational Framework for Aesthetical Navigation in Musical Search Space. In *AISB symposium on computational creativity*. Sheffield, UK.

Arshi, S., & Davis, D. N. (2017). Capturing the Dynamics of Cellular Automata, for the Generation of Synthetic Persian Music, Using Conditional Restricted Boltzmann Machines. In Bramer & Petridis (Eds.), *Artificial Intelligence XXXIV. SGAI 2017. Lecture Notes in Computer SciencE* (Vol. 10630). Springer, Cham. http://doi.org/https://doi.org/10.1007/978-3-319-71078-5_6

Bank, B. (2006). *Physics-based Sound Synthesis of String Instruments Including Geometric Nonlinearities, Budapest University of Technology and Economics Department of Measurement and Information Systems, Ph.D. thesis*.

Beigzadeh, B., & Koochesfahani, M. B. (2016). Classification of Iranian traditional musical modes (DASTGÄH) with artificial neural network. *Journal of Theoretical and Applied Vibration and Acoustics*, *2*(2), 107–118.

Beyls, P. (1988). The Musical Universe of Cellular Automata. *Proc. Int. Comp. Music Conf. (ICMC88)*.

Bilbao, S. (2006). *Direct Numerical Simulation*. Retrieved from https://ccrma.stanford.edu/~bilbao/booktop/node17.html#735

Biles. (2014). GenJam. Retrieved October 1, 2014, from http://igm.rit.edu/~jabics/GenJam.html

Biles, J. a. (1994). GenJam: A genetic algorithm for generating jazz solos. *Proceedings of the International Computer Music Conference*, 131–137. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:GenJam:+A+genetic+algorithm+for+generating+jazz+solos#0\nhttp://igm.rit.edu/~jabics/GenJam94/Paper.html

Boden, M. (2009). Computer models of creativity. *Artificial Intelligence Magazine*, *30*(3), 23–34.

Boden, M. a. (2000). Computer models of creativity. *Psychologist*, *13*(2), 72–76. http://doi.org/10.1609/aimag.v30i3.2254

Boersma, P., & Weenink, D. (1995). Praat. Retrieved from www.praat.org

Boulanger. (2000). *The Csound Book: Perspectives in Software Synthesis, sound design, signal processing, and programming*. MIT Press.

Boulanger_Lewandowski. (2015). Modeling and generating sequences of polyphonic music with the RNN-RBM. In *DeepLearning Tutorial* (pp. 149–158). LISA lab, University of Montr_eal. Retrieved from http://deeplearning.net/tutorial/deeplearning.pdf.

Boulanger-Lewandowski, Bengio, Y., & Vincent, P. (2012). Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription. In *Proceedings of the 29th International Conference on Machine Learning* (pp. 1159–1166). Edinburgh.

Briot, & Pachet. (2017). Music Generation by Deep Learning - Challenges and Directions. Retrieved from http://arxiv.org/abs/1712.04371

Broersen, P. M. T. (2006). *Automatic autocorrelation and spectral analysis*. *Automatic Autocorrelation and Spectral Analysis*. Springer. http://doi.org/10.1007/1-84628-329-9

Buckles, B. P., & Petry, F. (1992). *Genetic Algorithms*.

Burges, C. J. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167. http://doi.org/10.1023/A:1009715923555

Burraston, & Edmonds. (2005). Cellular automata in generative electronic music and sonic art: a historical and technical review. *Digital Creativity*, *16*(3), 165–185. http://doi.org/10.1080/14626260500370882

Burraston, Edmonds, Livingstone, & Miranda. (2004). Cellular Automata in MIDI based Computer Music. *Proceedings of the International Computer Music Conference*, *4*, 71–78. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.3882&rep=rep1&type=pdf

Burton, a R., & Vladimirova, T. (1999). Generation of musical sequences with genetic techniques. *Computer Music Journal*, *23*(4), 59–73. http://doi.org/10.1162/014892699560001

Cardoso, Veale, & Wiggins. (2009). Converging on the Divergent: The History (and Future) of the International Joint Workshops in Computational Creativity. *AI Magazine*, *30*(3), 15–22. http://doi.org/10.1609/aimag.v30i3.2252

Carreira-Perpiñán, M., & Hinton, G. (2005). On contrastive divergence learning. *Artificial Intelligence and Statistics*.

Colton. (2008). Creativity versus the perception of creativity in computational systems. In *AAAI Symposium on Creative Systems* (pp. 14–20).

Cont, A., & Gluck, B. (2008). *Electronic Music in Iran*. *Electronic Music Foundation*.

Cook, P. R., & Scavone, G. P. (2008). The Synthesis ToolKit in C++ (STK). Retrieved from http://ccrma.stanford.edu/software/stk/

Cope, D. (1987). An expert system for computer assisted music composition. *Computer Music Journal*, *11*(4), 30–46.

Cope, D. (1996). *EMI: Experiments in musical intelligence*. A-R Editions.

Coughlin, & Baran. (1995). *Neural Computation in Hopfield Networks and Boltzmann Machines*. University of Delaware Press.

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines : and other kernel-based learning methods*. Cambridge University Press.

Dahlstedt, P. (2007). Autonomous evolution of complete piano pieces and performances. *Proceedings of the 9th European Conference on Artificial Life (ECAL)*, (1999).

Dārābi, N., Azimi, N., & Nojumi, H. (2006). Recognition of Dastgah and Maqam for Persian Music with Detecting Skeletal Melodic Models. *The Second Annual IEEE BENELUX/ …*. Retrieved from http://www.cs.sunysb.edu/~navid/papers/recognition.pdf

Darwin, C. (1906). The Origins of Species. In *Chapter 4, Natural selection*. London, England: OUP Oxford; Revised edition edition (13 Nov. 2008).

Davis, D. N. (2016). Computer and Artificial Music: Liquid Persian Music. Retrieved September 1, 2015, from http://www2.dcs.hull.ac.uk/NEAT/dnd/music/lpm.html

Diaz-Jerez, G. (2011). Composing with Melomics: Delving into the computational world for musical inspiration. *Leonardo Music Journal*, *21*, 13–14.

Dixon, S. (2007). Evaluation of the Audio Beat Tracking System BeatRoot. *Journal of New Music Research*, *36*(1), 39–50. http://doi.org/10.1080/09298210701653310

During, J. (2011). Dastgah, Encyclopedia Iranica. In *Iranica*. Retrieved from http://www.iranicaonline.org/articles/dastgah

Eerola, T., & Toiviainen, P. (2004). *MIDI Toolbox: MATLAB Tools for Music Research*. *MATLAB tools for music research. Jyväskylä, Finland: …*. University of Jyväskylä:

Kopijyvä, Jyväskylä, Finland. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:MIDI+Toolbox#6

Erkut, C., Välimäki, V., & Karjalainen, M. (2000). Extraction of physical and expressive parameters for model-based sound synthesis of the classical guitar. In *AES 108th CONVENTION* (pp. 19–22). Paris.

Essl, G. (2002). Physical Wave Propagation Modeling For Real Time Synthesis of Natural Sounds, PhD Thesis, Princeton University. *Trends in Food Science and Technology*.

Farhat, H. (1990). *The Dastgah Concept in Persian Music*. Cambridge University Press.

Fernández, J. D., & Vico, F. (2013). AI methods in algorithmic composition: A comprehensive survey. *Journal of Artificial Intelligence Research*, *48*, 513–582. http://doi.org/10.1613/jair.3908

Fletcher, N., & Rossing, T. (1991). *The Physics of Musical Instruments*. New York, USA: Springer-Verlag.

Freeman, & Skapura. (1991). *Neural Networks Algorithms, Applications, and Programming Techniques*. Addison Wesley; First Printing edition.

Frerking, M. E. (1994). *Digital Signal Processing in Communications Systems*. Chapman & Hall.

Gehler, P. V., Holub, A. D., & Welling, M. (2006). The rate adapting poisson model for information retrieval and object recognition. *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*, 337–344. http://doi.org/10.1145/1143844.1143887

Georgaki, A., Solomos, I. M., Zervos, G., & Proceedings, D. (2006). the Grain of Xenakis ' Technological Thought in the Computer Music Research of Our Days. *Electronics*, 1–7.

Gibson, & Byrne. (1991). Neurogen, musical composition using genetic algorithms and cooperating neural networks. In *In Proceedings of the 2nd International Conference in Artificial Neural Networks*.

Gluck, B. (2013). A New East-West Synthesis: Conversations with Iranian Composer Alireza Mashayekhi. In *eContact! 14.4 — TES 2011: Toronto Electroacoustic Symposium / Symposium électroacoustique de Toronto*. Montreal, CEC.

Goldberg, D. E. (1989). *Genetic Algorithm in search, optimization, and machine learning.*

Gustaffson, Kreiss, & Oliger. (1995). *Time Dependent Problems and Difference Methods*. John Wiley and Sons, New York, New York, USA.

Gutmann. (1999). John Cage and the Avant-Garde: The Sounds of Silence.

Hajimolāhoseini, H., Amirfattāhi, R., & Zekri, M. (2012). Real-time Classification of Persian Musical Dastgah s Using Artificial Neural Network. *The 16th CSI International Symposium on Artificial Intelligence and Signal Processing*, (Aisp), 157–160.

Hall, M. (1999). Correlation-based Feature Selection for Machine Learning. *Methodology*, *21i195-i20*(April), 1–5. http://doi.org/10.1.1.149.3848

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA Data Mining Software: An Update; SIGKDD Explorations, *11*(1).

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (Third edit).

Morgan Kaufmann.

Hein, G. E. (1991). Constructivist Learning Theory. *The Museum and the Needs of People - CECA (International Committee of Museum Educators) Conference*, (October), 15–22.

Heydarian. (2016). Automatic recognition of Persian musical modes in audio musical signals. *PhD Thesis, Music Technology Group, Faculty of Art, Architecture and Design, London Metropolitan University*, (June).

Heydarian, P., & Reiss, J. D. (2005). A DATABASE FOR PERSIAN MUSIC. In *DMRN summer conference*. Glasgow,UK.

Heydarian, P., & Reiss, J. D. (2005). The Persian Music and the Santur Instrument. *Proceedings of the 6th International Society for Music Information Retrieval Conference (ISMIR-2005)*, 524–527.

Heydarian, P., & Reiss, J. D. (2007). Extraction of Long-Term Rhythmic Structures Using the Empirical Mode Decomposition. *Audio Engineering Society Convention*, *122*.

Hinton. (2007). Boltzmann Machines. *Tutorial*, (1), 1–7.

Hinton. (2010). A Practical Guide to Training Restricted Boltzmann Machines A Practical Guide to Training Restricted Boltzmann Machines. *Computer*, *9*(3), 1. http://doi.org/10.1007/978-3-642-35289-8_32

Hinton, G. (2012). CSC321: Introduction to Neural Networks and machine Learning Lecture 16: Hopfield nets and simulated annealing.

Hinton, G., & Sejnowski, T. J. (1983). Optimal perceptual inference. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C.

Hinton, & Salakhutdinov. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, *313*(5784), 504–507. http://doi.org/10.1126/science.1127647

Hoffmann, P. (2002). Towards an Automated Art: Algorithmic process in Xenakis' Composition. *Contemporary Music Review*, *21*(2–3), 121–131.

Horner, A., & Goldberg, D. E. (1991). Genetic algorithms and computer-assisted music composition. In *In Proceedings of the International Conference on Genetic Algorithms*.

Iannis Xenakis webpage. (n.d.). Retrieved from http://www.iannis-xenakis.org/

Jacob, B. L. (1995). Composing with genetic algorithms. *In Proceedings of the 1995 International Computer Music Conference*, (September), 452–455.

Jaffe, D. a., & Smith, J. O. (1983). Extensions of the Karplus-Strong Plucked-String Algorithm. *Computer Music Journal*. http://doi.org/10.2307/3680063

Jordanous. (2010). A fitness function for creativity in jazz improvisation and beyond. *Proceedings of the International Conference on Computational Creativity, ICCC-10*, 223–227.

Jordanous. (2011). Evaluating Evaluation: Assessing Progress in Computational Creativity. *International Conference on Computational Creativity*, 102–107. Retrieved from http://sro.sussex.ac.uk/17942/

Jordanous. (2012a). A Standardised Procedure for Evaluating Creative Systems: Computational Creativity Evaluation Based on What it is to be Creative. *Cognitive Computation*, *4*(3), 246–279. http://doi.org/10.1007/s12559-012-9156-1

Jordanous. (2012b). *Evaluating computational creativity: a standardised procedure for* *evaluating creative systems and its application*. PhD Thesis, Department of Informatics, University of Sussex. Retrieved from http://sro.sussex.ac.uk/44741/

K.F. Man, Tang, K. S., & Kwong, S. (1999). *Genetic algorithms : concepts and designs*. London: Springer.

Karegowda, Manjunath, & Jayaram. (2010). Comparative Study of Attribute Selection Using Gain Ratio and Correlation Based Feature Selection. *International Journal of Information Technology and Knowledge Management*, *2*(2), 271–277.

Karjalainen, M., & Smith, J. (1996). Body modeling techniques for string instrument synthesis. In *ICMC Proceedings* (pp. 232–239).

Kauffman, S. a. (1969). Homeostasis and differentiation in random Generic Control Networks. *Nature*, *224*(5215), 177–178. http://doi.org/10.1038/224177a0

Khaleghi, R. (2013). *A Vision to Music (Persian)* (Eight). Jourchin.

Kira, K., & Rendell, L. (1992). The feature selection problem: Traditional methods and a new algorithm. *Aaai*, 129–134. http://doi.org/10.1016/S0031-3203(01)00046-2

Kirke, & Miranda. (2007). Eva[1] A. Kirke and E. R. Miranda, "Evaluating mappings for cellular automata music," Proc. ECAL Work. Music Artif. Life (MusicAL, Lisbon, Port., p. 12, 2007.luating mappings for cellular automata music. *Proceedings of ECAL Workshop on Music and Artificial Life (MusicAL, Lisbon, Portugal)*, 12. Retrieved from http://cmr.soc.plymouth.ac.uk/Musical2007/papers/Kirke.pdf

Klapuri, A., & Davy, M. (2010). *Signal Processing Methods for Music Transcription*.

Kononenko, I., & Kukar, M. (2007). *Machine Learning and Data Mining: introduction to*

Principles and Algorithms.

Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, *7*(1), 39–55. http://doi.org/10.1023/A:1008280620621

Kononenko, & Kukar. (2013). *Machine Learning and Data Mining*. Woodhead Publishing.

Laine, P., & Kuuskankare, M. (1994). Genetic algorithms in musical style oriented generation. *The First IEEE Conference on Evolutionary Computation*, 858–861.

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, *42*(1–3), 12–37. http://doi.org/doi: DOI: 10.1016/0167-2789(90)90064-V

Larochelle, H., & Bengio, Y. (2008). Classification using discriminative restricted Boltzmann machines. *Icml*, 536–543. http://doi.org/10.1145/1390156.1390224

Lartillot, O. (2013). *MIR toolbox I.5, User's Manual*. Finnish Centre of Excellence in Interdisciplinary Music Research.

Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction from Audio. In *International Conferene on Digital Audio Effects*. Bordeaux.

Lartillot, O., & Toiviainen, P. (2008). A Matlab Toolbox for Music Information Retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and KnowledgeOrganization*. Springer-Verlag.

Lattner, S., Grachten, M., & Widmer, G. (2016). Imposing higher-level Structure in

Polyphonic Music Generation using Convolutional Restricted Boltzmann Machines and Constraints. Retrieved from http://arxiv.org/abs/1612.04742

Lauly, S. (2007). Music generation using Dynamically Linked Boltzmann Machines, 1–8.

Lāyegh, M. A., Haghipour, S., & Sarem, Y. N. (2013). Classification of the Radif of Mirza Abdollah a Canonic Repertoire of Persian Music Using SVM Method, *1*(4), 57–66.

Lerch, A. (2012). *An Introduction to Audio Content Analysis*. John Wiley & Sons.

Li; Packard. (1990). The Structure of the Elementary Cellular Automata Rule Space. *Complex Systems*, *4*, 281–297.

Lo. (2012). Evolving Cellular Automata for Music Composition with Trainable Fitness Functions. *Electronic Engineering*, (March).

Lo, M., & Lucas, S. M. (2006). Evolving Musical Sequences with N-Gram Based Trainable Fitness Functions. *In 2006 IEEE International Conference on Evolutionary Computation*, (1), 601–608,. http://doi.org/10.1109/CEC.2006.1688365

Loeckx, A. J., & Bultheel, J. (2015). Forward Conditional Restricted Boltzmann Machines for the Generation of Music. In *ICCC 2015 : 17th International Conference on Computational Creativity*. Zurich, Switzerland.

Mahmoodan, & Banooshi. (2012). Automatic classification of Iranian music by an artificial neural network. In *2nd International Conference on Acoustics and Vibration*.

Manaris, Machado, P., Mccauley, C., Romero, J., & Krehbiel, D. (2005). Developing Fitness Functions for Pleasant Music : Zipf's Law and Interactive Evolution Systems. *EvoWorkshops*, 498–507.

Manaris, Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's Law, Music Classification, and Aesthetics. *Computer Music Journal*, *29*(1), 55–69.

Manaris, Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2011). Music Zipf ' s Classification ,. *Computer Music Journal*, *29*(1), 55–69.

Manaris, Roos, P., Machado, P., Krehbiel, D., Pellicoro, L., & Romero, J. (2007). A corpus-based hybrid approach to music analysis and composition. *Proceedings of the National Conference on Artificial Intelligence*, *22*, 7.

Mandel, Eck, D., & Bengio, Y. (2010). Learning tags that vary within a song. *In Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)*, (Ismir), 399–404. Retrieved from http://mr-pc.org/work/ismir10.pdf

Mandel, Pascanu, R., Larochelle, H., & Bengio, Y. (2011). Autotagging music with conditional restricted Boltzmann machines. Retrieved from http://arxiv.org/abs/1103.2832

Manning, P. (1985). *Electronic and Computer Music*. Clarendon Press, Oxford.

Marple, S. L. (1999). Computing the DiscreteTime Analytic Signal via FFT. *IEEE Transactions on Signal Processing*, *47*, 2600–2603.

Matlab. (2016). Matlab. *The MathWorks, Inc., Natick, Massachusetts, United States*.

McCormack, J. (2007). Facing the Future: Evolutionary Possibilities for Human-Machine Creativity. *The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music*, 417–451. http://doi.org/10.1007/978-3-540-72877-1_19

Miranda. (2001). Evolving cellular automata music: From sound synthesis to composition. *Proceedings of the Workshop on Artificial Life Models for Musical Applications-ECAL*, 12. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.4.1231&rep=rep1&type=pdf

Miranda. (2007). Cellular Automata Music: From Sound Synthesis to Musical Forms. In *Evolutionary Computer Music*. London: Springer-Verlag.

Miranda, E. R. (2002). Sounds of artificial life. In *Proceedings of the 4th Conference on Creativity & and Cognition, ACM* (pp. 173–177).

Mnih, Larochelle, & Hinton. (2011). Conditional Restricted Boltzmann Machines for Structured Output Prediction. *Uai*, 514–522. Retrieved from http://arxiv.org/abs/1202.3748\nhttp://uai.sis.pitt.edu/papers/11/p514-mnih.pdf\nhttp://arxiv.org/abs/1202.3748

Moffat, & Kelly. (2006). An investigation into people's bias against computational creativity in music composition. *The Third Joint Workshop on Computational Creativity*.

Muller, M. (2015). *Fundamentals of Music Processing, Audio, Analysis, Algorithms, Applications*. Springer.

Nettl, B. (1987). The Radif of Persian music: Studies of Structure and Cultural Context. In *Asian Music*. University of Texas Press.

Novaković, Strbac, & Bulatović. (2011). Toward optimal feature selection using ranking methods and classification algorithms. *Yugoslav Journal of Operations Research*, *21*(1), 119–135. http://doi.org/10.2298/YJOR1101119N

Ogihara, L. T. M., & Tzanetakis, G. (Eds.). (2012). *Music Data mining*. CRC Press, Taylor & Francis Group.

Oppenheim, A. V, & Schafer, R. W. (2013). *Discreet-Time Signal Processing* (Third). Pearson.

Paivar, F., & Sabet, S. (2004). *Elementary Course of Santur*. Mahour.

Pease, Winterstein, & Colton. (2001). Evaluating machine creativity. In *In Proceedings of Workshop Program of ICCBR-Creative Systems: Approaches to Creativity in AI and Cognitive Science* (pp. 129–137).

Peragine, M. (2013). *The Universal Mind: The Evolution of Machine Intelligence and Human Psychology*. Xiphias Press.

Pereira, Mendes, & Gervas. (2001). Experiments With Assessment of Creative Systems : An Application of Ritchie ' s Criteria. *Creativity*.

Powers, D. M. W. (2007). Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation, (December), 24. http://doi.org/10.1.1.214.9232

Powley, E. (2009). Global properties of cellular automata. *PhD Thesis*.

Proakis, J. G., & Manolakis, D. G. (2007). *Digital Signal Processing: Principles, Algorithms and Applications* (4th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Puckette, M. (2002). Max at seventeen. *Computer Music Journal*, *26*(4), 31–43.

Quinlan. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

Rahm, E., & Do, H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, *23*(4), 3–13. http://doi.org/10.1145/1317331.1317341

Rault, L. (2000). *Musical Instruments: A worldwide survey of traditional music making*.

Ritchie. (2001). Assessing creativity. In *Proceedings of the AISB symposium on AI and creativity in arts and science* (pp. 3–11). York.

Ritchie, G. (2007). Some Empirical Criteria for Attributing Creativity to a Computer Program. *Minds and Machines*, *17*(1), 67–99. http://doi.org/10.1007/s11023-007-9066-2

Ritchie, Manurung, Pain, Waller, Black, & O'Mara. (2007). A practical application of computational humour. In *4th International Joint Workshop on Computational Creativity* (pp. 91–98). London.

Roads, C. (1978). Automated granular synthesis of sound. *Computer Music Journal*, *2*(2), 61–62.

Roads, C. (1988). Introduction to Granular synthesis. *Computer Music Journal*, *12*(2), 11–13. http://doi.org/150.237.96.59

Rojas, R. (1996). The Hopfield Model. *Neural Networks*, 338–372. http://doi.org/10.1007/978-3-642-22253-5_3

Sadie, S., & Tyrrell, J. (2001). *The New Grove Dictionary of Music and Musicians*.

Salakhutdinov. (2010). Learning Deep Boltzmann Machines using Adaptive MCMC. *Icml*, *1*, 943–950. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.168.2418&rep=rep1&type=pdf

Salakhutdinov. (2012). Learning Deep Boltzmann Machines Software. Retrieved from http://www.cs.toronto.edu/~rsalakhu/code.html

Salakhutdinov, & Hinton. (2009a). Deep Boltzmann Machines. *Aistats*, *1*(3), 448–455. http://doi.org/10.1109/CVPRW.2009.5206577

Salakhutdinov, & Hinton. (2009b). Replicated softmax: An undirected topic model. *Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference*, 1607–1614. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-77956556686&partnerID=40&md5=4ad6020a17c32ccdab6547a8ea62d83a

Salakhutdinov, & Hinton. (2012). An Efficient Learning Procedure for Deep Boltzmann Machines. *Neural Computation*, *24*(8), 1967–2006. http://doi.org/10.1162/NECO_a_00311

Salakhutdinov, & Larochelle. (2010). Efficient Learning of Deep Boltzmann Machines. *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS 2010)*, *9*, 693–700. http://doi.org/10.1162/NECO_a_00311

Salakhutdinov, Tenenbaum, & Torralba. (2013). Learning with hierarchical-deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1958–71. http://doi.org/10.1109/TPAMI.2012.269

Smith III, J. O. (2004). *Physical Audio Signal Procesing, draft version, Stanford, CA.*

Retrieved from http://ccrma.stanford.edu/~jos/pasp04/

Smola, a J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199–222. http://doi.org/10.1023/B:STCO.0000035301.49549.88

Smolensky, P. (1986). Information Processing in Dynamical systems: Foundations of Harmony Theory. In D. E. Rumelhart & J. L. Mcclelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations.* (pp. 194–281). MIT Press.

Solomos, M. (2005). Cellular Automata in Xenakis ' Music . Theory and Practice. *Symposium A Quarterly Journal In Modern Foreign Literatures*, 145–152.

Srivastava, & Salakhutdinov. (2012). Multimodal Learning with Deep Boltzmann Machines. *Advances in Neural Information Processing Systems (NIPS)*, *15*, 2222–2230. http://doi.org/10.1109/CVPR.2013.49

Steinwart, I., & Christmann, A. (2008). *Support Vector Machines*. New York: Springer-Verlag.

Strikwerda. (1989). *Finite Difference Schemes and Partial Differential Equations*. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, California, USA.

Tanaka, M. (2016). Deep Neural Network, Mathworks, Matlab. Retrieved from https://uk.mathworks.com/matlabcentral/fileexchange/42853-deep-neural-network

Taylor, G. W. (2010). Matlab implementation of Implicit mixtures of Conditional Restricted Boltzmann Machines. Retrieved from https://github.com/gwtaylor/imCRBM

Taylor, G. W., & Hinton, G. E. (2009). Factored conditional restricted Boltzmann Machines for modeling motion style. *Proceedings of the 26th International Conference on Machine Learning (ICML 09)*, 1025–1032. http://doi.org/10.1145/1553374.1553505

Taylor, G. W., & School. (2012). Learning Representations of Sequences With Applications To Motion Capture and Video Analysis. *Cvpr*.

Theodoridis, S. (2009). *Pattern Recognition* (4th editio).

Tieleman, T. (2008). Training restricted Boltzmann machines using approximations to the likelihood gradient. *Proceedings of the 25th International Conference on Machine Learning - ICML '08*, 1064–1071. http://doi.org/10.1145/1390156.1390290

Tom Fawcett. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters, Elsevier*, 861–874. http://doi.org/10.1016/j.patrec.2005.10.010

Tretter, S. A. (2008). *Communication System Design Using DSP Algorithms*. Springer.

Turner, C. (2008). *Liquid Brain Music*. Computer Science, University of Hull.

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, *10*(5), 293–302. http://doi.org/10.1109/TSA.2002.800560

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag.

Vapnik, V. N. (1998). *Statistical Learning Theory*. New York: John Wiley & Sons.

Voss, R. F., & Clarke, J. (1978). "1/f noise" in music: Music from 1/f noise. *Journal of the Acoustical Society of America*. http://doi.org/10.1121/1.381721

Wang, C., & Gao, Z. (2012). Image denoising method based on v-support vector regression and noise detection. In *Proceedings of the 2012 2nd International Conference on Computer and Information Application*.

Weisstein, E. W. (2016). Cellular Automaton. Retrieved January 1, 2016, from http://mathworld.wolfram.com/CellularAutomaton.html

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining : practical machine learning tools and techniques*. (M. M. K. Amsterdam; Boston, Ed.) (3rd ed.).

Wolfram, S. (1986). Tables of cellular automaton properties. *Theory and Applications of Cellular Automata*, *57*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Tables+of+Cellular+Automaton+Properties#0

Wolfram, S. (2002). *A New Kind of Science*.

Woods. (2009). Liquid Brain Music: Phase II. *Available at: http://www2.dcs.hull.ac.uk/NEAT/dnd/*, (May), Computer Science, University of Hull.

Wuensche. (1999). Classifying cellular automata automatically. *Complexity*, *4*(3), 1–26. http://doi.org/10.1002/(SICI)1099-0526(199901/02)4:3<47::AID-CPLX9>3.0.CO;2-V

Wuensche. (2009). Discrete dynamics lab: Tools for investigating cellular automata and discrete dynamical networks. *Artificial Life Models in Software (Second Edition)*, 215–258. http://doi.org/10.1007/978-1-84882-285-6_8

Wuensche, A. (2004). Basins of attraction in network dynamics: A Conceptual Framework for Biomolecular Networks. *Modularity in Development and Evolution*, (figure 2), 288–

314. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.8.7599&rep=rep1&type=pdf

Wuensche, & Lesser. (1992). *The Global Dynamics of Cellular Automata*. *SIAM Review* (Vol. 35). http://doi.org/10.1137/1035161

Wulff, N. H., & Hertz, J. a. (1993). Learning Cellular Automation Dynamics with Neural Networks. *Advances in Neural Information Processing Systems 5*, (October 1993), 631–638. Retrieved from http://papers.nips.cc/paper/703-learning-cellular-automation-dynamics-with-neural-networks.pdf\nfiles/2942/Wulff ? Hertz - 1993 - Learning Cellular Automation Dynamics with Neural .pdf\nfiles/2943/703-learning-cellular-automation-dynamics-with-neural-netw

Yin, S., Wu, F., Luo, H., & Gao, H. (2015). Support vector regression based approach for key index forecasting with applications. *Proceeding - 2015 IEEE International Conference on Industrial Informatics, INDIN 2015*, 591–596. http://doi.org/10.1109/INDIN.2015.7281800

Zeiler, M. D., Taylor, G. W., Troje, N. F., & Hinton, G. E. (2009). Modeling pigeon behaviour using a Conditional Restricted Boltzmann Machine. *European Symposium on Artificial Neural Networks ESANN2009*, (April), 22–24.

Zhu, X., Xu, Z., & Knot, T. (2009). How creative is your writing? a linguistic creativity measure from computer science and cognitive psychology perspectives. In *In Proceedings of NAACL HLT Workshop on Computational Approaches to Linguistic Creativity (ACL)* (pp. 87–93). Boulder, Colorado.

Zicarelli. (2002). How I learned to love a program that does nothing. *Computer Music Journal*, *26*(4), 44–51.

Zipf. (1949). *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley.

Zonis, E. (1965). Contemporary Art Music in Persia. *The Musical Quarterly*, *51*(4), 636–648.

Appendix A. Training a Three-Channelled Multimodal Deep Boltzmann Machine

This appendix provides further information for training a Deep Boltzmann Machine (DBM), and a three-channelled Multimodal Deep Boltzmann Machines (Multimodal DBM) followed from chapter 3. The presented algorithm is applied in chapter 6.

A.1 Training Restricted Boltzmann Machine

The probability of a particular configuration over v visible input vector and h hidden vector is obtained by equation 15, which is also expanded to its fractional representation. Z is known as the partition function which is recognized as the normalization term. In theory, the Z normalising value is resulted by computing the value of $e^{-E(v,h)}$ for all the configurations for v, and h .

$$p(v,h) = \frac{e^{-E(v,h)}}{Z} = e^{-(h^TWv+ a^Tv+b^Th)}/_Z \tag{15}$$

Obtaining Z is intractable due to the large number of existing binary units in the network. Inferring the joint probability distribution of v, and h can be achieved by more practical approaches rather than computing the value of the partition function directly. In order to tract the value of the $p(v,h)$ a procedure is taken by first determining the conditional probabilities $p(v|h)$, and $p(h|v)$. This step would eventually assist in marginalizing out h from the joint probability $p(v,h)$ (Mnih et al., 2011).

The procedure for achieving $p(h|v)$ is given in sources (Hinton, 2007, 2010; Salakhutdinov & Hinton, 2012). Since RBM is an undirected network with symmetrical weights, no more complications are involved for gaining the conditional probability of observing visible layer given the hidden layer. The conditional probability of observing hidden units having visible units; ($p(h|v)$) is the product of conditional probabilities of individual hidden units. This is another confirmation that hidden units are conditionally independent given the value of the visible units. Independency of the units in the two groups of hidden and visible layers is caused by having no inter-layer connections. Obtaining $p(v|h)$ can be similarly approached.

$$p(h|v) = \prod_n p(h_n|v) \tag{16}$$

$$p(v|h) = \prod_m p(v_m|h) \tag{17}$$

According to formula 18 and 19, $p(h_n = 1|v)$ and $p(v_m = 1|h)$ can be achieved which are of the form of the sigmoid function $\frac{1}{1+e^{-x}}$. In the formulations below, W_n is the n^{th} row of the W, while W_m is the m^{th}column of W from its matrix notation.

$$p(h_n = 1|v) = \frac{1}{1 + e^{-(b_n + W_n v)}} \tag{18}$$

$$p(v_m = 1|h) = \frac{1}{1 + e^{-(a_m + h^T W_m)}} \tag{19}$$

The marginal probability distribution of a visible vector equals the sum of the exponential of the energy function over all possible configurations as in equation 20. The procedure for deriving $p(v)$ is available in (Hinton & Sejnowski, 1983).

$$p(v) = \sum_{h} e^{-E(v,h)} = \sum_{\mathrm{h}} e^{-(h^T W v + a^T v + b^T h)} \tag{20}$$

The target would be to maximize the logarithm likelihood of v (which is expressed as $\log\big(p(v)\big)$) for the training vectors presented in the visible units. Differentiating the logarithm likelihood with respect to $W, a,$ and b would provide a basis for employing gradient descent method for training RBM weights and biases. The weights parameter update (Salakhutdinov & Hinton, 2012) can be achieved from the following equation. Likewise, the values for bias changes can be achieved similarly:

$$\frac{\partial logP(v)}{\partial W} = (E_{P_{data}}(vh^T) - E_{P_{model}}(vh^T)) \tag{21}$$

Equation 21 comprises of two terms known as data dependent and data independent terms. $E_{P_{data}}(.)$ stands for the data dependent term and obtains the expectation over the data distribution. $E_{P_{model}}(.)$ is the data independent term which presents the expectation over the model distribution. The direct computation of these expectations are infeasible, however, there are approaches which enable the approximation of these values. Alternating Gibbs sampling between the visible and hidden units provide an opportunity for estimating the expectation values.

The training procedure for RBM can be pursued by utilisation of the contrastive divergence algorithm. The contrastive divergence algorithm (Hinton, 2010; Tieleman, 2008) works according to the following steps for each of the training sample vectors:

1. A training vector is presented to the visible units as input vector.
2. The probabilities of the hidden units are computed and their states are specified

accordingly and by the following stochastic rule:

$$s(h_m) = \begin{cases} 1, & p(h_{m=1}|v) > Random\ number\ from\ a\ uniform\ distribution \\ -1, & otherwise \end{cases} \quad (22)$$

3. Gibbs sampling is performed in k-levels: The vector $\tilde{v}$ is reconstructed by sampling from the hidden units and the hidden units $\tilde{h}$ are resampled from $\tilde{v}$. This alternate procedure takes place in k step times (A schematic of Gibbs sampling is illustrated in figure A-1).

4. The learning rules determine how the weights and biases should be updated.

$$W^t = W^{t-1} + \gamma(h^t {v^t}^T - \tilde{h}^t {\tilde{x}^t}^T)$$

$$a^t = a^{t-1} + \gamma(v^t - \tilde{v}^t) \quad (23)$$

$$b^t = b^{t-1} + \gamma(h^t - \tilde{h}^t)$$

5. The algorithm iterates from step 1, until a stopping criteria is met.

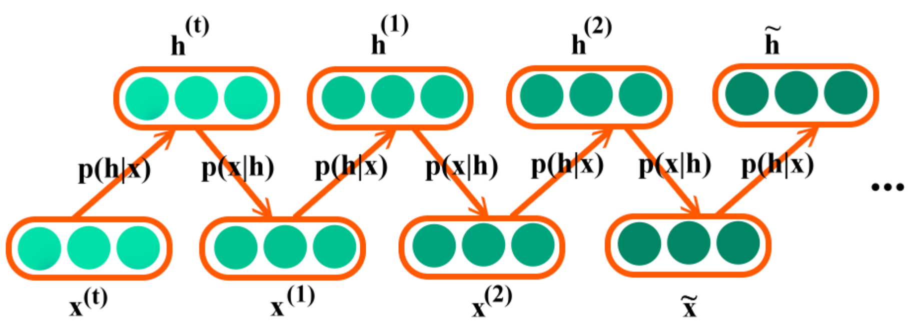


Figure A-1. A schematic view of the Gibbs sampling.

The updating of the weights and biases in the contrastive divergence algorithm consists of a positive phase and a negative phase. The positive phase can be assimilated to the memorization phase, however, the negative phase is a forgetting process; the contrastive divergence algorithm keeps the equilibrium between the two phases. Performing the storage of memories causes the weights to grow which may cause the system to crash eventually.

However, the right amount of forgetting is controlled automatically throughout the contrastive divergence algorithm (Hinton, 2012).

A.2 Procedure for Training a Deep Boltzmann Machine

The procedure for training a two-layered DBM is presented in this Appendix subsection. The training procedure followed in DBM is a basis for training a multi-channelled DBM that is the case for Appendix A.2. The formulations are adapted from (Salakhutdinov & Hinton, 2009a, 2012).

By having v as a set of units in the visible layer and $H = \{h^{(1)}, h^{(2)}\}$ as a set of units in the first and second hidden layers, the model parameters are expressed as: $\tau = \{W^{(1)}, W^{(2)}, b^{(0)}, b^{(1)}, b^{(2)}\}$. $W^{(1)}, W^{(2)}$ are the weights between visible to first hidden layer and the first hidden layer to the second hidden layers. The $b^{(0)}, b^{(1)}, b^{(2)}$ parameters are the biases of the visible, first hidden and the second hidden layers. The energy of the joint configuration can be obtained as:

$$E(v, h; \tau) = -\sum_{l=1}^{N_0} \sum_{m=1}^{N_1} W_{lm}^{(1)} v_l h_m^{(1)} - \sum_{m=1}^{N_1} \sum_{r=1}^{N_2} W_{mn}^{(2)} h_m h_n^{(2)} - \sum_{l=1}^{N_0} b_l^{(0)} v_l - \sum_{m=1}^{N_1} b_m^{(1)} h_m^{(1)} - \sum_{n=1}^{N_2} b_n^{(2)} h_n^{(2)} \tag{24}$$

The probability that the model assigns to visible units with respect to the model parameters τ is expressed as:

$$P(v; \tau) = \frac{\left(\sum_h \exp(-E(v, h^{(1)}, h^{(2)}; \tau))\right)}{Z(\tau)} \tag{25}$$

where Z is the partition function. The derivative of the logarithm likelihood with respect to $W^{(1)}, and\ W^{(2)}$ takes the following forms:

$$\frac{\partial logP(v;\tau)}{\partial W^{(1)}} = E_{P_{data}}\left[vh^{(1)^T}\right] - E_{P_{model}}\left[vh^{(1)^T}\right]$$
$$\frac{\partial logP(v;\tau)}{\partial W^{(2)}} = E_{P_{data}}\left[h^{(1)}h^{(2)^T}\right] - E_{P_{model}}\left[h^{(1)}h^{(2)^T}\right] \tag{26}$$

The data dependant expectation and model expectations are denoted as $E_{P_{data}}[.]$, and $E_{P_{model}}[.]$ respectively. The exact computation of the data dependent expectation and model's expectation are intractable. Some applicable approaches are mean field inference for achieving data-dependant expectations and Markov chain Monte Carlo method for approximating the data independent expectation.

Variational inference method works by replacing the true posterior probability distribution $P(h|v;\tau)$ by an approximation probability $Q(h|v;\mu)$. The approximate probability is written in its factorized form as:

$$Q(h|v;\mu) = \left(\prod_{i=1}^{N^{(1r)}} q\left(h_i^{(1r)}\middle|v\right) \prod_{j=1}^{N^{(2r)}} q\left(h_j^{(2r)}\middle|v\right)\right) * \prod_{k=1}^{N^{(3)}} q(h_k^{(3)}|v) \tag{27}$$

The above probability is obtained by approximating its constituent factorized marginal probabilities as $q\left(h_n^{(l)} = 1\middle|v\right) = \mu_n^{(l)}$ in which l represents the layer numbers and n is the neuron number. The μ are called mean field parameters that are obtained by considering the following maximization boundary problem:

$$\log\big(P(v;\tau)\big) \geq logP(v;\tau) - KL(Q(h|v;\mu)||P(h|v;\tau)) \tag{28}$$

In the above formulations, KL stands for Kullback-Liebler divergence method which is used for measuring the divergence between the true posterior probability and the approximate posterior probability distributions. The target is to minimize the divergence between the true and approximate probabilities. Therefore, in the variational inference approach the inference problem is converted to an optimization problem. The approximations are obtained based on the solution to the optimization problem. The mean field parameters for a two-layered deep Boltzmann machine can be obtained as:

$$\mu_{\mathrm{e}}^{(1)} = G\left(\sum_{\mathrm{d}=1}^{\mathrm{N}^{(0)}} W_{\mathrm{de}}^{(1)} \mathrm{v}_{\mathrm{d}} + \sum_{\mathrm{f}=1}^{\mathrm{N}^{(2)}} W_{\mathrm{ef}}^{(2)} \mu_{\mathrm{f}}^{(2)}\right), \mu_{\mathrm{f}}^{(2)} = G\left(\sum_{\mathrm{e}=1}^{\mathrm{N}^{(2)}} W_{\mathrm{ef}}^{(2)} \mu_{\mathrm{e}}^{(1)} + \sum_{\mathrm{m}=1}^{\mathrm{N}^{(3)}} W_{\mathrm{fm}}^{(3)} \mu_{\mathrm{m}}^{(3)}\right) \tag{29}$$

Variational inference tries to maximize the log-likelihood mean while minimizing the divergence between the true and approximate probabilities. The mean-field equations are obtained while the model parameters are kept fixed. The stochastic approximations are then applied to update the model parameters. The data dependent expectation can then be achieved:

$$\begin{aligned} E_{P_{data}}\left[v^r h^{(1r)^T}\right] &= \frac{1}{N}\sum_{n=1}^{N} v_n^m \mu_n^{(1m)^T} \\ E_{P_{data}}\left[h^{(1m)} h^{(2m)^T}\right] &= \frac{1}{N}\sum_{n=1}^{N} \mu_n^{(1m)} \mu_n^{(2m)^T} \end{aligned} \tag{30}$$

By once obtaining the mean field values, the parameters of the model are yielded by applying Markov chain Monte Carlo based stochastic approximation. The model parameters and states are updated in a sequence of Markov chains. The new states $\tilde{x}_{t+1} = \{\tilde{v}_{t+1}, \tilde{h}_{t+1}^{(1)}, \tilde{h}_{t+1}^{(2)}\}$ are sampled from the previous state x_t using Gibbs sampling. A gradient step is used to update the model parameters and the intractable model's expectation in the gradient equations in 26 is replaced by the point estimate at sample x_{t+1} - an average over the particles are the weight updates in the next step (t+1) for a DBM can be obtained from the following formulations in 31. γ_t is the learning parameter. $W_{t+1}^{(1)}$, and $W_{t+1}^{(2)}$ are the weights in the next time step.

$$
\begin{aligned}
W_{t+1}^{(1)} &= W_t^{(1)} + \gamma_t(\tfrac{1}{N}\textstyle\sum_{n=1}^{N} v_n \left(\mu_n^{(1)}\right)^T - \tfrac{1}{W}\textstyle\sum_{w=1}^{W} \tilde{v}_{t+1,w}\left(\tilde{h}_{t+1,w}^{(2)}\right)^T) \\
W_{t+1}^{(2)} &= W_t^{(2)} + \gamma_t(\tfrac{1}{N}\textstyle\sum_{n=1}^{N} \mu_n^{(1)} \left(\mu_n^{(2)}\right)^T - \tfrac{1}{W}\textstyle\sum_{w=1}^{W} \tilde{h}_{t+1,w}^{(1)}\left(\tilde{h}_{t+1,w}^{(2)}\right)^T)
\end{aligned}
\tag{31}
$$

The formulations in this section are expandable for building a structure with more hidden layers. The point to make here is that stacking restricted Boltzmann machines on top of each other does not produce a deep Boltzmann machine architecture; instead it produces a deep belief network structure (Salakhutdinov & Hinton, 2012). For the construction to form a deep Boltzmann machine, further considerations should be taken into account. In a deep belief network, the upper layers receive their input from previous bottom layer. Lower layers are not provided with feedback from higher layers. However, in a deep Boltzmann machine the bottom up and top down passes both exist in the united structure. Some considerations should be taken into account for the pre-training of a DBM.

The pre-training of a deep Boltzmann machine plays an important role in the convergence of the mean-field inference (Salakhutdinov & Hinton, 2012). Therefore sensible initialization of the weights are desirable rather that randomly initializing the weights. During the pretraining phase, all the weights are doubled to compensate for the lack of top-down pass. This is due to the fact that the stack of RBMs have not yet been combined as a DBM structure and the bottom layers provide inputs for the higher level layers. However, this procedure does not work quite right unless further modifications are performed on the magnitudes of the weights. The fact is that the first layer and last layer receive input from only one layer, while all the other intermediate layers accept inputs from their bottom and top adjacent layers. The states of all the hidden layers except the last hidden layer are inferred by geometrical means of their immediate bottom and top layer distributions. For a hidden layer $h^{(1)}$, the states of the units are influenced by both $W^{(1)}$, and $W^{(2)}$ weights equally. The states of the intermediate layers are resampled considering the sum of top and bottom layers input influences. The Gibbs sampling requires symmetrical weights in both directions for the first and last layers. However, this cannot happen since they receive input from only one layer. Overcoming this problem is accomplished by constraining the bottom up weight to be twice the top-down weights in the first layer. The top-down weights in the last layer is constrained to be twice the bottom-up weight in the last layer. The rest of the weights are doubled in the rest of the layers in both directions. This guarantees the symmetrical condition for the weights. The pretraining algorithm for DBM is in the following (Salakhutdinov & Hinton, 2012):

1. The first RBM is trained using contrastive divergence algorithm with mean-field reconstruction while the bottom-up weights are constrained to be $2W^{(1)}$ and top-down weights are $W^{(1)}$.

2. The second RBM uses the states $h^{(1)}$ inferred from $P(h^{(1)}|v, 2W^{(1)})$ as input vectors. This RBM is trained using contrastive divergence algorithm with mean-field reconstructions. The weights in both bottom-up and top-down directions are kept to be

$2W^{(2)}$.

3. The procedure performed in the second step is repeated for layers $L = 2, \dots, L-1$. The states of the hidden units in the $(L-1)^{th}$ layer are inferred from $P\left(h^{(L-1)} \middle| v, h^{(L-2)}, h^{(L-3)}, \dots, h^{(1)}, 2W^{(L-1)}, \dots, 2W^{(1)}\right)$. One step contrastive divergence takes place with both the bottom-up and top-down weights to be $2W^{(L-1)}$.
4. The last layer is trained by one-step contrastive divergence with top-down weights being constrained to be twice the bottom-up weights.

The weights obtained from the pre-training procedure are employed for initializing the network. The main training algorithm is discussed in the next section related to Multimodal DBM.

A.3 Algorithm for Training a Three Channelled Multimodal Deep Boltzmann Machine

Figure 3.12 in chapter 3 demonstrates the architecture of a Multimodal DBM. This structure consists of three pathways each of which has two hidden layers. The presented Multimodal DBM architecture is employed in chapter six. The accompanied formulations in appendix A.2 are tailored for the Multimodal DBM with three pathways. The original formulation can be found in (Srivastava & Salakhutdinov, 2012) where the Multimodal DBM were introduced. The pathways are tagged with English alphabets that assist in the presented formulations in this subsection. The set of units in the visible layer and the first and second hidden layers of the first pathway are denoted as: $path1 = \{v^{(r)}, h^{(1r)}, h^{(2r)}\}$. These sets for the other existing pathways in the presented structure are $path2 = \{v^{(s)}, h^{(1s)}, h^{(2s)}\}$,

$path3 = \{v^{(t)}, h^{(1t)}, h^{(2t)}\}$. The model parameters including the weights and biases related to the three pathways are expressed as: τ^r, τ^s, τ^t. The parameters of the conjoint layer is depicted as τ^m.

The probability that the deep Boltzmann machine assigns to the visible units in a pathway (r) can be achieved from equation 32. Since all the pathways have binary units for the visible layer, a similar formulation for $P(v^s; \tau^s)$, and $P(v^t; \tau^t)$ can be inferred.

$$
\begin{aligned}
P(v^r; \tau^r) = \sum_{h^{(1r)} h^{(2c)}} & P\left(v^r, h^{(1r)}, h^{(2r)}; \tau^r\right) \\
& = \frac{1}{Z(\tau^r)} \sum_{h^{(1r)} h^{(2r)}} \exp\Big(\sum_l \sum_m W_{lm}^{(1r)} h_m^{(1r)} v_l^r \\
& + \sum_m \sum_n W_{mn}^{(2r)} h_n^{(2r)} h_m^{(1r)} + \sum_l b_l^{(0r)} v_l^r + \sum_m b_m^{(1r)} h_m^{(1r)} \\
& + \sum_n b_n^{(2r)} h_n^{(1r)}\Big)
\end{aligned}
\tag{32}
$$

The probability that is assigned to the visible layer in the conjoint Multimodal DBM architecture can be achieved from equation 33. $L = \{v^r, v^s, v^t, h^{(1r)}, h^{(2r)}, h^{(1s)}, h^{(2s)}, h^{(1t)}, h^{(2t)}, h^{(3)}, h^{(4)}\}$ denotes all the visible and hidden variables. $\theta = \{\tau^r, \tau^s, \tau^t, \tau^m\}$ consists of all the model parameters of Multimodal DBM .

$$P(v^r, v^s, v^t; \theta) = \sum_{h^{(2r)}, h^{(2s)}, h^{(2t)}, h^{(3)}} P\left(h^{(2r)}, h^{(2s)}, h^{(2t)}, h^{(3)}\right) *$$

$$\left(\sum_{h^{(1r)}} P\left(v^r, h^{(1r)} \middle| h^{(2r)}\right)\right)\left(\sum_{h^{(1s)}} P\left(v^s, h^{(1s)} \middle| h^{(2s)}\right)\right)\left(\sum_{h^{(1t)}} P\left(v^t, h^{(1t)} \middle| h^{(2t)}\right)\right) =$$

$$\frac{1}{Z(\theta)} \sum_h \exp\Bigg($$

$$\overbrace{\sum_d \sum_e W_{de}^{(1r)} h_e^{(1r)} v_d^r + \sum_e \sum_f W_{ef}^{(2r)} h_f^{(2r)} h_e^{(1r)} + \sum_d b_d^{(0r)} v_d^r + \sum_e b_e^{(1r)} h_e^{(1r)} + \sum_f b_f^{(2r)} h_f^{(2r)}}^{First\ pathway} +$$

$$\overbrace{\sum_g \sum_h W_{lh}^{(1s)} h_h^{(1s)} v_g^s + \sum_h \sum_i W_{hn}^{(2s)} h_i^{(2s)} h_h^{(1s)} + \sum_g b_g^{(0s)} v_g^r + \sum_h b_h^{(1s)} h_h^{(1s)} + \sum_i b_i^{(2s)} h_i^{(2s)}}^{Second\ pathway} + \tag{33}$$

$$\overbrace{\sum_j \sum_k W_{jk}^{(1t)} h_k^{(1t)} v_j^r + \sum_k \sum_l W_{kl}^{(2t)} h_l^{(2t)} h_k^{(1t)} + \sum_j b_j^{(0t)} v_j^t + \sum_k b_k^{(1t)} h_k^{(1t)} + \sum_l b_l^{(2t)} h_l^{(2t)}}^{Third\ pathway} +$$

$$\overbrace{\sum_f \sum_m W_{fm}^{(3r)} h_m^{(3)} h_f^{(3r)} + \sum_i \sum_m W_{im}^{(3s)} h_m^{(3)} h_i^{(3s)} + \sum_l \sum_m W_{lm}^{(3t)} h_m^{(3)} h_l^{(3t)} + \sum_m b_m^{(3)} h_m^{(3)}\Bigg)}^{Joint\ layer}$$

The required formulations for training a Multimodal DBM can be induced similar to the deep Boltzmann machines. After all, a Multimodal DBM consists of several deep Boltzmann machine pathways. In a Multimodal DBM, the maximum likelihood learning approach is intractable. In fact maximum likelihood learning is an NP hard problem for the case of RBM families. Therefore, variation inference and Gibbs sampling are followed for training the model. The naive mean field approximation employs a distribution that is fully factorized. The approximate posterior probability in its factorial form for a Multimodal DBM with three pathways is written as in:

$$Q(h|v;\mu) = \left(\prod_{i=1}^{N^{(1r)}} q\left(h_i^{(1r)}\middle|v\right)\prod_{j=1}^{N^{(2r)}} q\left(h_j^{(2r)}\middle|v\right)\right)\left(\prod_{i=1}^{N^{(1s)}} q\left(h_i^{(1s)}\middle|v\right)\prod_{j=1}^{N^{(2s)}} q\left(h_j^{(2s)}\middle|v\right)\right) * \left(\prod_{i=1}^{N^{(1t)}} q(h_i^{(1t)}|v)\prod_{j=1}^{N^{(2t)}} q(h_j^{(2t)}|v)\right)\prod_{k=1}^{N^{(3)}} q(h_k^{(3)}|v) \tag{34}$$

The mean field parameters can be derived relatively and given in equations 35. The procedure for training a Multimodal DBM is shown in the algorithm presented in the following. This algorithm is applied in chapter 6 for training a Multimodal DBM.

The following algorithm is repeated for $t = 0, \dots, T$ number of iterations:

1. For each of the training samples $v_n = \{v_n^r, v_n^s, v_n^t\}$, and $n = 1, \dots, N$ the mean-field parameters are updated. In 35 formulation, the mean field parameters are presented for one pathway. The mean-field parameters for the second and third pathways are achieved similar to those of the first pathway.

$$\begin{cases} \overbrace{\mu_{\mathrm{e}}^{(1\mathrm{r})} = G\left(\sum_{\mathrm{d}=1}^{\mathrm{N}^{(0\mathrm{r})}} W_{\mathrm{de}}^{(1\mathrm{r})}\mathrm{v}_{\mathrm{d}}^{(\mathrm{r})} + \sum_{\mathrm{f}=1}^{\mathrm{N}^{(2\mathrm{r})}} W_{\mathrm{ef}}^{(2\mathrm{r})}\mu_{\mathrm{f}}^{(2\mathrm{r})}\right), \mu_{\mathrm{f}}^{(2\mathrm{r})} = G\left(\sum_{\mathrm{e}=1}^{\mathrm{N}^{(2\mathrm{r})}} W_{\mathrm{ef}}^{(2\mathrm{r})}\mu_{\mathrm{e}}^{(1\mathrm{r})} + \sum_{\mathrm{m}=1}^{\mathrm{N}^{(3)}} W_{\mathrm{fm}}^{(3\mathrm{r})}\mu_{\mathrm{m}}^{(3)}\right)}^{\text{Mean field parameters for the musical note pitches pathway (first pathway)}} \\ \overbrace{\mu_{\mathrm{m}}^{(3)} = G\left(\sum_{\mathrm{f}=1}^{\mathrm{N}^{(2\mathrm{r})}} W_{\mathrm{fm}}^{(3\mathrm{r})}\mu_{\mathrm{f}}^{(3\mathrm{r})} + \sum_{\mathrm{i}=1}^{\mathrm{N}^{(2\mathrm{s})}} W_{\mathrm{im}}^{(3\mathrm{s})}\mu_{\mathrm{i}}^{(3\mathrm{s})} + \sum_{\mathrm{l}=1}^{\mathrm{N}^{(2\mathrm{t})}} W_{\mathrm{lm}}^{(3\mathrm{t})}\mu_{\mathrm{l}}^{(3\mathrm{t})}\right)}^{\text{Mean field parameter for the Conjoint layer:}} \end{cases} \tag{35}$$

2. k-steps of alternate Gibbs sampling is performed. The vector $\tilde{v}$ is reconstructed by sampling from the hidden units and the hidden units $\tilde{h}$ are resampled from $\tilde{v}$. The stochastic approximations of the conditional probabilities are achieved for one pathway by the following formula. The related conditional probabilities for the other pathways are achieved similarly.

$$
\begin{cases}
p\left(h_e^{(1r)} = 1 | v^r, h^{(2r)}\right) = G\left(\sum_{d=1}^{N^{(0r)}} W_{de}^{(1r)} v_d^r + \sum_{f=1}^{N^{(2r)}} W_{ef}^{(2r)} h_f^{(2r)} + b_e^{(1r)}\right) \\
p\left(h_f^{(2r)} = 1 | h^{(1r)}, h^{(3)}\right) = G(\sum_{f=1}^{N^{(1r)}} W_{ef}^{(2r)} h_e^{(1r)} + \sum_{m=1}^{M} W_{fm}^{(3r)} h_m^{(3)} + b_f^{(2r)}) \\
p\left(h_m^{(3)} = 1 | h^{(2)}\right) = G(\sum_{f=1}^{N^{(2r)}} W_{fm}^{(3r)} h_f^{(2r)} + \sum_{i=1}^{N^{(2s)}} W_{im}^{(3s)} h_m^{(2s)} + \sum_{l=1}^{N^{(2t)}} W_{lm}^{(3t)} h_m^{(2t)} + b_m^{(3)} \\
p\left(v_d^{(r)} = 1 | h^{(1r)}\right) = \frac{\exp(\sum_{e=1}^{E} W_{de}^{(1r)} h_e^{(1r)} + b_d^{(r)})}{\sum_{d=1}^{D^{(r)}} \exp(\sum_{e=1}^{E} W_{de}^{(1r)} h_e^{(1r)} + b_d^{(r)})}
\end{cases}
\tag{36}
$$

3. The model parameters are updated. In the formulations below the parameter update for weights in the first pathway is given. The parameter update for the other pathways are obtained in a similar way.

$$
\begin{cases}
\text{Parameter update for weights in the first pathway:} \\
W_{t+1}^{(1r)} = W_t^{(1r)} + \gamma_t(\frac{1}{N}\sum_{n=1}^{N} v_n^{(r)} \left(\mu_n^{(1r)}\right)^T - \frac{1}{W}\sum_{w=1}^{W} \tilde{v}_{t+1,w}^{(r)} \left(\tilde{h}_{t+1,w}^{(2r)}\right)^T \\
W_{t+1}^{(2r)} = W_t^{(2r)} + \gamma_t(\frac{1}{N}\sum_{n=1}^{N} \mu_n^{(1r)} \left(\mu_n^{(2r)}\right)^T - \frac{1}{W}\sum_{w=1}^{W} \tilde{h}_{t+1,w}^{(1r)} \left(\tilde{h}_{t+1,w}^{(2r)}\right)^T \\
\text{Parameter Update for weights in the conjoint section:} \\
W_{t+1}^{(3r)} = W_t^{(3m)} + \gamma_t(\frac{1}{N}\sum_{n=1}^{N} \mu_n^{(2r)} \left(\mu_n^{(3)}\right)^T - \frac{1}{S}\sum_{s=1}^{S} \tilde{h}_{t+1,s}^{(2r)} \left(\tilde{h}_{t+1,s}^{(3)}\right)^T \\
W_{t+1}^{(3s)} = W_t^{(2s)} + \gamma_t(\frac{1}{N}\sum_{n=1}^{N} \mu_n^{(2s)} \left(\mu_n^{(3)}\right)^T - \frac{1}{S}\sum_{s=1}^{S} \tilde{h}_{t+1,s}^{(2s)} \left(\tilde{h}_{t+1,s}^{(3)}\right)^T \\
W_{t+1}^{(3t)} = W_t^{(2t)} + \gamma_t(\frac{1}{N}\sum_{n=1}^{N} \mu_n^{(2t)} \left(\mu_n^{(t)}\right)^T - \frac{1}{S}\sum_{s=1}^{S} \tilde{h}_{t+1,s}^{(2t)} \left(\tilde{h}_{t+1,s}^{(3)}\right)^T
\end{cases}
\tag{37}
$$

Appendix B. The Human Evaluation in Terms of Musicality, and Rhythmic Structures

The following material adds further detail to the material given in Chapter 4, and presents the human evaluation of LPM audio and visual outputs derived from 88 CA rules and 20 pattern-matching rules. The tables in Appendix B illustrate the 88 CA rules based on the Wolfram four classes and Li and Packard (Li; Packard, 1990) extensions on the second class.

Table B-1.The human evaluation on Wolfram's class one on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class	Rule	Pattern Metric (Frequency)																				Human evaluation	
		PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 1: Homogeneous	0	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	8	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	32	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	40	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	128	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	136	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	160	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	168	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y

Tables B-1 through to B-6 present the human evaluation on the extended Wolfram CA classes. B-1 shows Class1. Tables B-2 to B-4 the three subtypes for Class2. Class3 is given in Table B-5 and the fourth Class in Table B-6. Table C-7 shows how the 256 CA rules are mapped onto 88 through conjugation and reflection as explained in Chapter 3.

Table B-2. The human evaluation on first subclass of Wolfram’s second class on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class		Rule	Pattern Metric (Frequency)- for each																				Human evaluation	
			PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 2	Fixed Point	2	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		4	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		10	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		12	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		13	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		24	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		34	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		36	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		42	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		44	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		46	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		56	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		57	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		58	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		72	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		76	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		77	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		78	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		104	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		130	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		132	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		138	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		140	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		152	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		162	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		164	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		170	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		172	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		184	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		200	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		204	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
		232	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y

In tables B-1 through C-6, the third column (PM1) through to twenty-second (PM20) columns show the pattern-matching rules investigated. The last two columns show the opinion of a human subject on the LPM musical outputs in terms of melody, and rhythm. All the melodies have been considered to be rhythmic, due to the sound synthesizer changes in each CA time step. Yellow cells show constant audio tone, dark blue oscillatory, and red disordered audio fluctuations. ‘y’ (yes), and ‘n’ (no) indicate the musicality or non-musicality of the audio. ‘l’ stands for low level of musicality.

Table B-3.The human evaluation on second subclass of Wolfram's second class on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class		Rule	Pattern Metric (Frequency)																				Human evaluation	
			PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 2	Periodic	1	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		3	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		5	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		6	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		7	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		9	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		11	n	n	n	n	n	n	n	n	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		14	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		15	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		19	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		23	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		25	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		27	n	y	n	y	n	y	n	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		28	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		29	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		33	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		35	y	y	y	y	y	n	y	y	y	y	n	y	y	y	y	n	n	n	n	n	l	y
		37	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		38	y	y	y	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		41	y	n	y	n	y	y	n	n	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		43	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		50	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	n	y	y	y	y	l	y
		51	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	y	y	l	y
		62	y	y	y	y	y	y	y	y	y	y	y	y	n	n	n	n	y	y	y	y	l	y
		74	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	l	y
		94	n	y	n	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		108	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		134	n	n	n	n	n	n	n	n	y	n	n	n	y	n	n	n	n	n	n	n	l	y
		142	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		156	n	n	n	n	y	n	n	n	y	n	n	n	y	n	n	n	n	n	n	n	l	y
		178	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y

Table B-4.The human evaluation on third subclass of Wolfram's second class on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class		Rule	Pattern Metric (Frequency)																				Human evaluation	
			PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 2	Locally Chaotic	26	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		73	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
		154	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y

Table B-5.The human evaluation on Wolfram's third class on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class	Rule	Pattern Metric (Frequency)																				Human evaluation	
		PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 3: Globally Chaotic	18	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	y
	22	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	30	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	l	y
	45	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	60	y	y	y	n	n	n	n	y	y	n	n	y	y	y	y	y	y	y	y	y	y	y
	90	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	105	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	106	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	122	n	y	n	y	n	y	n	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	126	n	y	n	y	n	y	n	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	146	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	150	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	y	y	l	y

Table B-6. The human evaluation on Wolfram's fourth class on twenty pattern-matching rules in terms of musicality, and rhythmic structures.

Class	Rule	Pattern Metric (Frequency)																				Human evaluation	
		PM1	PM2	PM3	PM4	PM5	PM6	PM7	PM8	PM9	PM10	PM11	PM12	PM13	PM14	PM15	PM16	PM17	PM18	PM19	PM20	Melodic	Rhythmic
Class 4: Complex	54	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y
	110	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y

Table B-7. List of Equivalent CA rule numbers

The left most column shows the 88 fundamental CA rules. Column 2 the equivalent rule due to conjugation (of cell order in rules). Column 3 shows the equivalent rule due to reflection. Column 4 equivalent rules due to both conjugation and reflection (as explained in Chapter 3). The rules are grouped vertically (and colour coded) according to the extended Wolfram classes (Class 1, Class2.1, Class 2.2, Class 2.3, Class 3, Class 4).

Rule	Conjugate	Reflection	C & R
0	255	0	255
8	239	64	253
32	251	32	251
40	235	96	249
128	254	128	254
136	238	192	252
160	250	160	250
168	234	224	248
2	191	16	247
4	223	4	223
10	175	80	245
12	207	68	221
13	79	69	93
24	231	66	189
34	187	48	243
36	219	36	219
42	171	112	241
44	203	1	217
46	75	101	89
56	227	98	185
57	99	99	57
58	163	114	177
72	237	72	237
76	205	76	205
77	77	77	77
78	141	92	197
104	233	104	233
130	190	144	246
132	222	132	222
138	174	208	244
140	206	196	220
152	230	194	188
162	186	176	242
164	218	164	218
170	170	240	240
172	202	228	216
184	226	226	184
200	236	200	236
204	204	204	204
232	232	232	232
1	127	1	127
3	63	17	119
5	95	5	95

6	159	20	215
7	31	21	87
9	111	65	125
11	47	81	117
14	143	84	213
15	15	85	85
19	55	19	55
23	23	23	23
25	103	67	61
27	39	83	53
28	199	70	157
29	71	71	29
33	123	33	123
35	59	49	115
37	91	37	91
38	155	52	211
41	107	97	121
43	43	113	113
50	179	50	179
51	51	51	51
62	131	118	145
74	173	88	229
94	133	94	133
108	201	108	201
134	158	148	214
142	142	212	212
156	198	198	156
178	178	178	178
26	167	82	181
73	109	73	109
154	166	210	180
18	183	18	183
22	151	22	151
30	135	86	149
45	75	101	89
60	195	102	153
90	165	90	165
105	105	105	105
106	169	120	225
54	147	54	147
110	137	124	193

Appendix C. Examples of LPM Musical Outputs, and their Zipfian Distribution

In this appendix, some examples are presented for illustrating the behaviours of pattern-matching rule outputs and their related Zipfian slopes. Figures C-1, C-3, C-5, C-7, C-9, C-11, C-13, C-15, C-17, and C-19 illustrate examples of pattern-matching rules outputs behaviours for CA rule numbers 168, 184, 11, 27, 38, 51, 73, 22, 146, and 110, respectively. The values of 20 pattern-matching rules for the CA rules over 10000 iterations were extracted. The initial seeds for CA were selected randomly. The horizontal coordinates show CA iterations, the vertical coordinate stand for pattern-matching outputs.

The graphs suggest that for some CA rules, there are oscillations occurring at the beginning of the CA progression before the CA reach a stable state. Converging to stable states requires more iteration in some of the graphs (e.g. CA rule 110, pattern-matching 5) while this happens very quickly for other rules (e.g. CA rule 11, pattern-matching 2). The behaviour of CA were studied after a certain number of progressions. After becoming stable, the CA progressions for various rules show different behaviours. Some stay on the same value while others start fluctuating between different values. In the cases where CA progressions converge to a value, the result would be a monotonous audio. The oscillation ranges from two values to more than thirty in some cases. In the auditory tests, these phenomena show themselves as oscillations between two or more pitches.

In this section, Zipfian metrics are employed as an aesthetical measurement to study LPM outputs. The values achieved from the pattern-matching rules for the CA rules were applied for investigating the behaviour of LPM in terms of Zipfian distribution. The pattern-matching rules outputs for each of the CA rules were ranked in compliance with their redundancy (a stage in the procedure for obtaining Zipfian slopes). Linear regression was applied on the rank and frequency of occurrence of the pattern-matching rule values. The obtained slopes and R-squared measurements characterize the Zipfian distribution and the precision of the linear regression fit respectively, (the procedure for determining Zipfian slopes are described in chapter 2).

The values of the twenty pattern-matching rules were extracted from 10000 iterations of CA progression. The Zipfian distribution characteristics of LPM outputs were studied from the five-hundredth to ten-thousandth CA iterations. This time delay is to let the CA progressions to reach stability after the initial state. Figures C-2, C-4, C-6, C-8, C-10, C-12, C-14, C-16, C-18, C-20 illustrate the linear regression lines fitted to Zipfian data distribution of LPM outputs for CA rules 168, 184, 11, 27, 38, 51, 73, 22, 146, and 110, respectively. The horizontal coordinates show the logarithm rank, the vertical coordinate stand for logarithm frequency of occurrence of the pattern-matching outputs.

It should be emphasized that in this appendix, only 1 out of 2^{100} possible initial CA configurations was selected. The initial seed for the CA progressions was selected randomly and was employed for all the CA rules. Other initial configurations would result in different emerging patterns, therefore the Zipfian slopes would change accordingly. The author of this thesis examined further initial seeds and studied the pattern-matching outputs through visual investigation. The result on the three grouping behaviours persisted for those cases as well.

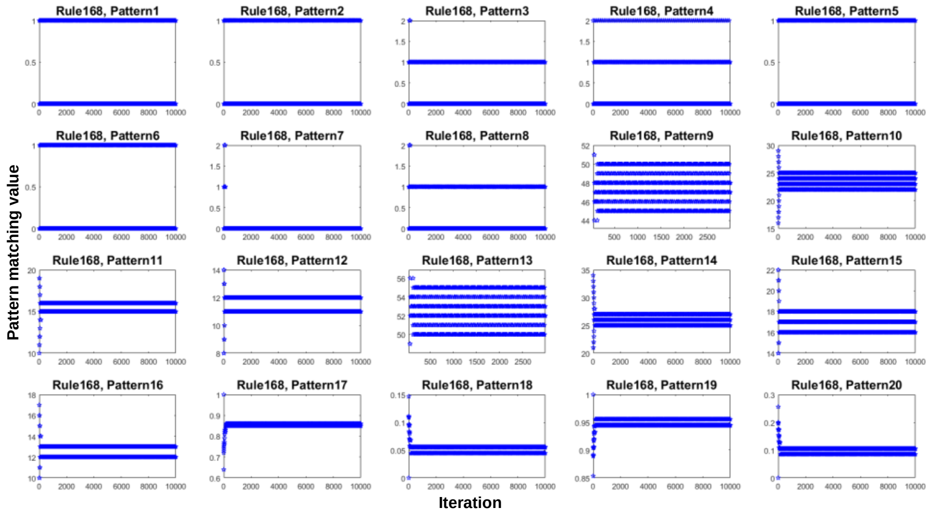


Figure C-1. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 168 (class 1).

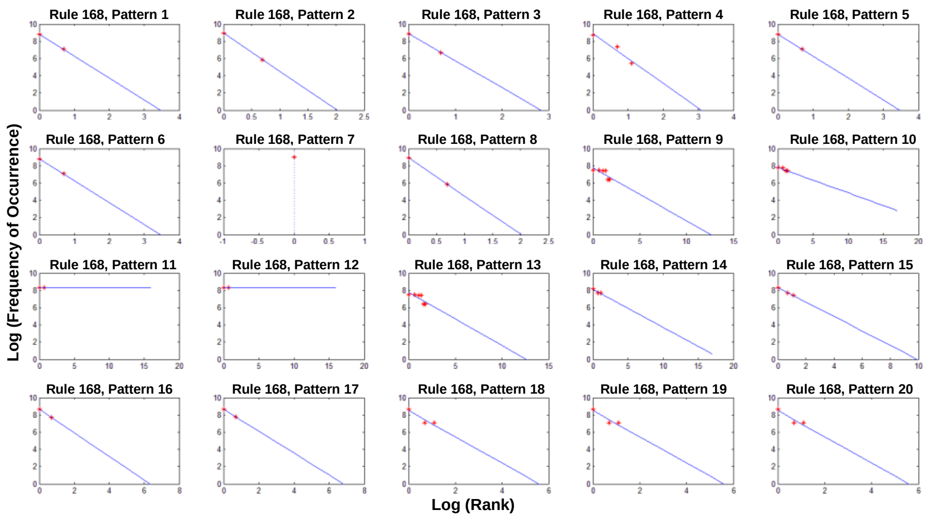


Figure C-2. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 168 (class 1).

168	-2.53751324419525	-4.43438994288255	-3.12674678123601	-2.87761417283276	-2.53751324419525	-2.53751324419525	-Inf	-4.43438994288255	-0.614249741117966	-0.294492353313452	-0.000303693305171152	-0.000303693305171152	-0.614249741117966	-0.4442229764708973	-0.838756695588165	-1.36654943131832	-1.26479880214243	-1.51893711767086	-1.51893711767086	-1.51893711767086

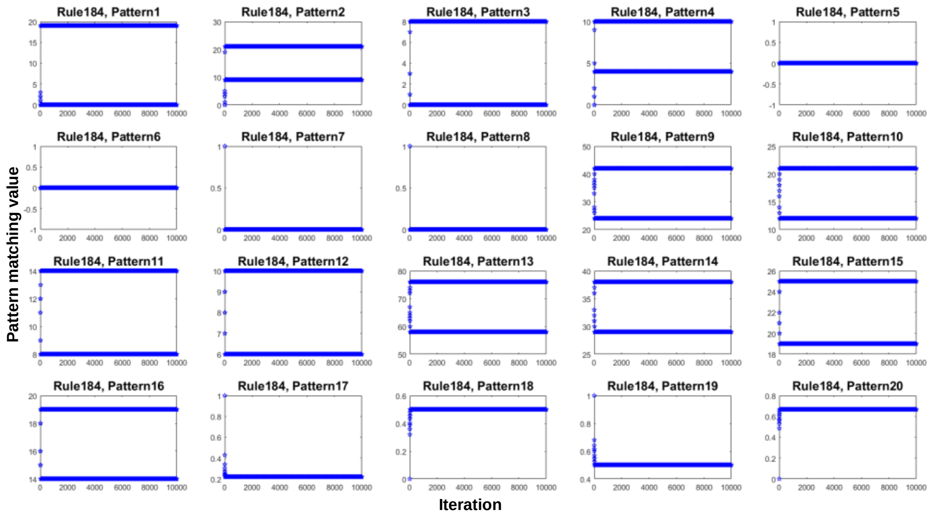


Figure C-3. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 184 (class 2, category 1).

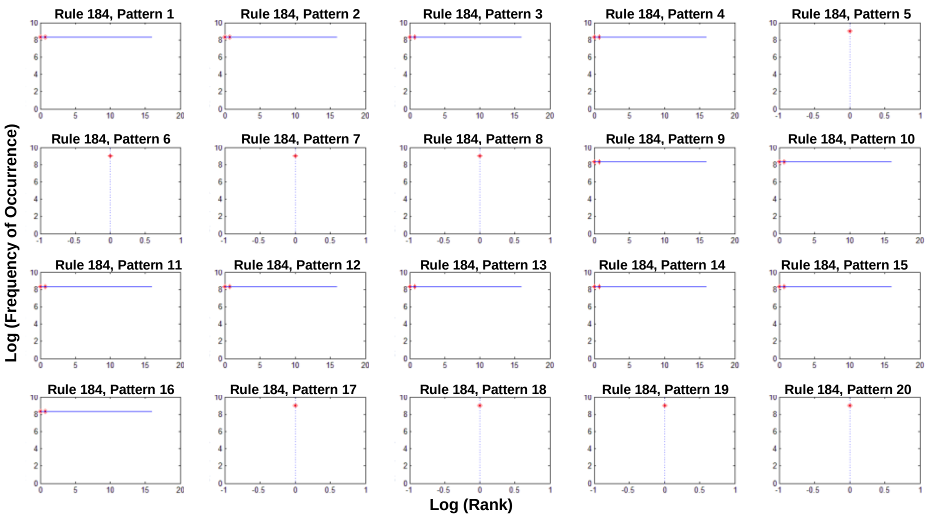


Figure C-4. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 184 (class 2, category 1).

184	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf

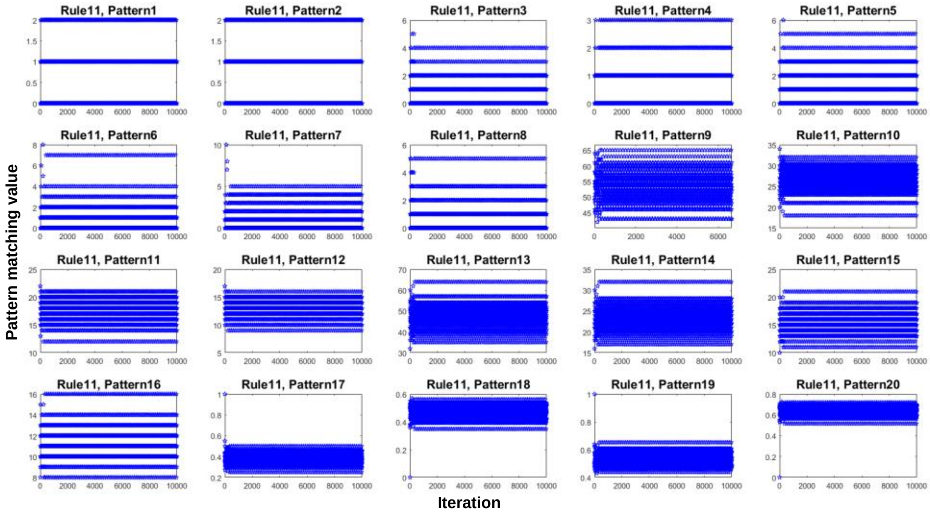


Figure C-5. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 11 (class 2, category 2).

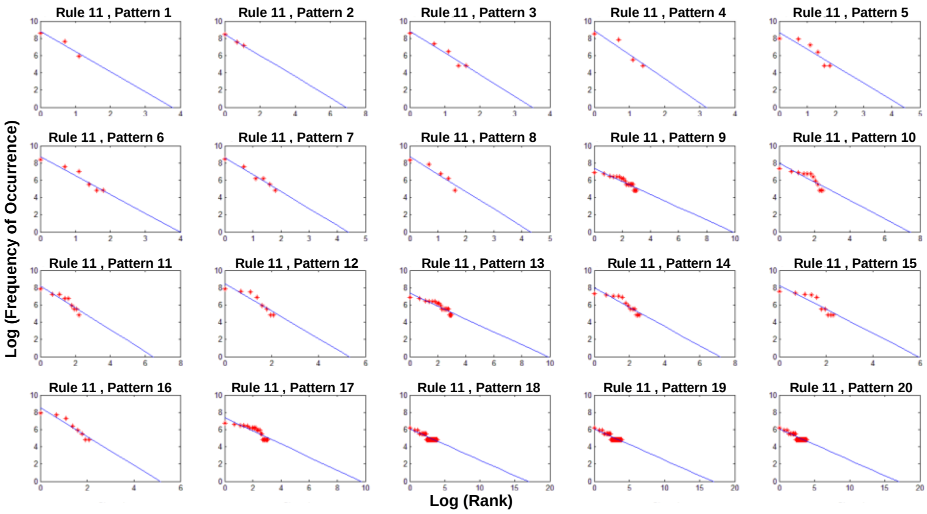


Figure C-6. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 11 (class 2, category 2).

11	-2.33101666717715	-1.21888659915224	-2.52025681136131	-2.78806821736908	-1.94225888389328	-2.19626152955936	-1.96875598929180	-2.00947142971970	-0.75090790460261	-1.07047403221575	-1.27432621722710	-1.59538219534022	-0.75090790460261	-1.12029113936953	-1.39865224441519	-1.66288401551120	-0.76149156777882	-0.35996748020838	-0.35996748020838	-0.35996748020838

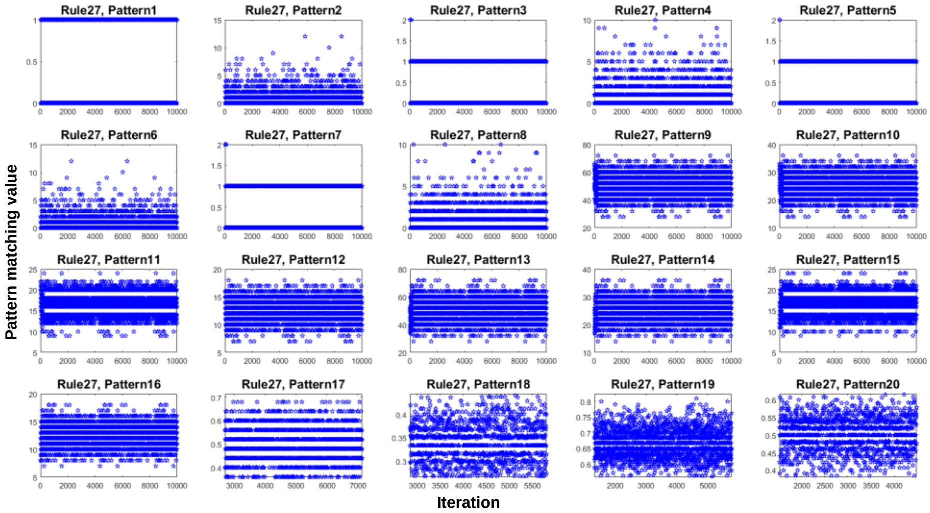


Figure C-7. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 27 (class 2, category 2).

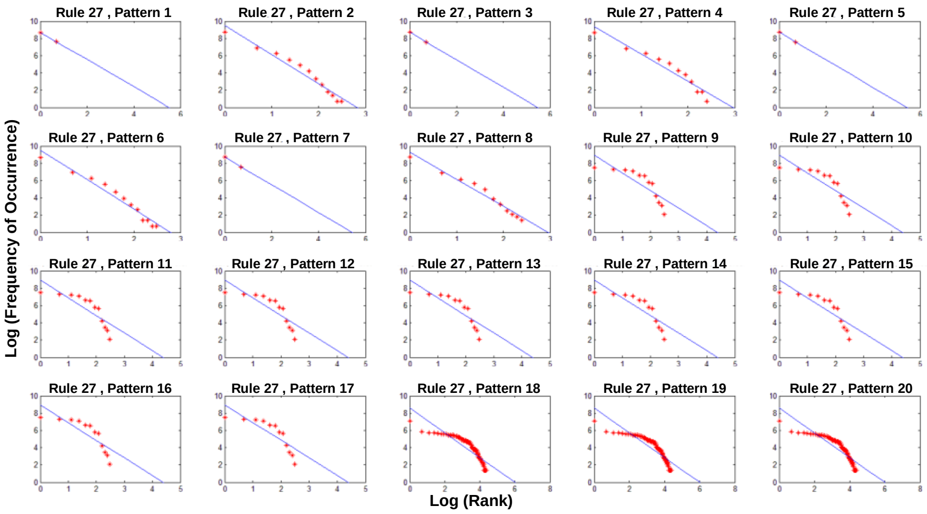


Figure C-8. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 27 (class 2, category 2).

27	- 1.54 255 383 970 956	- 3.37 298 267 621 419	- 1.57 142 924 534 303	- 3.15 979 658 523 868	- 1.60 630 002 317 112	- 3.30 450 000 143 676	- 1.61 447 065 696 810	- 3.13 806 027 447 450	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 781 351 689 618	- 1.99 783 831 287 390	- 1.40 718 003 725 778	- 1.40 718 003 725 778	- 1.40 718 003 725 778

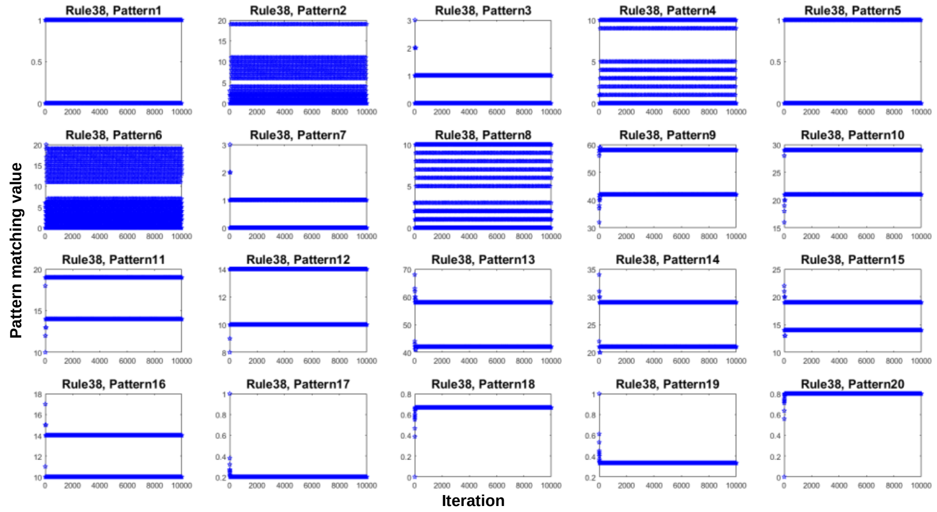


Figure C-9. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 38 (class 2, category 2).

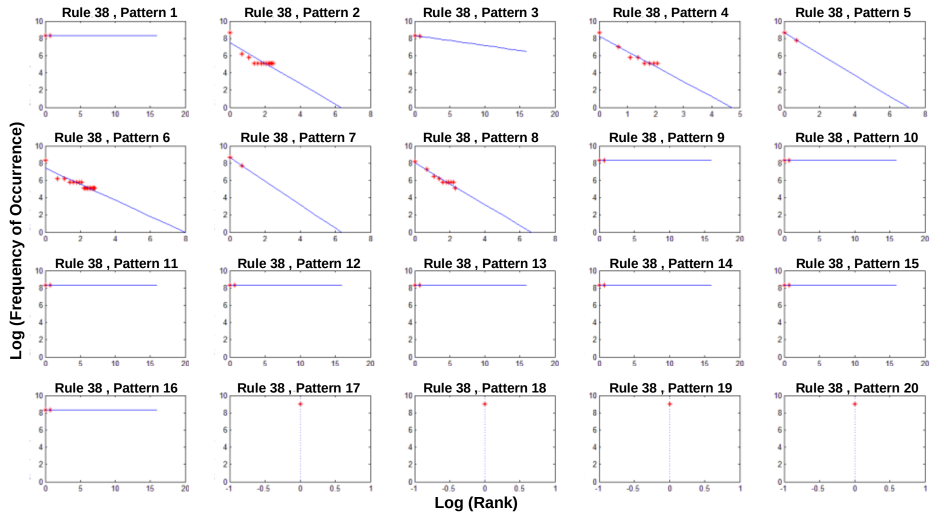


Figure C-10. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 38 (class 2, category 2).

38	0.00 0303 6933 0517 1152	1.18 238 434 419 007	0.11 576 923 139 540 7	1.74 945 125 372 261	1.22 260 935 164 900	0.93 043 435 525 140 9	1.36 278 098 429 567	1.19 995 939 158 856	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	0.00 0303 6933 0517 1152	- Inf	- Inf	- Inf	- Inf

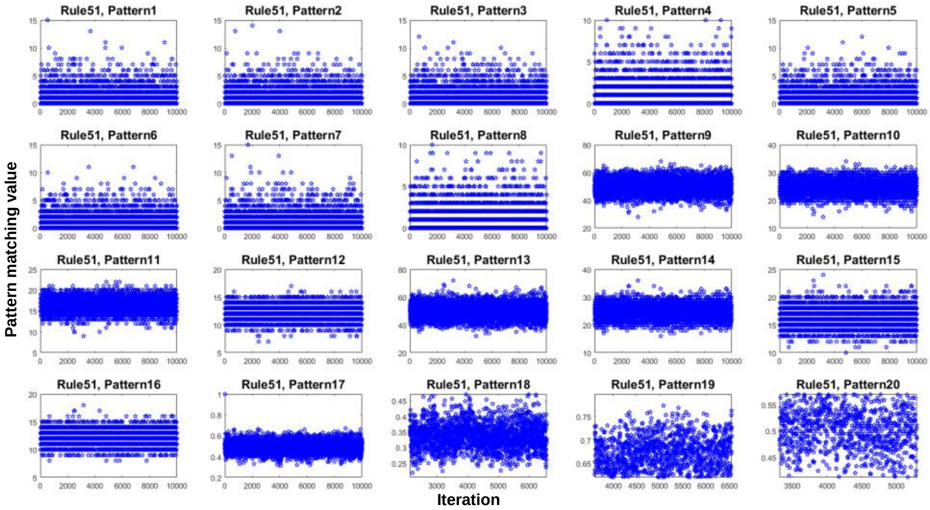


Figure C-11. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 51 (class 2, category 2).

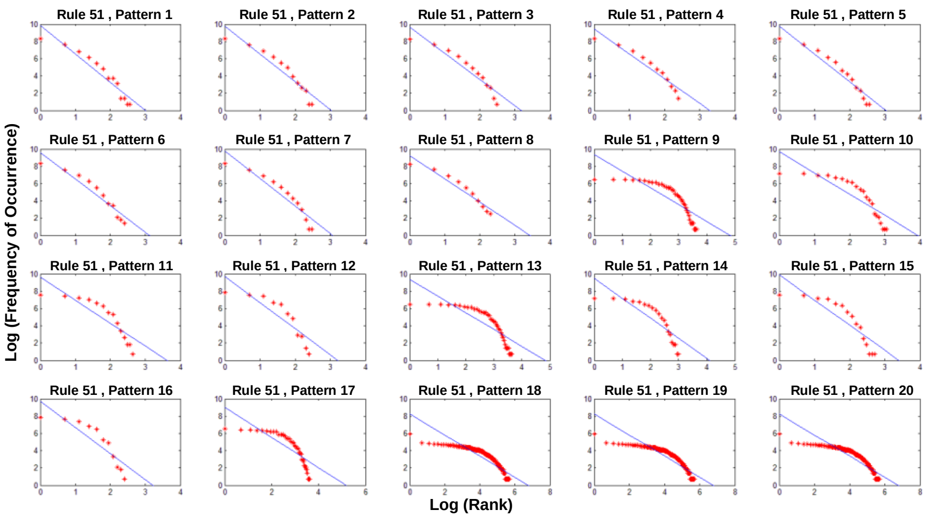


Figure C-12. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 51 (class 2, category 2).

51	- 3.30 2248 7899 6813	- 3.18 1294 1866 2382	- 3.19 3609 3927 8328	- 2.84 8768 5258 588	- 3.28 0021 9257 903	- 3.28 1435 4699 777	- 3.34 1797 7010 6825	- 3.00 1788 9397 788	- 1.95 3744 1091 2868	- 2.51 7933 4325 214	- 2.71 3897 3736 3692	- 3.08 8272 1404 185	- 1.95 3744 1091 286	- 2.37 1999 6230 1500	- 2.96 5704 8640 266	- 3.05 9717 8447 451	- 1.73 0367 9990 949	- 1.26 1785 2115 082	- 1.26 1785 2115 082	- 1.26 1785 2115 0829

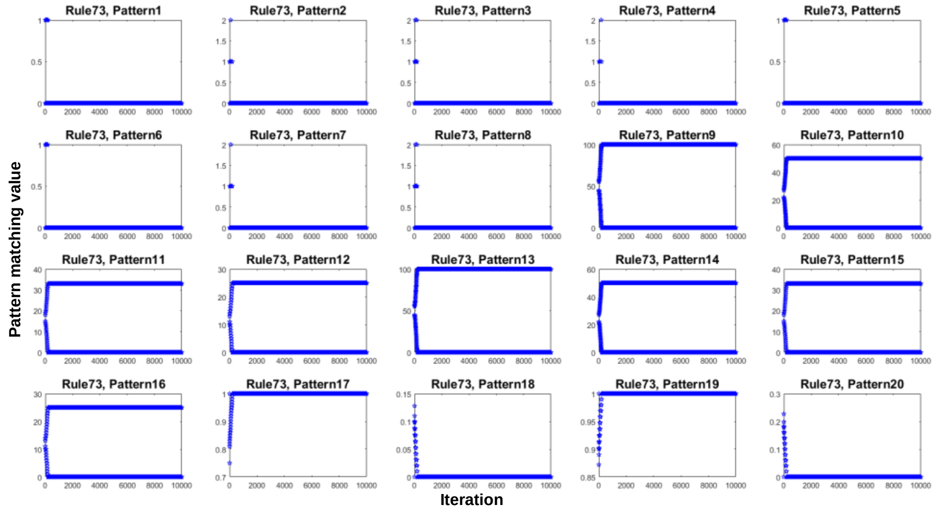


Figure C-13. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 73 (class 2, category 3).

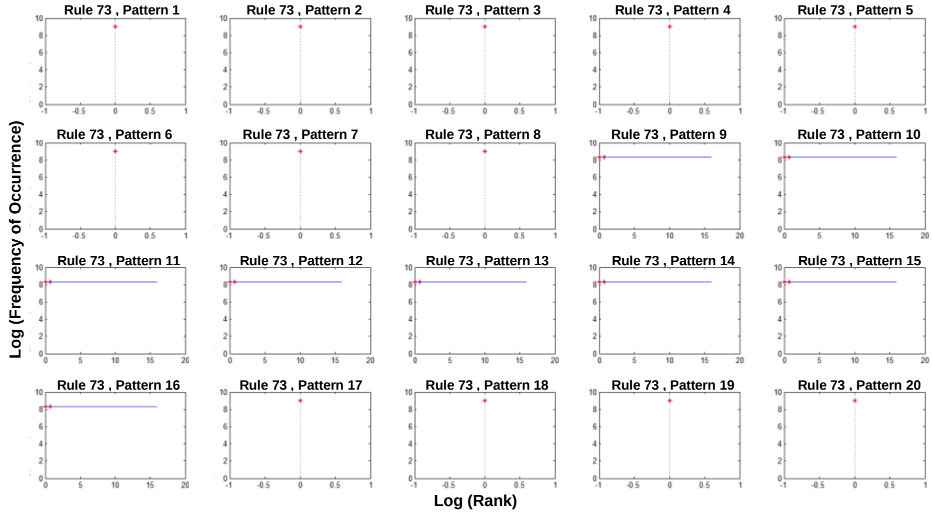


Figure C-14. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 73 (class 2, category 3).

73	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171115	-0.00030369330517 11	-0.000303693305171115	-0.00030369330517 11	-0.00030369330517 11	-0.000303693305171115	-0.00030369330517 11	-0.00030369330517 11	-Inf	-Inf	-Inf	-Inf

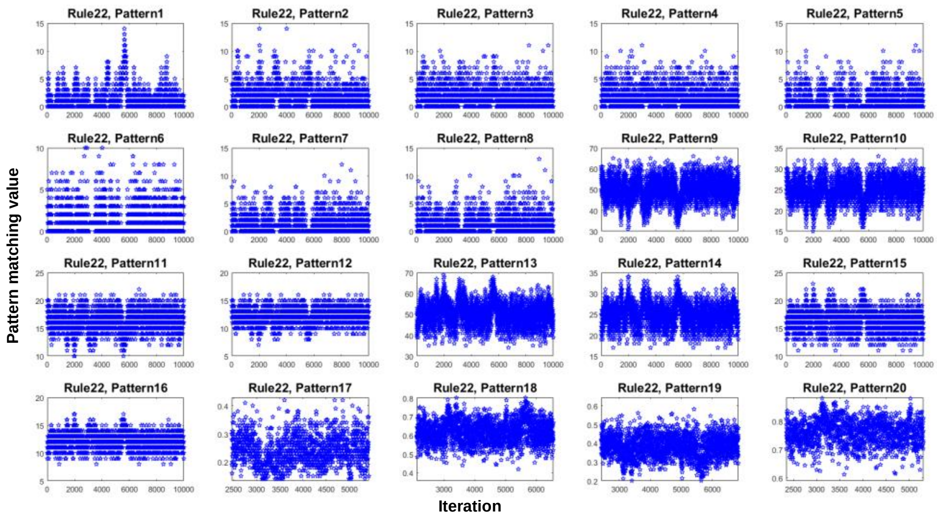


Figure C-15. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 22 (class 3).

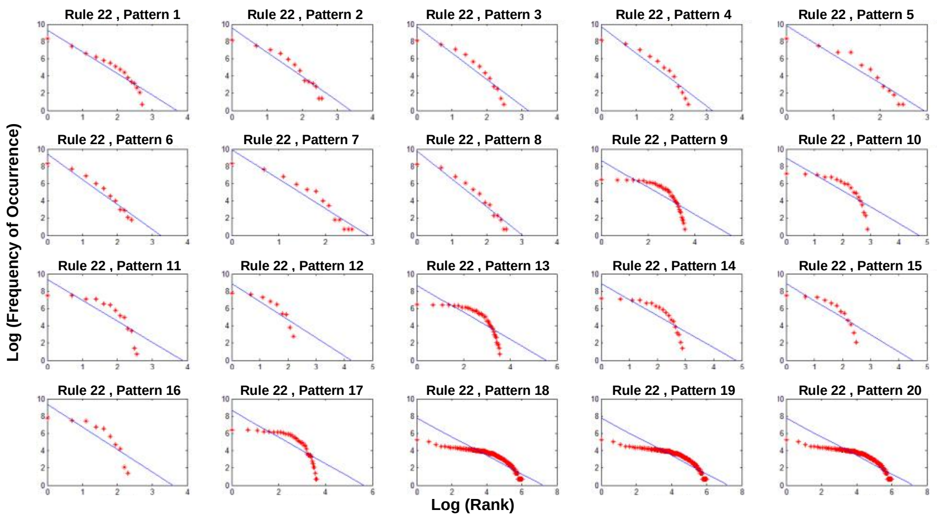


Figure C-16. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 22 (class 3).

22	-2.60019049261838	-2.86788998064152	-2.96650875807881	-3.07443282168699	-3.26887751907091	-2.94249708593886	-3.44920881688320	-3.29082449243973	-1.64334689960123	-2.08858113629721	-2.39825282075211	-2.64609410214504	-1.64334689960123	-1.81808201498454	-2.35274614485377	-2.55506250231225	-1.53074292594956	-1.13303637590092	-1.13303637590092	-1.13303637590092

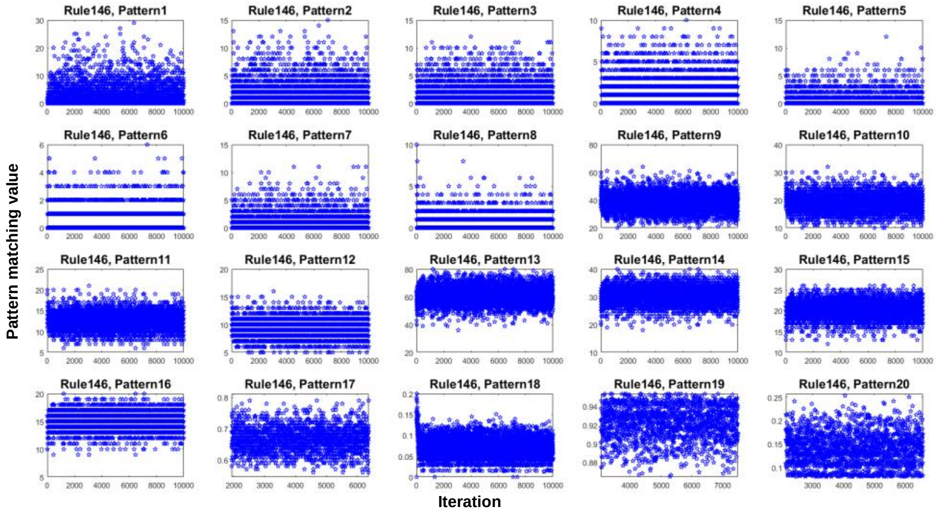


Figure C-17. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 146 (class 3).

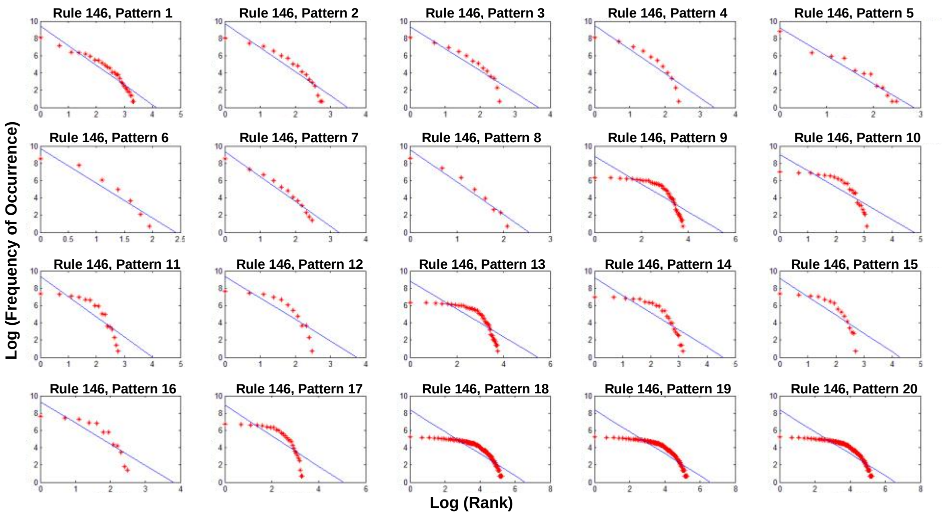


Figure C-18. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 146 (class 3).

146	- 2.28 877 760 572 064	- 2.81 724 385 729 942	- 2.57 304 890 049 883	- 2.83 129 335 346 907	- 3.30 843 965 588 891	- 4.07 261 479 492 897	- 2.96 589 822 307 375	- 3.81 596 240 837 952	- 1.62 519 774 986 656	- 1.89 375 689 501 333	- 2.34 018 030 992 557	- 2.52 095 930 419 100	- 1.62 519 774 986 656	- 2.02 991 700 426 741	- 2.15 328 102 199 933	- 2.46 128 864 020 528	- 1.82 287 378 467 168	- 1.31 209 323 534 482	- 1.31 209 323 534 482	- 1.31 209 323 534 482

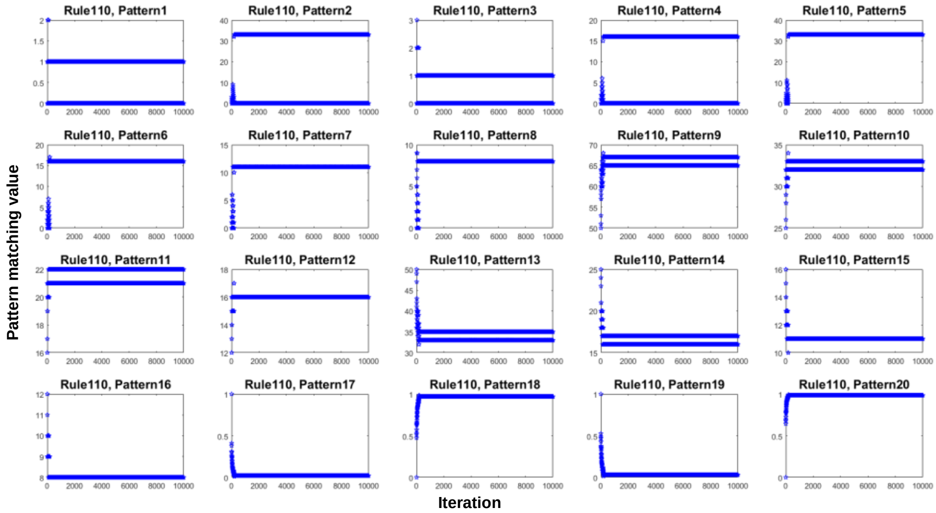


Figure C-19. The twenty pattern-matching rule values over 10000 generation of CA progression for rule 110 (class 4).

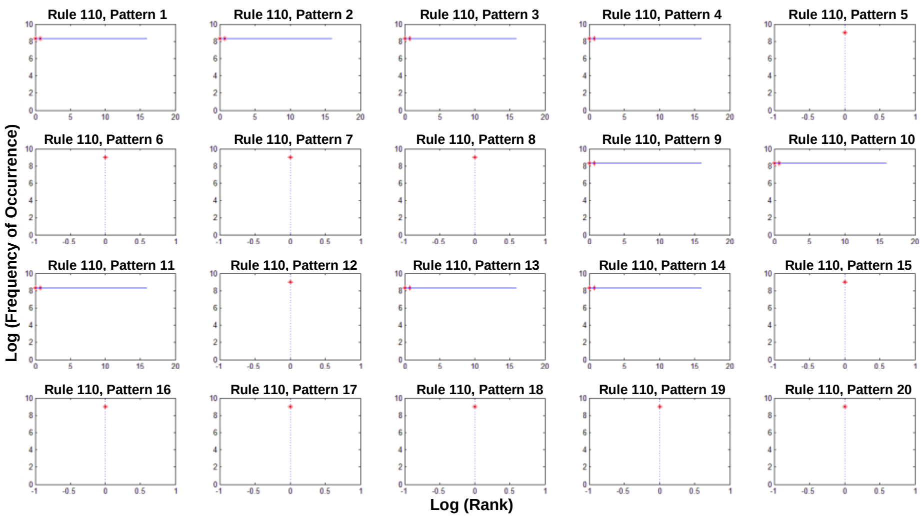


Figure C-20. Rank-frequency distribution of twenty different pattern-matching rules for CA rule 110 (class 4).

110	0.000303693305171152	0.000303693305171152	0.000303693305171152	0.000303693305171152	-Inf	-Inf	-Inf	-Inf	0.000303693305171152	0.000303693305171152	0.000303693305171152	-Inf	0.000303693305171152	0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf

Appendix D. Zipfian Slopes for 88 Rules on 20 Pattern-matching Rule Values

Zipf's law slopes for 88 CA rules on 20 pattern-matching rule values are presented in table D-1. The orange cells indicate the conditions where the distribution follows Zipf's law. The rusty orange cells show the cases (LPM outputs) where the distribution follows Zipf's law (slopes between -0.6 and -2.1), and the auditory tests and studies on the LPM output graphs demonstrated musical LPM outputs for these cases. The yellow cells show the distributions with minus infinity slopes. The cells coloured in dark green colour illustrate other monotonous outputs. The light green cells demonstrate the situations in which the author would expect Zipfian ideal parameters, despite of the obtained parameters that are far away from ideal cases. The light blue cells depict the cases in which the parameters have nearly ideal Zipfian distribution; however, the graphs show the contrary. In this case, the Zipf's distribution is not sufficient for showing the musicality of the data distribution, due to the limited diversity of events occurred. The remaining cells coloured with dark purple colour show slopes near to zero, although the graphs suggest tedious outputs, which are categorized in dark green group.

Table D-1.Zipf's law slopes for 88 rules on twenty pattern-matching rule values.

Rules	Zipfian distribution slopes																			
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
40	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.00036933051171152	-Inf	-Inf	-0.000303693305171152	-0.00036933051171152	-0.000303693305171152	-0.00036933051171152	-0.000303693305171152	-0.00036933051171152	-0.00036933051171152	-0.000303693305171152	-0.00036933051171152	-0.00036933051171152	-Inf	-Inf	-Inf	-Inf
128	-2.86889808374502	-0.349427073858380	-2.87033235510793	-2.10992439316323	-2.86889808374502	-Inf	-2.86889808374502	-0.708447637336429	-0.000303693305171152	-0.00036933051171152	-0.000303693305171152	-0.00036933051171152	-0.00036933051171152	-0.000303693305171152	-0.00036933051171152	-0.00036933051171152	-Inf	-Inf	-Inf	-Inf

136	-3.36984456825047	-3.99693746538514	-5.04142448440311	-3.02547152961185	-3.36984456825047	-3.36607237529102	-Inf	-3.99693746538514	-1.52604199798443	-1.27245760257797	-1.72474457042438	-1.43946749913903	-1.52604199798443	-1.92798142532153	-1.72388284208476	-0.000303693305171152	-2.22306860300799	-2.16162183116543	-2.16162183116543	-2.16162183116543
160	-2.87464167877286	-3.03361470341010	-2.87320415067956	-2.48961878421995	-2.87464167877286	-Inf	-2.87464167877286	-3.35480425113405	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
168	-2.53751324419525	-4.43438994288255	-3.12674678123601	-2.87761417283276	-2.53751324419525	-2.53751324419525	-Inf	-4.43438994288255	-0.614249741117966	-0.294492353313452	-0.000303693305171152	-0.000303693305171152	-0.614249741117966	-0.442229764708973	-0.838756695588165	-1.36654943131832	-1.26479880214243	-1.51893711767086	-1.51893711767086	-1.51893711767086
104	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
184	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
232	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
6	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-Inf	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
7	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-Inf	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
9	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-Inf	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
11	-2.33101666717715	-1.21888659915224	-2.52025681136131	-2.78806821736908	-1.94225888389328	-2.19626152955936	-1.96875598929180	-2.00947142971970	-0.750907904602611	-1.07047403221575	-1.27432621722710	-1.59538219534022	-0.750907904602611	-1.12029113936953	-1.39865224441519	-1.66288401551120	-0.761491567778824	-0.359967480208384	-0.359967480208384	-0.359967480208384
14	-2.27319901534975	-2.68180488365263	-2.98759436467126	-3.44369801268132	-3.30815755567885	-3.80949231838372	-3.30376564228905	-3.82231084769804	-2.12368457986986	-2.68180488365263	-3.16408268120939	-3.44369801268132	-2.12368457986986	-2.70827347266011	-3.16235905723356	-3.12170683452579	-2.05282292282017	-1.41407819856508	-1.41407819856508	-1.41407819856508
15	-1.08045619333033	-1.08045619333033	-1.73325462793220	-0.822210193572998	-1.17094865198241	-1.17094865198241	-2.80822281298247	-0.000303693305171152	-0.299859072645684	-0.473664799012211	-0.464117464353114	-0.675490511572363	-0.299859072645684	-0.473664799012211	-0.463708743461478	-0.675490511572363	-0.549436374296284	-0.000517589024665243	-0.000517589024665243	-0.000517589024665243

19	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
27	-1.54255383970956	-3.37298267621419	-1.57142924534303	-3.15979658523868	-1.60630002317112	-3.30450000143676	-1.61447065696810	-3.13806027447450	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99781351689618	-1.99783831287390	-1.40718003725778	-1.40718003725778	-1.40718003725778
28	-3.11658887973393	-3.31030153135869	-3.10922245346264	-3.24401361878060	-3.76389309060875	-4.21477859258668	-3.86281748643988	-4.13821020237721	-1.58414529242897	-1.98817000202825	-2.07346166002390	-2.24397735679391	-1.58414529242897	-1.79385665733203	-2.08652915552256	-2.18383442571434	-1.58058800925066	-1.23494101201238	-1.23494101201238	-1.23494101201238
29	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
33	-0.597759551203193	-1.95848935179438	-6.48619242956088	-4.00857986537141	-3.16672219522220	-3.13004457920014	-1.82528048060362	-1.52964731007292	-2.98531099717254	-3.52189043347750	-4.18096909788839	-4.82523008838578	-2.98531099717254	-3.89093037551792	-4.89622472015814	-5.01843413311558	-3.51720825525152	-3.32028811388282	-3.32028811388282	-3.32028811388282
35	-2.62161058782897	-5.03084319991532	-2.62161058782897	-2.12386779047051	-2.62161058782897	-Inf	-2.62161058782897	-5.63044141150364	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
37	-2.28716868518440	-2.62861614793166	-3.10139298212567	-3.49270036537138	-1.85624551229643	-2.86212406234970	-2.63454739720836	-3.48382753849939	-1.72929900267323	-1.78825577435479	-1.94605344085146	-2.09000260209059	-1.72929900267323	-1.91835387211787	-2.06925494390181	-2.10964331442877	-2.29506064119269	-1.77876326424888	-1.77876326424888	-1.77876326424888
38	-0.000303693305171152	-1.18238434419007	-0.115769231395407	-1.74945125372261	-1.22260935164900	-0.930434355251409	-1.36278098429567	-1.19995939158856	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
41	-5.04142448440311	-Inf	-5.04142448440311	-Inf	-5.04142448440311	-5.04142448440311	-Inf	-Inf	-1.75049433156596	-0.505614743724921	-0.478698845133068	-0.505614743724921	-1.75049433156596	-0.400668535821333	-0.000303693305171152	-0.858378644662466	-3.99693746538514	-3.35130598959770	-3.35130598959770	-3.35130598959770
50	-2.98113430741395	-3.45928165988972	-2.52134193949427	-1.09678607476421	-1.38026102910982	-1.78926141556167	-4.27473977235319	-2.65138357286510	-0.685642103763799	-0.730704767336060	-0.000911079942429926	-1.78838909575076	-0.685642103763799	-1.79336583346929	-2.25146418945067	-0.000911079942429926	-1.36954847791075	-1.36954847791075	-1.36954847791075	-1.36954847791075
51	-3.30224878996813	-3.18129418662382	-3.19360939278328	-2.84876852585882	-3.28002192579038	-3.28143546997772	-3.34179770106825	-3.00178893977887	-1.95374410912868	-2.51793343252148	-2.71389737363692	-3.08827214041852	-1.95374410912868	-2.37199962301500	-2.96570486402663	-3.05971784474519	-1.73036799909498	-1.26178521150829	-1.26178521150829	-1.26178521150829
62	-2.35805663204202	-0.149605437946640	-2.60328899239272	-0.201685003174232	-2.19356492778795	-2.74176538399531	-1.44131478720283	-1.44131478720283	-2.11456049824046	-2.50662440484349	-5.06280873905695	-0.504659223763414	-2.11456049824046	-1.04050659009615	-1.04050659009615	-1.04050659009615	-1.91617568975284	-1.91617568975284	-1.91617568975284	-1.91617568975284

74	-1.01753863141668	-1.01185680626990	-1.15072674175094	-1.20180501177213	-1.28856244055444	-1.10011454087994	-0.734552415302815	-1.04568736588146	-0.543357704044288	-0.500367124541545	-0.509727374268955	-0.675905715075423	-0.543357704044288	-0.484262465073295	-0.498969248750465	-0.558450951369431	-2.13798688647960	-1.93830709783437	-1.93830709783437	-1.93830709783437
94	-1.36127481811410	-1.09055949392195	-1.35976934291519	-1.20110231725917	-1.36127481811410	-Inf	-1.36127481811410	-1.43993128749925	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
108	-1.36278098429567	-2.47424958573593	-2.18781218962630	-1.82522947508621	-2.39136859338521	-2.42261227015281	-1.50842967863111	-1.64192555572561	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
134	-2.90800953105414	-6.06316392337829	-3.12674678123601	-3.65224062988376	-2.90654644125553	-2.90947375662252	-Inf	-6.06316392337829	-0.882328984332036	-0.401749393494719	-0.402113265677256	-0.711059948357693	-0.882328984332036	-0.402113265677256	-0.000303693305171152	-0.711627118929712	-1.70088653434480	-1.81170294312123	-1.81170294312123	-1.81170294312123
142	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf
156	-4.58512068229479	-2.25898562603246	-2.59518588004966	-2.72299863544517	-1.24551476238484	-2.43877858698340	-2.53448448084190	-2.59540176212918	-1.17908142368564	-1.62198202569483	-0.407511224593251	-3.73430153744630	-1.17908142368564	-1.62222365317765	-0.406891660541950	-3.73430153744630	-1.58465102267741	-1.58465102267741	-1.58465102267741	-1.58465102267741
178	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
26	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
73	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf
22	-2.60019049261838	-2.86788998064152	-2.96650875807881	-3.07443282168699	-3.26887751907091	-2.94249708593886	-3.44920881688320	-3.29082449243973	-1.64334689960123	-2.08858113629721	-2.39825282075211	-2.64609410214504	-1.64334689960123	-1.81808201498454	-2.35274614485377	-2.55506250231225	-1.53074292594956	-1.13303637590092	-1.13303637590092	-1.13303637590092
30	-0.671152330575478	-0.000303693305171152	-0.670060953358673	-1.17055555654074	-0.670606591411588	-1.08034130862258	-0.000577709074479331	-1.08022637486090	-0.908189561669714	-1.17006095335867	-1.17006095335867	-2.80683437540181	-0.908189561669714	-0.908189561669714	-1.73407439557115	-1.58516497002763	-0.908189561669714	-0.300613358275328	-0.300613358275328	-0.300613358275328
45	-1.66771363159420	-1.89820118168093	-2.10796565752421	-2.29718235334832	-2.23366160179347	-2.26395636069072	-2.29678229974850	-2.26153032252612	-0.810052571438297	-1.26545682360352	-1.45584783169286	-1.74813766301986	-0.810052571438297	-1.26311467846781	-1.37328499017712	-1.82740584881356	-0.877162371697132	-0.401935979648954	-0.401935979648954	-0.401935979648954

60	-0.784973781742093	-1.01705601332141	-0.767803875841366	-1.87708531871201	-Inf	-2.42646374077157	-2.68490153410911	-2.22967130289218	-0.784973781742093	-1.01705601332141	-0.767803875841366	-1.87708531871201	-0.784973781742093	-1.03996880491776	-1.40648352905190	-0.850069856382143	-0.771306580593175	-0.424545030172469	-0.424545030172469	-0.424545030172469
90	-3.07875225632659	-3.19187368904462	-3.32929180218929	-3.26134542721928	-3.26443563495683	-3.08297665566262	-3.25481428260823	-3.09823307230625	-1.62633970120641	-2.20871197491538	-2.77523421901807	-2.43803969109496	-1.62633970120641	-2.22011458456462	-2.43006124889180	-2.94578370012931	-1.61647203532139	-1.21920009010964	-1.21920009010964	-1.21920009010964
105	-3.16346367531278	-3.39596331882193	-3.21953724057831	-3.12804241373739	-3.10399703749138	-3.35330329647718	-3.21945731830437	-2.81830845537188	-1.74252228958708	-2.17321986515998	-2.44818749038918	-2.65766487663125	-1.74252228958708	-2.13984437692477	-2.69672090815339	-2.75030077919646	-1.67679158740110	-1.10436847427392	-1.10436847427392	-1.10436847427392
106	-3.20331838620634	-3.38359741159628	-3.23390900900368	-3.05434097150653	-3.32838135724730	-3.04785630459658	-3.35087778440372	-3.27086594933731	-1.85240083685171	-2.35403215950906	-2.74685173338411	-2.98378497383134	-1.85240083685171	-2.32889299547020	-2.85578457445673	-3.11699668163897	-1.57163915621431	-1.19555906163936	-1.19555906163936	-1.19555906163936
122	-1.57626971371862	-3.02247315475384	-1.59571303723928	-2.83146105924361	-1.57950118805099	-3.20282232394671	-1.54574822048767	-3.15642542292032	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19290851976649	-2.19330986627133	-1.52993384283325	-1.52993384283325	-1.52993384283325
126	-1.52187396507045	-3.33278668650215	-1.53697197692095	-3.00839810388689	-1.60711603307757	-3.44423806670153	-1.66483042333304	-2.98557406426486	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19150457801175	-2.19146787607384	-1.43207643654978	-1.43207643654978	-1.43207643654978
146	-2.28877760572064	-2.81724385729942	-2.57304890049883	-2.83129335346907	-3.30843965588891	-4.07261479492897	-2.96589822307375	-3.81596240837952	-1.62519774986656	-1.89375689501333	-2.34018030992557	-2.52095930419100	-1.62519774986656	-2.02991700426741	-2.15328102199933	-2.46128864020528	-1.82287378467168	-1.31209323534482	-1.31209323534482	-1.31209323534482
150	-3.26243048221058	-3.23182151945605	-3.27185798475379	-2.94018833858932	-2.89809231961069	-2.65922552010022	-3.12684405546412	-3.27653462110087	-1.80665918853615	-2.31726972312488	-2.41857944428260	-2.58982456348769	-1.80665918853615	-2.11354314203007	-2.64209195881055	-2.72379951136418	-1.67679158740110	-1.13125574243462	-1.13125574243462	-1.13125574243462
110	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-0.000303693305171152	-0.000303693305171152	-0.000303693305171152	-Inf	-0.000303693305171152	-0.000303693305171152	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
*****	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf

In the above table the asterisk in the last row represent CA rule numbers: 0, 8, 32, 2, 4, 10, 12, 13, 24, 34, 36, 42, 44, 46, 56, 57, 58, 72, 76, 77, 78, 130, 132,138, 140, 152, 162, 164, 170, 172, 204, 1, 3, 5, 23, 25, 154, 18, 54, 43.

Appendix E. Standardizing the Musical Parameters

This appendix presents the required information for standardizing the Persian music pitch frequencies achieved in chapter 5 by the application of auto-correlation algorithm. This appendix also provides further standardization ranges for LPM sequences, while preparing the training datasets for the machine learning fitness function. The auto-correlation algorithm may achieve slightly different results for the note pitches as they occur in the musical piece. However, obtaining Zipfian metrics for musical events requires the calculation of the number of occurrence of those events. Therefore, further refinements on the achieved pitch values were required. On this account, the concept of pitch bins was introduced for justifying the pitch values. The midpoints of the pitch bins were selected to be the standard pitches of nine-bridged Persian Santur with their various tunings for performance in different Dastgāh. A table of the 27 Santur strings and their pitches (44 notes or pitches in different Dastgāh tunings) are provided here in table E-1.

Table E-1. Different tunings for Santur musical instrument in different Dastgāh. This table is applied for standardizing different tone frequencies during the musical information retrieval from Persian music.

Note Number	Notes in the bass Position	Pitch	Note Number	Notes in the Middle Position	Pitch	Note Number	Notes in the Treble range	Pitch
1	E3𝄳	158	10	E4𝄳	292	19	E5𝄳	584
1	E3	164	10	E4	310	19	E5	624
2	F3	174	11	F4	348	20	F5	696
3	G3	197	12	G4	393	21	G5	787
4	A3𝄳	214	13	A4𝄳	428	22	A5♭	835
4	A3	220	13	A4	440	22	A5	880
5	B3♭	233	14	B4♭	466	23	B5♭	936
5	B3	246	14	B4	493	23	B5	987
6	C4	262	15	C5	523	24	C6	1046
7	D4𝄳	284	16	D5𝄳	570	25	D6𝄳	1076
7	D4	293	16	D5	587	25	D6	1121
8	E4♭	311	17	E5♭	622	26	E6♭	1175
8	E4𝄳	319	17	E5	659	26	E6	1318
8	E4	329	18	F5	697	27	F6	1396
9	F4	349						
9	F4#	370						

In the remainder of this appendix, there would be a look towards how the LPM sequences dataset were prepared. The LPM sequences consist of a consecutive arrangement of voices. Each of the voices can have 11 governing dimensions mentioned in the end of section 4.2 in chapter 4, and the production of databases can continue to 11 dimensions. Table E-2 shows the Lower/Upper ranges values of the musical dimensions for the interested reader. The Attack, Decay, Sustain, Release, Loop filter, and loop gain value ranges are selected in a way that would contribute to the stability of the synthesizer filters. The normalization of the data in each of the dimensions is an important step and is performed in accordance with the synthesizer stability and the feasibility of the produced audio from musical aspects. The synthesizers are often based on a series of digital filters that accept a set of parameters. Certain configurations of the filters may make the filter unstable. This condition is often heard as burst of noise of fluffy sounds from the designed filter. Therefore, the thresholds in table F-2 have been suggested; some of which have been discovered through trial and error in relation to each other and in the application.

Table E-2.The Normalization value ranges for the musical dimensions in LPM sequences

This table shows all the governing dimensions of LPM sequences and their Lower/Upper ranges value. In this thesis, The LPM sequences are evolved based on their pitch frequency, note durations, and note onset times.

Musical Dimension	**Metric**	**Lower Range**	**Upper Range**
Pitch Frequency	Hz	100	3000
Duration	Seconds	1/16	2
Onset	Seconds	1/16	2
Attack	Seconds	0.001	0.5
Decay	Seconds	0.001	4
Sustain	Seconds	0	4
Release	Seconds	0.5	0.9
Loop Filter	Coefficient	0.01	0.05
Loop Gain	Coefficient	0.01	0.05
Interval	Musical Tones	-20	20
Music Speed	Coefficient	1	20

The evolution of melodic structures of LPM sequences is the focus of the research in the first place. Therefore, the pitch, duration, and onset times are taken into account as 3 dimensions out of the 11 dimensions; the other 8 musical dimensions are left for further research. Evolving the audio sequences based on the remaining dimensions is subject for future investigations and are out of the scope of the thesis.

Appendix F. Attribute Selection from Dastgāh Music, and LPM Sequences

This appendix shows how the attribute selection is performed on the cleaned databases consisting of Persian music, and LPM sequences in chapter 5. After performing the initial data cleaning, the second phase would be the attribute selection phase. The attributes are evaluated in Weka by different algorithms such as *ReliefF*, *Gain Ratio*, *InfoGain*, and *Symmetrical Uncertainty* by the application of Ranker search method (Kononenko & Kukar, 2007). The attributes are ranked in descending order by Weka and are assigned scores between zero and one. The given values assist in the detection of competent features. The underlying attribute selection process starts with marking the attributes with evaluation values below 0.1 as undesirable ones. The attributes are sorted by their attribute numbers by the help of Excel custom sort. This would arrange different evaluators' value for each attribute in one row, so that the evaluations can be compared with each other. Colour coding assisted throughout the attribute selection procedure. Those attributes, which have not been marked as undesirable in any of the evaluators' categories, have won the selection process and will enter the final phase of cleaning before the training is performed. 10 attributes remain which have the evaluation values greater than 0.1 among all the rankers. Table F-1 shows the complete attribute selection criteria (yellow coded attributes retained).

Table F-1. The complete table of the attribute Selection procedure. Values less than 0.1 are colour coded. Attributes coded yellow have values greater than 0.1 for all metrics.

	Attribute name	**ReliefF**	**Gain Ratio**	**InfoGain**	**Symmetrical Uncertainty**
1	Pitches	0.0638	0.0862	0.1304	0.1038
2	Chromatic	0.0854	0.1504	0.2489	0.1875
3	Pitch_Distance1	0.0581	0.2671	0.5951	0.3687
4	Contour_Melody_Pitch	0.1122	0.2683	0.5462	0.3599
5	Melodic_Bigrams	0.0232	0.0524	0.0474	0.0498
6	Melodic_Trigrams	0.0205	0.0984	0.1373	0.1146
7	Melodic_4grams	0.0211	0.1193	0.1864	0.1455
8	Contour_Melody_Pitch_d1	0.194	0.3681	0.7087	0.4845
9	Contour_Melody_Duration_d1	0.1309	0.2843	0.6016	0.3862
10	Melodic_Bigram_d1	0.0298	0.1456	0.2104	0.1721
11	Chromatic_DataSet_d2	0.1428	0.3007	0.5296	0.3836
12	Pitch_Distance1_d2	0.0585	0.346	0.6392	0.449
13	Contour_Melody_Pitch_d2	0.0677	0.2372	0.4624	0.3136
14	Contour_Melody_Duration_d2	0.0311	0.0955	0.1797	0.1247
15	Chromatic_DataSet_d3	0.0425	0.1201	0.19	0.1472
16	Pitch_Distance1_d3	0.0365	0.1833	0.3586	0.2426
17	Contour_Melody_Pitch_d3	0.0584	0.1817	0.3993	0.2498
18	Contour_Melody_Duration_d3	0.0613	0.1684	0.3724	0.2319
19	Chromatic_DataSet_d4	0.1394	0.2269	0.4473	0.301
20	Pitch_Distance1_d4	0.0719	0.3735	0.6729	0.4803
21	Contour_Melody_Pitch_d4	0.0739	0.1731	0.391	0.2399
22	Contour_Melody_Duration_d4	0.0535	0.1507	0.3001	0.2007
23	pitches_LV	0.0405	0.0502	0.0446	0.0473
24	Combined_Pitch_Duration_LV	0.0513	0.1957	0.5037	0.2818
25	Contour_Melody_Pitch_LV	0.0389	0.1356	0.2403	0.1734
26	Contour_Melody_Duration_LV	0.1242	0.2161	0.4945	0.3008
27	Melodic_Interval_LV	0.1235	0.2762	0.4584	0.3447
28	Melodic_Bigram_LV	0.0368	0.039	0.0379	0.0384
29	Melodic_Trigrams_LV	0.0374	0.0419	0.0362	0.0389
30	Melodic_4grams_LV	0.0317	0.0287	0.0354	0.0317
31	Rests_LV	0.0378	0.1532	0.2156	0.1791
32	pitches_d1_LV	0.0244	0.0396	0.0626	0.0485
33	Chromatic_DataSet_d1_LV	0.1653	0.2829	0.5545	0.3747
34	Pitch_Distance1_d1_LV	0.0509	0.0958	0.1856	0.1264
35	Contour_Melody_Pitch_d1_LV	0.0886	0.2367	0.4754	0.316
36	Contour_Melody_Duration_d1_LV	0.1644	0.3564	0.7508	0.4833
37	Melodic_Bigram_d1_LV	0.0254	0.0528	0.0877	0.0659
38	Melodic_Trigrams_d1_LV	0.0241	0.052	0.0883	0.0654
39	Melodic_4grams_d1_LV	0.0224	0.0545	0.0903	0.0679

40	Rests_d1_LV	0.0888	0.1801	0.3937	0.2472
41	pitches_d2_LV	0.0709	0.1597	0.2901	0.206
42	Chromatic_DataSet_d2_LV	0.1838	0.4275	0.7631	0.548
43	Pitch_Distance1_d2_LV	0.0525	0.0517	0.0766	0.0617
44	Contour_Melody_Pitch_d2_LV	0.0306	0.0156	0.0148	0.0152
45	Contour_Melody_Duration_d2_LV	0.0639	0.1	0.1757	0.1275
46	Melodic_Bigram_d2_LV	0.0668	0.1478	0.2774	0.1929
47	Melodic_Trigrams_d2_LV	0.0657	0.1422	0.2697	0.1862
48	Melodic_4grams_d2_LV	0.0586	0.1618	0.2441	0.1946
49	Rests_d2_LV	0.0447	0.0439	0.0385	0.041

Table F-2 shows the performance of various machine learning tools trained with different number of attributes in various stages of data cleaning, and attribute selection. The number of features are 248, 49, 10, and 8 in the experiment. The result provided in the table witnesses that the classifiers are still able to perform outstandingly after the dramatic shrinkage of the initial attribute sets.

Table F-2. Classification performance for different subsets of selected attributes

Classifiers SVM, JRip, Ridor,Decision Table,J48, Logistic, RBF, and NaiveBayes for different subsets of selected attributes for 1532 instances in the training data.

	Number of Features	Correctly Classified	Sensitivity	Specificity	PPV	NPV
SVM	248	100	100	100	100	100
rules.JRip	248	100	97.91	98.17	98.17	97.92
rules.Ridor	248	99.93	100	99.87	99.87	100
rules.DecisionTable	248	99.93	100	99.87	99.87	100
trees.J48	248	100	99.086	98.56	98.57	99.08
Logistic	248	100	100	100	100	100
RBF	248	100	99.48	100	100	99.48
bayes.NaiveBayes	248	100	99.61	100	100	99.61
SVM	49	100	100	100	100	100
rules.JRip	49	98.04	97.91	98.17	98.17	97.92
rules.Ridor	49	98.30	98.43	98.17	98.18	98.43
rules.DecisionTable	49	96.93	97.65	96.21	96.27	97.62
trees.J48	49	98.82	99.09	98.56	98.57	99.08
Logistic	49	99.86	99.87	99.87	99.87	99.87
RBF	49	99.73	99.48	100	100	99.48
bayes.NaiveBayes	49	99.80	99.61	100	100	99.61
SVM	10	99.67	99.35	100	100	99.35
rules.JRip	10	98.62	98.43	98.83	98.82	98.44
rules.Ridor	10	98.49	98.43	98.83	98.82	98.44
rules.DecisionTable	10	97.78	98.17	97.39	97.41	98.16
trees.J48	10	98.95	99.22	98.69	98.70	99.21
Logistic	10	99.73	99.61	99.87	99.87	99.61
RBF	10	93.53	99.61	99.87	99.87	99.61
bayes.NaiveBayes	10	99.80	99.61	100	100	99.611
SVM	8	99.73	99.48	100	100	99.48
rules.JRip	8	98.10	97.78	98.43	98.42	97.79
rules.Ridor	8	99.80	99.74	99.87	99.87	99.74
rules.DecisionTable	8	97.78	98.17	97.39	97.41	98.16
trees.J48	8	98.95	99.22	98.69	98.70	99.21
Logistic	8	99.80	99.79	99.87	99.87	99.74
RBF	8	99.86	99.87	99.87	99.87	99.87
bayes.NaiveBayes	8	97.78	99.61	99.87	99.87	99.61

This experiment is performed in Weka, with the abundant choices of rules, decision trees, functions, and Bayes methods that it provides. The first section of the table shows the results of *SVM, JRip, Ridor, Decision Table, J48, Logistic, RBF*, and *Naïve-Bayes* algorithm on the original dataset with 248 features. The second section of the table shows the results of the classifiers with 49 features which are left after performing the attribute cleaning phase (The attributes with redundant values, so many Nan or minus infinite, or zero values have been removed. In this phase, the non-overlapping attributes for the two Persian music and LPM pieces were eliminated.) The third phase of attribute selection leaves us with ten features (ReliefF, Gain Ratio, Symmetrical uncertainty, and Information gain methods have been used for performing attribute selection). The last phase of attribute selection after performing CfSSubset evaluation leaves us with 8 features. Although the attributes have been shrinked by a significant amount, the different classifiers maintain their performance. Working with a small number of attributes saves processing time in the computations of the evolutionary algorithm.

Appendix G. First Survey Design and Results

In this appendix, various components in the design of the first auditory survey are presented. The design of the surveys is the subject of seventh chapter. Figure G-1 shows the introductory text appearing at the beginning of the survey.

LPM Survey

Liquid Persian Music Survey

Hello,

You are invited to complete the following survey in the fulfilment of a research project for generating Persian like music. This research is being conducted by Sahar Arshi at the Faculty of Engineering and Computer Science, University of Hull. The questionnaire includes the ranking of 9 audio pieces. Each piece lasts between 10 to 30 seconds. The whole test takes about 10 minutes to complete.

Your participation is voluntary, and you are free to withdraw at any time by simply closing the window without submission. This survey will be conducted in an anonymous manner. Any information regarding participants and the filled questionnaire will be saved in locked filings. The statistical results are intended to be published. However, the identifying information will not be included in the results. No levels of harm or inconvenience have been anticipated by doing the test.

If you have any questions concerning your participation in the survey please feel free to contact the organizers: Sahar Arshi [S.Arshi@2014.hull.ac.uk], or Doctor Darryl N. Davis [D.N.Davis@hull.ac.uk]. Or simply contact the Computer Science Office , University of Hull, Robert Blackburn, Cottingham Road, HU6 7RX, T. +44 (0)1482 465951 or +44 (0)1482 465067 , F. +44 (0)1482 466666, Email: dcs-ug@hull.ac.uk

By completing the questionnaire you are giving your consent of participation.

You can play the audio pieces in turn and select an answer for each of the three questions. Please adjust the volume of your system to a comforting level. I invite you to listen to https://www.youtube.com/watch?v=rPvLWeUlmZc (about two minute long) in relation to the third question (opens in separate tab).

Thank you very much in advance for your time and energy for rating the following audio.

Figure G-1.The introduction appearing on the first public survey page.

This text consists of a brief introduction of the research at hand and its purpose. It guides the audience to listen to a short sample of Persian Dastgāh music for those people who are not familiar with traditional Persian music. It also invites them to leave the survey if they do not want to commit to completing the survey. The audience are also suggested to contact the creators of the survey for further questions, assistance and the final results. The information provided on the welcoming page are in consistence with the ethical rules agreed apon with the University of Hull before conducting the survey.

Figure G-2 shows a sample question for evaluating an audio piece as they appear in the first public survey. This question is repeated for all the nine audio pieces present in the auditory survey. In the first LPM survey, nine audio clips are assessed according to their musicality, Persian Dastgāh-likeness, and whether the respondents liked the audio piece.

Figure G-2. A sample question for evaluating an audio piece as they appear in the first public survey.

In the first LPM survey nineaudio clips are evaluated against their musicality, Persian Dastgāh like-ness, and whether the respondents liked the audio piece.

Table G-1 demonstrates the result of the ratings of the audio clips for the first public survey. The left most column shows the evaluation criteria based on musicality, Persian-likeness, and the preferences of the participants. The ratings ranges between Totally disagree to strongly agree.

Table G-1.The result of the ratings of the audio clips for the first public survey.

The left most column presents the evaluation criteria on the basis of musicality, Persian-likeness, and the preferences of the participants. The ratings ranges between Totally disagree to strongly agree. The ratings are presented for nine audio pieces as appears in the remaining columns of the table. For each of the evaluation criteria, the total number of collected answeres together with the weighted average are presented.

		Audio Pieces								
Question	**Ratings**	**Piece 1**	**Piece 2**	**Piece 3**	**Piece 4**	**Piece 5**	**Piece 6**	**Piece 7**	**Piece 8**	**Piece 9**
1-How musical do you think the piece is?	Very Low (%)	42.42	32.35	23.53	31.43	23.53	29.41	11.76	31.43	22.86
	Low(%)	30.30	17.65	26.47	42.86	32.35	32.35	29.41	28.57	28.57
	Moderate(%)	18.18	38.24	38.24	14.29	26.47	26.47	38.24	25.71	31.43
	High(%)	6.06	8.82	8.82	8.57	11.76	8.82	17.65	11.43	14.29
	Very High(%)	3.03	2.94	2.94	2.86	5.88	2.94	2.94	2.86	2086
	Total Answers	33/35	34/35	34/35	35/35	34/35	34/35	34/35	35/35	35/35
	Weighted Average	1.97	2.32	2.41	2.09	2.44	2.24	2.71	2.26	2.46
2-How do you like this piece?	Very Low(%)	54.55	35.29	35.29	42.86	32.35	41.18	23.53	42.86	25.71
	Low(%)	27.27	26.47	23.53	37.14	38.24	29.41	32.35	31.43	34.29
	Moderate(%)	12.12	35.29	32.35	17.14	20.59	26.47	32.35	22.86	34.29
	High(%)	6.06	2.94	8.82	2.86	5.88	2.94	11.76	2.86	5.71
	Very High(%)	0.00	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00
	Total Answers	33/35	34/35	34/35	35/35	34/35	34/35	34/35	35/35	35/35
	Weighted Average	1.70	2.06	2.15	1.80	2.09	1.91	2.32	1.86	2.20
3-How do you like this piece in terms of Persian likeness?	Very Low(%)	24.24	20.59	20.59	28.57	29.41	32.35	12.12	22.86	17.14
	Low(%)	33.33	35.29	29.41	42.86	38.24	32.35	30.30	37.14	28.57
	Moderate(%)	30.30	35.29	38.24	22.86	23.53	29.41	45.45	31.43	45.71
	High(%)	12.12	8.82	11.76	5.71	5.88	5.88	12.12	8.57	8.57
	Very High(%)	0.00	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00
	Total Answers	33/35	34/35	34/35	35/35	34/35	34/35	33/35	35/35	35/35
	Weighted Average	2.30	2.32	2.41	2.06	2.15	2.09	2.58	2.26	2.46

Table G-2 shows the participants' comments in the first public survey.

Table G-2. The respondent's feedback in the first public conducted evaluation.

1	There is lack of fluency in the musical pieces. To be honest none of them are bad, but it reminds of sounds produced by a child who just started learning an instrument. For example, song n.2 (s6) starts with a weird set of accelerated notes and then suddenly drops the tempo. Maybe if you could set some limits to the highest frequency between notes.
2	Good quality recording. Musicality not to my ear.
3	Sounded very amateur, like a child let loose on a piano. Didn't find any of it at all musical at least not in the sense I appreciate.
4	Real music has more repetition of notes and themes which seems lacking here. Don't be afraid of repetition! The tonal qualities are mostly good but some of the lower notes are warped and sound wrong, like a badly tuned instrument. Also, don't be afraid to leave more space between notes/themes. Your produced sound needs to 'think in phrases' more. Most of these do sound santoor-like, but are lacking the musical structure a little. Well done for getting this far!
5	I did not like my music, I am sorry if it is a rude comment.
6	It all sounds terrible. I've got a degree in Music and I've heard Persian music a number of times. I don't know what I've just heard in those samples - if your goal is to develop an algorithmic representation of Persian-like music in order to generate simulated audio (that sounds like it), then there must be other fields of research in computer science/audio that would be more interesting.
7	The audio visuals have no rhythm to follow. All of them are rather dull but soothing to a certain extent. If there are rhythms, I think I would enjoy it. Now most of the audio visuals sound like a ghost movie background! Creepy.
8	I have absolutely no idea what Persian music sounds like
9	I don't think any of these pieces are in tune, but this could be due to my lack of knowledge in Persian music.
10	Ears struggled with music not at 440hz. I must confess I'm not familiar with Persian music - my judgement should be tempered.
11	would have been nice to mix in some nice Persian music, this started off OKish but got worse. Not a good advert for Persian music, but also meant I couldn't compare or contrast the music
12	I didn't like any of the sounds and I'm not very sure what Persian music actually sounds like.
13	I have never really heard Persian music before, but I have to say it might as well rank as one of the worst music I have ever listened too. The music sounds so random and their doesn't seem to be a focus in the music
14	I prefer 25 Band!!! Particularly Hamishe Ba Hamim :)
15	N/A
16	I have very little knowledge of this particular type of music, and based upon what I heard I can say I can get excited for it. However like any music I appreciate the talent and work that goes into the making.
17	I have very little idea of what 'Persian' music sounds like, so I wouldn't necessarily take my answers as scientific.
18	The production of the fragments is good, but it would be helpful to provide examples of what it meant by Persian likeness.
19	Moderate audio is the best.
20	It sounds like someones got a broken guitar AND has no idea how to play it anyway i think tje persians need to select some new riddems may i suggest "in the whip" by casual
21	I don't understand what music this was - I could not work out what instrument was being used? It was difficult because there were no chords, and I could not understand the key. I have played with musicians from other places, and I have not had this problem before. If your music is your way of communicating, or your way of praying, then I apologise for my ignorance. I hope that you will continue with your music. Perhaps you can sing also?

Appendix H. Second Survey Design and Results

In this appendix, the various components in the design of the second auditory survey are presented. The design of the survey is the subject of seventh chapter. Figure H-1 depicts the welcoming introduction appearing on the public survey page. The text consists of a brief introduction of the research at hand and its purpose. It guides the audience to listen to a short sample of Persian Dastgāh music. The information provided on the introduction page is in accordance with the ethical rules agreed upon with the University of Hull before conducting the survey.

Hello Everyone,

You are invited to complete the following survey in the fulfilment of a research project for generating Persian Dastgah-like music. This research is being conducted by Sahar Arshi at the Faculty of Engineering and Computer Science, University of Hull. The questionnaire includes the ranking of seven audio pieces which are generated by computer. Each piece lasts between 8 to 28 seconds. The survey will take around 10 minutes to complete.

Your participation is voluntary, and you are free to withdraw at any time by simply closing the window without submission. This survey will be conducted in an anonymous manner. Any information regarding participants and the filled questionnaire will be saved in locked files and be inaccessible to anyone other than the participant. The statistical results are intended to be published. However, the identifying information will not be included in the results. No levels of harm or inconvenience have been anticipated by doing the test.

If you have any questions concerning your participation in the survey please feel free to contact the organizers: Sahar Arshi [S.Arshi@2014.hull.ac.uk], or Doctor Darryl N. Davis [D.N.Davis@hull.ac.uk]. Or simply contact the Computer Science Office , University of Hull, Robert Blackburn, Cottingham Road, HU6 7RX, T. +44 (0)1482 465951 or +44 (0)1482 465067 , F. +44 (0)1482 466666

By completing the questionnaire you are giving your consent of participation.

You can play the audio pieces in turn and select an answer for each of the three questions. Please adjust the volume of your system to a comforting level. I invite you to listen to https://www.youtube.com/watch?v=95amB7kUu-k to provide you with a taste of Persian traditional music (opens in separate tab).

Thank you very much in advance for your time and energy for rating the following audio.

Figure H-1. The welcomming text appearing at the begginning of the second public survey.

This text consists of a brief introduction of the research at hand and its purpose. It guides the audience to listen to a short sample of Persian Dastgāh music for those people who are not familiar with traditional Persian music. It also invites them to leave the survey if they don’t want to commit to completing the survey. The audience are suggested to contact the creators of the survey for further questions, assistance and the final results.

Figure H-2 shows the question appearing at the beginning of the survey (both second and third surveys) to identify the possible biases about computational creativity, and computers improvising Dastgāh music. The list of bias inquiries are listed in table H-1.

1. To what extent would you agree with the following statements?

	Totally Disagree	Disagree	Neutral	Moderate	Agree	Strongly Agree
a) Computers can produce creative outputs.	○	○	○	○	○	○
b) Computers can occasionally or randomly be creative.	○	○	○	○	○	○
c) Computers cannot be creative because they merely reflect the creativity of programmer.	○	○	○	○	○	○
d) The idea of computers being creative disturbs me.	○	○	○	○	○	○
e) Computers might be or can be creative in the future but currently are not creative.	○	○	○	○	○	○
f) Computers will never be creative.	○	○	○	○	○	○
g) Computers can not generate Persian music.	○	○	○	○	○	○
h) Dastgāh Persian music should not be a subject for computational creativity.	○	○	○	○	○	○
i) I like the idea of computers being creative.	○	○	○	○	○	○
j) I do not like the idea of computers generating Dastgāh-	○	○	○	○	○	○

Figure H-2.The first question appearing in the survey is to identify the possible biases about computational creativity, and computers being able to improvise Dastgāh music.

Table H-2 depicts public responses to the second survey questions regarding the bias criteria. The rows are related to the queries as they appear in Table H-1. The value percentages under each column shows the people's ratings. The ranges of ratings vary between Totally Agree, Disagree, Neutral, Moderate, Agree, and Strongly Agree. These ratings have correspondingly weighted as: -2,-1, 0,1,2,3 in the order of their appearance. These weights are utilized for calculating the weighted average represented in the last column (These weights are different from the weights obtained for calculating Ritchie's evaluation criteria. The weights in Ritchie's criteria are obtained from the extrapolations of survey results).

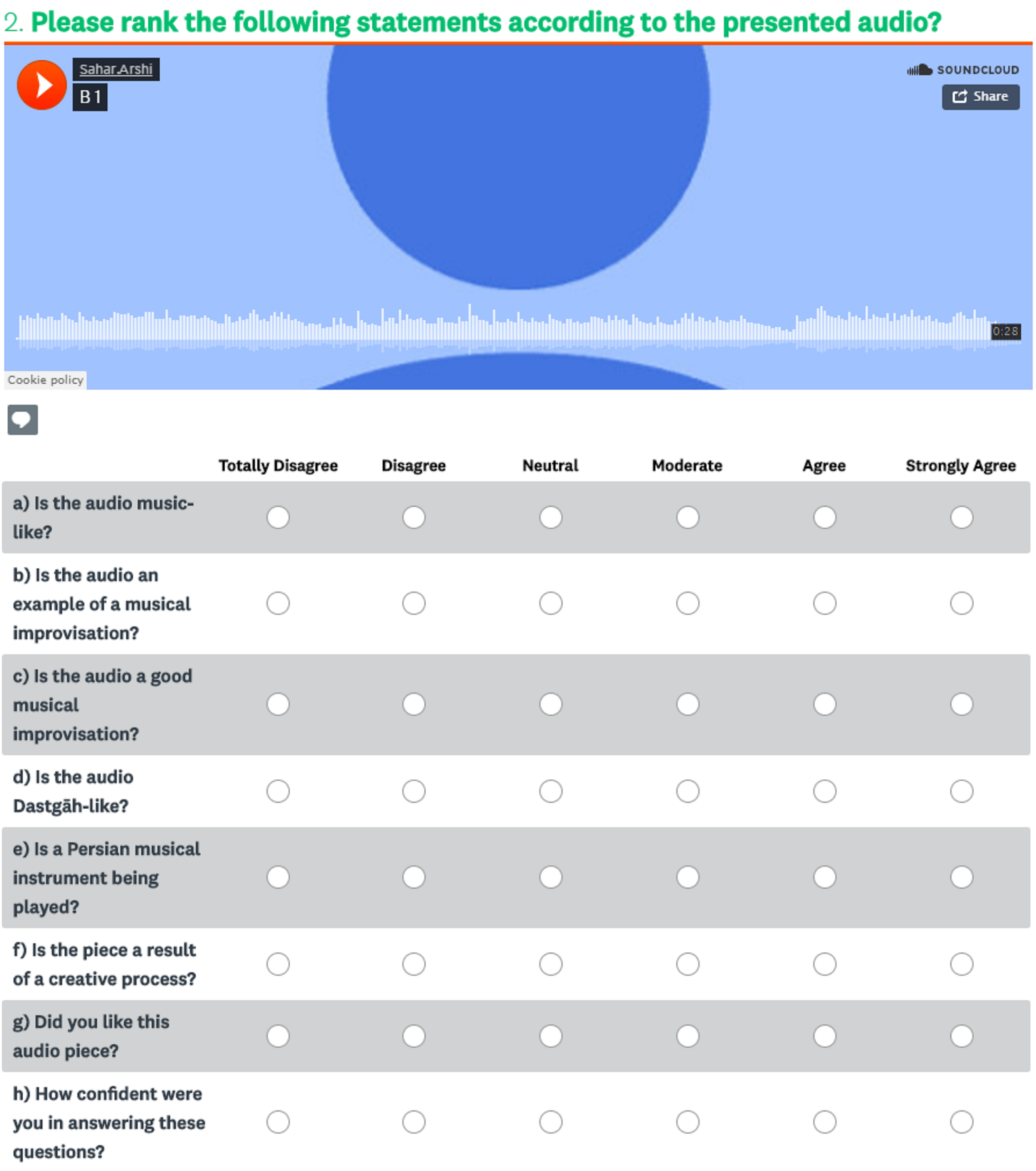
Table H-1.The possible bias criteria which asked about in the first question in the second and third surveys.

a	Computers can produce creative outputs.
b	Computers can occasionally or randomly be creative.
c	Computers cannot be creative because they merely reflect the creativity of programmer
d	The idea of computers being creative disturbs me.
e	Computers might be or can be creative in the future but currently are not creative.
f	Computers will never be creative.
g	Computers can not generate Persian music.
h	Dastgāh Persian music should not be a subject for computational creativity.
i	I like the idea of computers being creative.
j	I do not like the idea of computers generating Dastgāh-like music.

Table H-2.The public responses to the first survey questions regarding the bias criteria. The rows are associated to the subquestions as they appear in Table H-1. The value percentages under each column shows the proportion of people 's ratings.

	Totally Disagree (-2)	Disagree (-1)	Neutral (0)	Moderate (1)	Agree (2)	Strongly Agree (3)	Total	Weighted Average
a	5.66%	7.55%	20.75%	20.75%	24.53%	20.75%	53	1.13
b	7.55%	9.43%	20.75%	26.42%	16.98%	18.87%	53	0.92
c	7.55%	32.08%	7.55%	16.98%	24.53%	11.32%	53	0.53
d	13.21%	32.08%	26.42%	13.21%	7.55%	7.55%	53	-0.08
e	11.32%	18.87%	30.19%	22.64%	15.09%	1.89%	53	0.17
f	38.46%	32.69%	17.31%	3.85%	3.85%	3.85%	52	-0.87
g	28.30%	33.96%	24.53%	7.55%	1.89%	3.77%	53	-0.68
h	26.42%	41.51%	18.87%	5.66%	5.66%	1.89%	53	-0.072
i	11.32%	7.55%	20.75%	13.21%	28.30%	18.87%	53	0.96
j	28.30%	33.96%	26.42%	1.89%	1.89%	7.55%	53	-0.62

Figure H-3 shows a sample question for evaluating an audio piece as they appear in the second public survey. This question is repeated for all the 7 audio pieces present in the auditory survey. In the second LPM survey, audio clips are assessed according to criteria presented in table H-4.


2. **Please rank the following statements according to the presented audio?**

	Totally Disagree	Disagree	Neutral	Moderate	Agree	Strongly Agree
a) Is the audio music-like?						
b) Is the audio an example of a musical improvisation?						
c) Is the audio a good musical improvisation?						
d) Is the audio Dastgāh-like?						
e) Is a Persian musical instrument being played?						
f) Is the piece a result of a creative process?						
g) Did you like this audio piece?						
h) How confident were you in answering these questions?						

Figure H-3.The evaluation criteria for each of the audio pieces in the second public survey.

The subquestions a-h appear for each of the audio clips. There are 7 audio clips presented in this survey. The format of question two in the survey is replicated for all the questions up to the ninth question.

Table H-3 shows the ratings for the seven audio pieces according to the evaluation criteria in the second public survey.

Table H-3.The result of the ratings of the audio clips for the second public survey.

The left most column presents the evaluation criteria. The ratings ranges between Totally disagree to strongly agree. The ratings are presented for seven audio pieces as appears in the remaining columns of the table. For each of the evaluation criteria a-h, the total number of collected answeres together with the weighted average are presented.

Questions	Ratings	Piece 1	Piece 2	Piece 3	Piece 4	Piece 5	Piece 6	Piece 7
1- Is the audio music-like	Totally Disagree (%)	5.88	7.84	7.84	9.80	7.84	9.80	11.76
	Disagree (%)	17.65	9.80	7.84	19.61	11.76	17.65	19.61
	Neutral (%)	11.76	13.73	11.76	15.69	15.69	11.76	7.84
	Moderate (%)	33.33	19.61	29.41	29.41	33.33	29.41	27.45
	Agree (%)	21.57	39.22	33.33	21.57	25.49	25.49	27.45
	Strongly Agree (%)	9.80	9.80	9.80	3.92	5.88	5.88	5.88
	Total Answers	51	51	51	51	51	51	51
	Weighted Average	0.76	1.02	1.02	0.45	0.75	0.61	0.57
2- Is the audio an example of a musical improvisation ?	Totally Disagree (%)	5.77	5.77	3.85	11.54	5.77	9.62	7.69
	Disagree (%)	11.54	19.23	9.62	15.38	17.31	17.31	26.92
	Neutral (%)	23.08	23.08	26.92	25.00	23.08	25.00	17.31
	Moderate (%)	28.85	21.15	26.92	21.15	30.77	28.85	17.31
	Agree (%)	25.00	26.92	26.92	23.08	19.23	15.38	26.92
	Strongly Agree (%)	5.77	3.85	5.77	3.85	3.85	3.85	3.85
	Total Answers	52	52	52	52	52	52	52
	Weighted Average	0.73	0.56	0.81	0.40	0.52	0.35	0.40
3- Is the audio a good musical improvisation ?	Totally Disagree (%)	9.62	13.46	15.38	15.38	13.46	17.65	19.23
	Disagree (%)	44.23	28.85	21.15	34.62	21.15	33.33	23.08
	Neutral (%)	21.15	25.00	26.92	19.23	32.69	21.57	28.85
	Moderate (%)	11.54	19.23	17.31	19.23	25.00	17.65	19.23
	Agree (%)	7.69	9.62	13.46	9.62	3.85	5.88	5.77
	Strongly Agree (%)	5.77	3.85	5.77	1.92	3.85	3.92	3.85
	Total Answers	52	52	52	52	52	51	52
	Weighted Average	-0.19	-0.06	0.10	-0.21	-0.04	-0.27	-0.19
4- Is the audio Dastgāh-like?	Totally Disagree (%)	1.92	13.46	7.69	7.69	7.69	7.69	9.62
	Disagree (%)	11.54	42.31	7.69	26.92	13.46	23.08	23.08
	Neutral (%)	34.62	38.46	34.62	32.69	38.46	28.85	34.62
	Moderate (%)	32.69	3.85	19.23	23.08	21.15	25.00	17.31
	Agree (%)	15.38	1.92	23.08	7.69	15.38	13.46	13.46
	Strongly Agree (%)	3.85	0.00	7.69	1.92	3.85	1.92	1.92

	Total Answers	52	52	52	52	52	52	52
	Weighted Average	0.60	-0.62	0.65	0.02	0.35	0.19	0.08
5- Is a Persian musical instrument being played?	Totally Disagree (%)	9.62	21.15	7.69	11.54	11.54	15.38	9.62
	Disagree (%)	15.38	38.46	11.54	21.15	13.46	13.46	26.92
	Neutral (%)	23.08	30.77	17.31	26.92	32.69	30.77	28.85
	Moderate (%)	11.54	5.77	30.77	17.31	25.00	25.00	19.23
	Agree (%)	32.69	3.85	25.00	21.15	13.46	11.54	13.46
	Strongly Agree (%)	7.69	0.00	7.69	1.92	3.85	3.85	1.92
	Total Answers	52	52	52	52	52	52	52
	Weighted Average	-0.65	-0.67	0.77	0.21	0.27	0.15	0.06
6- Is the piece a result of a creative process?	Totally Disagree (%)	3.85	7.69	3.85	7.69	5.88	7.69	9.62
	Disagree (%)	17.31	13.46	7.69	15.38	13.73	13.46	13.46
	Neutral (%)	26.92	34.62	28.85	30.77	31.37	30.77	25.00
	Moderate (%)	25.00	23.08	23.08	19.23	23.53	32.69	28.85
	Agree (%)	21.15	15.38	30.77	23.08	21.57	11.54	19.23
	Strongly Agree (%)	5.77	5.77	5.77	3.85	3.92	3.85	3.85
	Total Answers	52	52	52	52	51	52	52
	Weighted Average	0.60	0.42	0.87	0.46	0.53	0.38	0.46
7- Did you like this audio piece?	Totally Disagree (%)	20.00	28.85	11.76	32.69	26.92	21.15	21.15
	Disagree (%)	44.00	21.15	33.33	28.85	25.00	36.54	40.38
	Neutral (%)	8.00	19.23	13.73	5.77	19.23	13.46	11.54
	Moderate (%)	14.00	15.38	23.53	15.38	13.46	13.46	17.31
	Agree (%)	10.00	7.69	11.76	13.46	7.69	11.54	5.77
	Strongly Agree (%)	4.00	7.69	5.88	3.85	7.69	3.85	3.85
	Total Answers	50	52	51	52	52	52	52
	Weighted Average	-0.38	-0.25	0.08	-0.40	-0.27	-0.31	-0.42
8- How confident were you in answering these questions?	Totally Disagree (%)	0.00	0.00	1.92	1.92	3.85	3.85	3.85
	Disagree (%)	11.54	15.38	9.62	7.69	9.62	7.69	9.62
	Neutral (%)	19.23	17.31	21.15	17.31	17.31	17.31	17.31
	Moderate (%)	19.23	17.31	25.00	21.15	26.92	26.92	25.00
	Agree (%)	25.00	21.15	21.15	25.00	21.15	19.23	19.23
	Strongly Agree (%)	25.00	28.85	21.15	26.92	21.15	25.00	25.00
	Total Answers	52	52	52	52	52	52	52
	Weighted Average	1.33	1.31	1.17	1.40	1.15	1.25	1.21

Table H-5 illustrates the the scorings associated to all the seven audio pieces over the evaluation criteria in the second public survey presented in table H-4.

Table H-4. The queries asked about the generated audio in the second auditory survey.

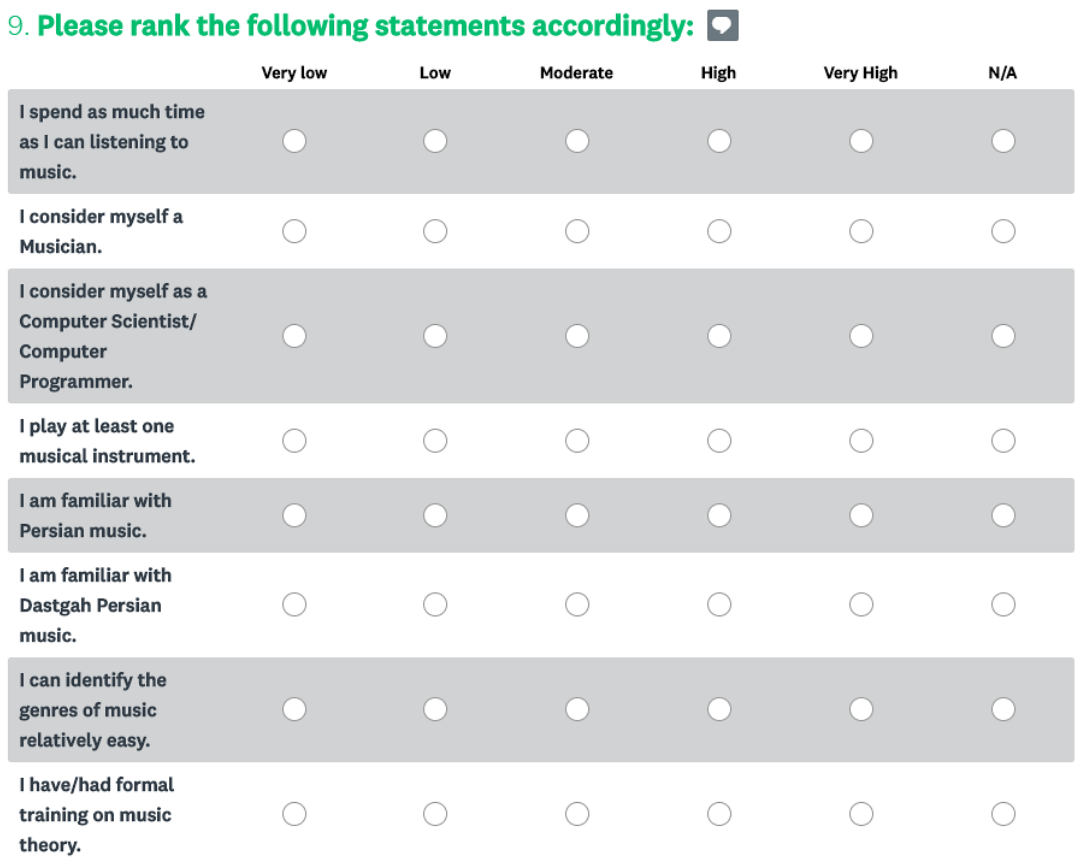
a	Is the audio music-like?
b	Is the audio an example of a musical improvisation?
c	Is the audio a good musical improvisation?
d	Is the audio Dastgāh-like?
e	Is a Persian musical instrument being played?
f	Is the piece a result of a creative process?
g	Did you like this audio piece?
h	How much confident were you in answering these questions?

Table H-5.The ratings for all the seven audio pieces over the evaluation criteria in the second public survey.

The percentages are recalculated according to the responses for all the seven audio clips rather than individually as they were presented in table H-4.

	Totally Disagree	Disagree	Neutral	Moderate	Agree	Strongly Agree	Total	Weighted Average
a	8.683%	14.84%	12.60%	28.85%	27.73%	7.28%	357	73.94
b	7.14%	16.75%	23.35%	25%	23.35%	4.39%	364	53.84
c	6.61%	29.47%	25.06%	18.45%	7.98%	4.13%	363	4.13
d	7.96%	21.15%	34.61%	20.87%	12.91%	3.02%	364	18.68
e	12.36%	20.05%	27.19%	19.23%	17.30%	3.84%	364	20.60
f	6.61%	13.49%	29.75%	25.06%	20.38%	4.68%	363	53.16
g	23.26%	32.68%	13.01%	16.06%	9.69%	5.26%	361	-27.97
h	2.19%	10.16%	18.13%	23.07%	21.7%	24.45	364	12.52

Figure H-4 shows question nine as it appears in the survey. Its purpose is to investigate the familiarity of the respondents with music, Persian music, computer programming, and the participants background in the related area.



9. **Please rank the following statements accordingly:**

	Very low	Low	Moderate	High	Very High	N/A
I spend as much time as I can listening to music.						
I consider myself a Musician.						
I consider myself as a Computer Scientist/ Computer Programmer.						
I play at least one musical instrument.						
I am familiar with Persian music.						
I am familiar with Dastgah Persian music.						
I can identify the genres of music relatively easy.						
I have/had formal training on music theory.						

Figure H-4.Question nine in the second public survey.

This question asks for the background and familiarity of respondents about music and computer science.

Table H-6 illustrates the sub queries presented in the ninth question of the second public survey. The collected rating for the sub queries in question nine are presented in table H-7.

Table H-6. The subqueries presented in the ninth question of the second public survey.

a	I spend as much time as I can listening to music.
b	I consider myself a Musician.
c	I consider myself as a Computer Scientist/ Computer Programmer.
d	I play at least one musical instrument.
e	I am familiar with Persian music.
f	I am familiar with Dastgāh Persian music.
g	I can identify the genres of music relatively easy.
h	I have/had formal training on music theory.

Table H-7. The results of the queries about the familiarity of respondents with music and computer science.

	Very Low	Low	Moderate	High	Very High	N/A	Total	Weighted Average
a	3.77%	13.21%	24.53%	26.42%	32.08%	0.00%	53	1.53
b	33.96%	22.64%	18.87%	3.77%	16.98%	3.77%	53	-0.14
c	43.40%	18.87%	11.32%	11.32%	11.32%	3.77%	53	-0.39
d	22.64%	20.75%	11.32%	13.21%	30.19%	1.89%	53	0.63
e	33.96%	20.75%	20.75%	11.32%	11.32%	1.89%	53	-0.12
f	41.51%	28.30%	16.98%	7.55%	3.77%	1.89%	53	-0.69
g	5.66%	18.87%	41.51%	22.64%	11.32%	0.00%	53	0.91
h	33.96%	24.53%	7.55%	11.32%	16.98%	5.66%	53	-0.12

Figure H-5 shows the tenth question as it appears in the second survey. The participants' comments are presented in table H-8.

Figure H-5. The last question in the second public survey was dedicated to collecting comments from audience.

Table H-8. The respondent's feedbacks in the second public survey conducted.

1	All the audio had the same thing in common- it didn't have a set time-signature, which made it uncomfortable to listen to and very different from almost all other types of music. Even the randomness of the notes might have been interesting if there was some semblance of rhythm. I'm not sure of some of these are computer generated and some done on real instruments- I didn't listen carefully enough- if that had been explicitly asked, I think I could have told the difference.
2	I didn't feel I could comment on whether or not the musical process was creative with little background on what process was being used. If it was totally random code then I would have said no, but if the computer was programmed to make a choice based on a preference then I would have agreed it was creative.
3	Much of the material very much felt computer generated. What I mean by that is it lacked the microscopic blemishes (in terms of time, pitch, attack etc.) you'd expect from any human performer. With the one exception being the track on prepared piano, I was unable to determine whether this was computer generated or actually performed. Good luck with the research!
4	the audio from later part are horrible. I don't think they can be categorised as music
5	Very interesting audio clips, I enjoyed the thought process of thinking through what is 'music' and what can be classed as a genre. I must admit that I go to the Youtube clip a few times to orient myself.
6	Very interesting experience to listen to. I think the timbres were often quite convincing but it definitely lacked a sense of mode in both repetitive shaping (formulas) and chromatic shading (microtones). I think you might consider using a word other than improvisation - it isn't possible to tell if something is improvised by listening to it! you need further - Information and a definition of what you mean by improvisation (this is part of my phd topic, sorry!)
7	I was very surprised with the rhythms because I find them surprisingly credible and nice. I immediately imagined Bossa Nova songs with such rhythms. There is something strange to my ears, though... there are very rich rhythms and the notes variations generally as well... I asked myself is there any way of extracting the rhythms and some notes and building on top of them with other genres? (on a poetical note, I would comment with this song https://www.youtube.com/watch?v=tCMhuN3053o) (on the musical spirit I felt on the rhythms, I would go with something like this: https://www.youtube.com/watch?v=XjJYeCYO-hA)
8	The audios do sound music-like but in a very computerised way. If I heard the audios with no context then I would still come to the conclusion that they were computer generated. In my opinion 'creativity' is subjective; therefore, what I may deem to be creative for a computer may not be what I would consider to be creative for a person.

9	The improvised nature of it meant that it lacked rhythm and melody and therefore failed to be convincingly ‘musical’.
10	sorry, I did not understand the audio as music, probably due to my ignorance regarding Persian / Dastgāh Persian music. To me it sounded like random playing of the same 3-4 notes (with one dominant one), without musicality to my ears
11	I enjoyed the process. Identifying non-Western musical sounds takes a while to become accustomed to, so I feel that my answers where impacted as the survey developed for me.
12	You have done your best and is appreciated and you have lot of jobs to do,and you are not far from the final result. I wish you all the best.
13	I admire your perservance and creativity.
14	In my opinion all of the audios are very creative and unique.
15	Well done ????
16	After listening to the first Youtube clip provided, which I really enjoyed listening to, moving to these clips seemed coming from a computer generated sound (an electronic version of the instrument). Being a musician I was very aware of the difference in sound, which is why I was a little hesitant at the question on whether a Persian instrument was played in the questionnaire clips. At times this was better than others though. Also, I am aware computers are able to generate sounds "in the style of" a certain music genre or composer, however the human creative process will still lack in these. This was a thought when listening to those clips, although some were better than others.
17	This was quite painful to listen to. There wasn't even one piece that I liked.
18	Some of them were very nice to listen to and although they weren't very familiar to me, but I think they were very creative and I even played them twice.
19	None
20	It all sounded like very basic computer generated music. It lacked a sense of cohesion, and was too messy to pick up on a melody. I felt there was too much dissonance and uncoordinated dissonance. It felt as if the notes were randomly selected by a computer. Could do with an algorithm to match up notes across different instruments, enabling more harmonic control.
21	https://www.youtube.com/watch?v=SacogDL_4JU https://www.youtube.com/watch?v=uiJAy1jDIQ0 https://www.youtube.com/watch?v=nA3YOFUCn4U
22	This kind of music is not my favourite. However, I'm Persian. Good luck Sahar
23	A lot of it sounded like when young children first play about keyboards
24	It is clearly produced using some form of digitised production technique. If we are studying the compositional process of intelligence it may be worthwhile to transfer these across to the real instruments. This would increase our ability to determine an opinion based solely on the composition. The production of the audio is therefore too unreflective of the acoustic sound of such zither instrumentation that it automatically adds connotations to our listening perceptions. The timing of the zither parts seems random, almost like it is electrical signals literally transferred into zither sound. This is unreflective if the inmates rhythmic tendencies of Persian and Dastgāh music, which tend to allay beats using subtle changes of stresses, but still retain the sense of tempo. The music also is not reflective of the realistic pitched range of the ‘instrument’ used. This could be a point of development for random computer generated music to be specified and limited by human intuition and/or musical expertise.
25	I am ignorant of Dastgāh Persian music. I have listened to the beautiful piece given in the link prior to the survey and I was captivated by the sense of musical narrative I felt whilst listening. About the samples, I have felt that it sounded like it, but because the samples where short, I didn't feel the sensorial story/narrative that I have felt with the Dastgāh player. Maybe I would need a full piece to get it. I also ask myself if a computer could play such narratives "ad eternum"? (instead of limited to one musical piece?). In some of the samples, I have also had a feeling of hearing fast and varied Jazz rhythms.

26	Having listened to the initial sample from the YouTube clip, I found the audio to sound like it had been created using a computer rather than a 'real life' sound od a Persian Instrument. The youtube clip was much more enjoyable to listen to that the audio clips in the survey - they didn't seem to have any structure or flow to them. I also didn't understand question 'a' in the survey which is why I didn't answer.
27	Some parts of the audio seemed unnatural and more randomised than created patterns
28	I focused too much on what I perceived as discontinuities in the music so I was convinced all samples were created artificially. I am not a musician so my appreciation is limited.
29	It seems that there is a far way to go for computers being creative. I'm happy to hear from this type of research and wish you a progressive results. Good luck
30	Not very experienced with Persian music or Dastgāh music but most of the samples sounded plausibly realistic, well done !
31	Not really my cup of tea ;)
32	I felt that all of the tracks sounded tinny and were not to my liking. In my opinion all of them sounded like a computer rather than an instrument. Sorry!

Appendix I. Third Survey Design and Results

In this appendix the associated various components in the design of the third auditory survey are presented. The design of the surveys are the subject of seventh chapter. Figure I-1 depicts the welcomming introduction appearing on the public survey page. The text consists of a brief introduction of the research at hand and its purpose. The information provided on the introduction page are in accordance with the ethical rules agreed apon with the University of Hull before conducting the survey.

Hello Everyone,

You are invited to complete the following survey in the fulfilment of a research project for generating Persian like music. This research is being conducted by Sahar Arshi at the Faculty of Engineering and Computer Science, University of Hull. The questionnaire includes the ranking of eight audio pieces which are generated by computer. Each piece lasts between 8 to 28 seconds. The survey will take around 10 to 15 minutes to complete.

Your participation is voluntary, and you are free to withdraw at any time by simply closing the window without submission. This survey will be conducted in an anonymous manner. Any information regarding participants and the filled questionnaire will be saved in locked files and be inaccessible to anyone other than the participant. The statistical results are intended to be published. However, the identifying information will not be included in the results. No levels of harm or inconvenience have been anticipated by doing the test.

If you have any questions concerning your participation in the survey please feel free to contact the organizers: Sahar Arshi [S.Arshi@2014.hull.ac.uk], or Doctor Darryl N. Davis [D.N.Davis@hull.ac.uk]. Or simply contact the Computer Science Office , University of Hull, Robert Blackburn, Cottingham Road, HU6 7RX, T. +44 (0)1482 465951 or +44 (0)1482 465067 , F. +44 (0)1482 466666.

By completing the questionnaire you are giving your consent of participation.

You can play the audio pieces in turn and select an answer for each of the three questions. Please adjust the volume of your system to a comforting level.

Thank you very much in advance for your time and energy for rating the following audio.

Figure I-1.The introduction appearing on the survey page for professionals.

This text consists of a brief introduction of the research at hand and its purpose. It lets the respondents to leave the survey if they do’nt want to commit to completing the survey. The audience are suggested to contact the creators of the survey for further questions, assistance and asking about the final results. The information provided on the welcoming page are in consistence with the ethical rules agreed apon with the University of Hull before conducting the survey.

Table I-2 depicts the professional responses to the third survey sub queries regarding the bias criteria. The rows are related to the queries as they appear in Table I-1. The value percentages under each column show the people's ratings. The ranges of ratings vary between Totally Agree, Disagree, Neutral, Moderate, Agree, and strongly Agree. These ratings have correspondingly weighted as -2,-1, 0, 1, 2, 3 in the order of their appearance. These weights are utilized for calculating the weighted average represented in the last column.

Table I-1. The possible bias criteria which asked about in the first question in the third survey.

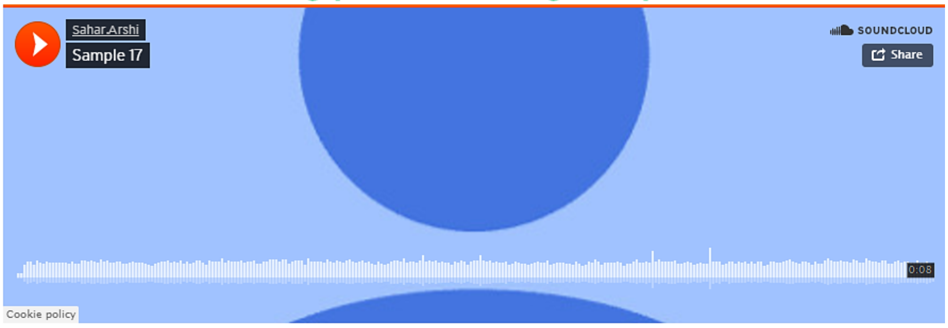
a	Computers can produce creative outputs.
b	Computers can occasionally or randomly be creative.
c	Computers cannot be creative because they merely reflect the creativity of programmer
d	The idea of computers being creative disturbs me.
e	Computers might be or can be creative in the future but currently are not creative.
f	Computers will never be creative.
g	Computers can not generate Persian music.
h	Dastgāh Persian music should not be a subject for computational creativity.
i	I like the idea of computers being creative.
j	I do not like the idea of computers generating Dastgāh-like music.

Table I-2. The professional responses to the first question regarding the bias criteria.

The rows are associated to the subquestions as they appear in Table I-1. The value percentages under each column shows the proportion of people 's ratings. The ranges of ratings vary between Totally Agree, Disagree, Neutral, Moderate, Agree, and strongly Agree.

	Totally Disagree (-2)	Disagree (-1)	Neutral (0)	Moderate (1)	Agree (2)	Strongly Agree (3)	Total	Weighted Average
a	28.57%	28.57%	0.00%	42.86%	0.00%	0.00%	7	-0.43
b	42.86%	28.57%	0.00%	28.57%	0.00%	0.00%	7	-0.86
c	14.29%	28.57%	0.00%	0.00%	14.29%	42.86%	7	1.00
d	42.86%	28.57%	14.29%	0.00%	0.00%	14.29%	7	-0.71
e	28.57%	0.00%	0.00%	42.86%	28.57%	0.00%	7	0.43
f	14.29%	42.86%	14.29%	0.00%	14.29%	14.29%	7	0.00
g	0.00%	57.14%	0.00%	14.29%	14.29%	14.29%	7	0.29
h	28.57%	57.14%	0.00%	0.00%	0.00%	14.29%	7	-0.71
i	14.29%	28.57%	14.29%	28.57%	14.29%	0.00%	7	0.00
j	14.29%	71.43%	0.00%	0.00%	0.00%	14.29%	7	-0.57

3. **Please rank the following questions according to the presented audio:**



	Very Low	Low	Neutral	Moderate	High	Very High
a) Is the audio music-like?	○	○	○	○	○	○
b) Would you agree that this is an example of musical improvisation?	○	○	○	○	○	○
c) Would you agree that this is a good musical improvisation?	○	○	○	○	○	○
d) To what extent is the generated audio an example of the Dastgāh Persian music genre?	○	○	○	○	○	○
e) To what extent is the produced audio a high quality example of Dastgāh?	○	○	○	○	○	○
f) Do you like this audio piece?	○	○	○	○	○	○
g) To what extent would you agree the piece was a result of a creative process?	○	○	○	○	○	○
h) Does the audio sound as if a Persian musical instrument is being played?	○	○	○	○	○	○
i) How confident were you in answering these questions?	○	○	○	○	○	○

Figure I-2. The evaluation criteria for each of the audio pieces in the third survey.

The subquestions a-i appear for each of the audio clips. There are eight audio clips presented in this survey. The format of question two in the survey is replicated for all the auditory questions.

This question is repeated for all the 8 audio pieces presented in the auditory survey.

Table I-3 shows the ratings for the eight audio pieces.

Table I-3. The result of the ratings of the audio clips for the survey from professionals.

The left most column presents the evaluation criteria. The ratings ranges between Totally disagree to strongly agree. The ratings are presented for eight audio pieces as appears in the remaining columns of the table. For each of the evaluation criteria a-i, the total number of collected answeres together with the weighted average are presented.

Question	Ratings	Piece 1	Piece 2	Piece 3	Piece 4	Piece 5	Piece 6	Piece 7	Piece 8
a)Is the audio music-like?	Totally Disagree (%)	14.29	0.00	28.57	28.57	57.14	57.14	57.14	71.43
	Disagree (%)	57.14	0.00	28.57	71.43	14.29	28.57	28.57	28.57
	Neutral (%)	0.00	28.57	28.57	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	28.57	57.14	14.29	0.00	14.29	14.29	14.29	0.00
	Agree (%)	0.00	14.29	0.00	0.00	14.29	0.00	0.00	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-0.57	0.86	-0.71	-1.29	-0.86	-1.29	-1.29	-1.71
b) Would you agree that this is an example of musical improvisation?	Totally Disagree (%)	57.14	28.57	42.86	71.43	85.71	57.14	57.14	85.71
	Disagree (%)	14.29	57.14	28.57	28.57	14.29	42.86	42.86	14.29
	Neutral (%)	14.29	0.00	28.57	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	14.29	14.29	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-1.14	-1	-1.14	-1.71	-1.86	-1.57	-1.57	-1.86
c) Would you agree that this is a good musical improvisation?	Totally Disagree (%)	85.71	57.14	57.14	57.14	85.71	71.43	85.71	85.71
	Disagree (%)	0.00	28.57	28.57	28.57	0.00	14.29	0.00	0.00
	Neutral (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Strongly Agree (%)	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-1.29	-1	-1	-1	-1.29	-1.14	-1.29	-1.29
d) To what extent is the generated audio an example	Totally Disagree (%)	85.71	71.43	42.86	71.43	100.00	85.71	85.71	100.00
	Disagree (%)	0.00	28.57	57.14	28.57	0.00	14.29	14.29	0.00
	Neutral (%)	14.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7

of the Dastgāh Persian music genre?	Weighted Average	-1.71	-1.71	-1.43	-1.71	-2	-1.86	-1.86	-2
e) To what extent is the produced audio a high quality example of Dastgāh?	Totally Disagree (%)	71.43	85.71	71.43	71.43	100.00	100	85.71	85.71
	Disagree (%)	14.29	14.29	28.57	28.57	0.00	0.00	14.29	14.29
	Neutral (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	14.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-1.43	-1.86	-1.71	-1.71	-2	-2	-1.86	-1.86
f) Do you like this audio piece?	Totally Disagree (%)	57.14	42.86	57.14	85.71	71.43	57.14	100.00	85.71
	Disagree (%)	42.86	28.57	28.57	14.29	14.29	28.57	0.00	14.29
	Neutral (%)	0.00	28.57	14.29	0.00	14.29	14.29	0.00	0.00
	Moderate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-1.57	-1.14	-1.43	-1.86	-1.57	-1.43	-2	-1.86
g) To what extent would you agree the piece was a result of a creative process?	Totally Disagree (%)	28.57	42.86	28.57	42.86	57.14	71.43	71.43	71.43
	Disagree (%)	42.86	42.86	57.14	28.57	28.57	14.29	14.29	14.29
	Neutral (%)	14.29	0.00	0.00	14.29	0.00	0.00	0.00	0.00
	Moderate (%)	14.29	14.29	0.00	14.29	14.29	14.29	0.00	14.29
	Agree (%)	0.00	0.00	14.29	0.00	0.00	0.00	14.29	0.00
	Strongly Agree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	-0.86	-1.14	-0.86	-1	-1.29	-1.43	-1.29	-1.43
h) Does the audio sound as if a Persian musical instrument is being played?	Totally Disagree (%)	0.00	57.14	0.00	14.29	14.29	14.29	57.14	42.86
	Disagree (%)	42.86	28.57	28.57	71.43	57.14	57.14	28.57	42.86
	Neutral (%)	0.00	14.29	14.29	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	42.86	0.00	28.57	0.00	14.29	14.29	14.29	14.29
	Agree (%)	14.29	0.00	14.29	14.29	14.29	14.29	0.00	0.00
	Strongly Agree (%)	0.00	0.00	14.29	0.00	0.00	0.00	0.00	0.00
	Total Answers	7	7	7	7	7	7	7	7
	Weighted Average	0.29	-1.43	0.71	-0.71	-0.43	-0.43	-1.29	-1.14
i) How much	Totally Disagree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Disagree (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

confident were you in answering these questions?	Neutral (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Moderate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Agree (%)	100	71.43	57.14	57.14	66.67	71.43	42.86	42.86
	Strongly Agree (%)	0.00	28.57	42.86	42.86	33.33	28.57	57.14	57.14
	Total Answers	7	7	7	7	6	7	7	7
	Weighted Average	2	2.29	2.43	2.43	2.33	2.29	2.57	2.57

Table I-5 illustrates the scorings associated to all the 8 audio pieces over the evaluation criteria in the third survey (presented in table I-4).

Table I-4. The queries asked about the generated audio in the third survey.

a	Is the audio music-like?
b	Would you agree that this is an example of musical improvisation?
c	Would you agree that this is a good musical improvisation?
d	To what extent is the generated audio an example of the Dastgāh Persian music genre?
e	To what extent is the produced audio a high quality example of Dastgāh?
f	Do you like this audio piece?
g	To what extent would you agree the piece was a result of a creative process?
h	Does the audio sound as if a Persian musical instrument is being played?
i	How much confident were you in answering these questions?

Table I-5. The ratings for all the seven audio pieces over the evaluation criteria in the third survey.

The percentages are recalculated according to the responses for all the eight audio clips rather than individually as they were presented in table I-4.

	Totally Disagree	Disagree	Neutral	Moderate	Agree	Strongly Agree	Total	Weighted Average
a	39.28%	32.14%	7.14%	16.07%	3.57%	0%	56	-0.87
b	60.71%	30.35%	5.35%	3.57%	0%	0%	56	-1.48
c	73.21%	12.5%	0%	0%	0%	14.28%	56	-1.16
d	80.35%	17.85%	1.78%	0%	0%	0%	56	-1.78
E	83.92%	10.71%	1.78%	1.78%	0%	0%	56	-1.76
f	69.64%	21.42%	8.92%	0%	0%	0%	56	-1.60
g	51.78%	30.35%	3.57%	10.71%	3.67%	0%	56	-1.16
h	25%	44.64%	3.57%	16.07%	8.92%	1.78%	56	-0.55
i	0%	0%	0%	0%	63.63%	36.36%	55	2.36

The participants' comments are presented in table I-6.

Table I-6. The comments provided by the participants in the third survey.

1	Hello and I am grateful from you. Many of these pieces seem to be the result of a random process, rather than innovative improvisation. (I wish you would have explained more about the creative process and its difference with a random process especially for the case of improvisation of traditional music) For me the thing that imposed gaps between the generated pieces and Persian Iranian music was the lack of modes and melodies based on the Dastgāh and vocal models of Persian music. The generated audio enlivened more of Persian Santur's tuning space. Of course in some of the cases the tonality quality represented the sound of Persian Santur. However, the lack of the formation of musical modes did not guide me through the space of improvisation of traditional music. In the end, I wish to thank you for your hard work. Please accept my apologies if you find this comment disturbing. I felt my truthful opinion will be more helpful in the process of your research. Thank you very much
2	Research in any field for music is worthy of praise appreciable. Good luck with your research
3	In the discussion of improvisation in Persian music, the mere performance of a random melody cannot be taken into account as an improvisation. Improvisation has its own coordinates and is based on a very strong infrastructure, including the full range of complete Iranian Radif, and Dastgāh musical pieces including all the folk songs and vocals.
4	Creativity in the music is nascent from human brain. Creativity is not acquired. What is acquired is experience that becomes a tool for manifesting creativity and flourishes the creativity itself. Similarly, it can be concluded that the transfer of creativity to the machine is not possible. At least not at the moment. You can work on the quality of the sounds at the moment to be better propagated. Work on music sounds to have better quality. Also, work on the composition of instruments. Especially in the performance of monotone Santur tones a sound of a second instrument could be heard which was not pleasant. All in all, it was an interesting research work and has the potential for further work. Good luck
5	Thanks to you, starting of any new work and paths requires long journeys in order to achieve the desired results, and certainly with effort and perseverance in the work you will achieve better results.

Appendix J. Evaluation of the Ritchie's Criteria for the Second Survey

The typicality and value questions in the second survey are:

- Is the audio an example of a musical improvisation (Typicality)?
- Is the audio a good musical improvisation (Value)?

The guidelines for implementing Ritchie's evaluation criteria are taken from (Jordanous, 2012b; Pereira et al., 2001; Ritchie, 2007). Jordanous provided a comprehensive manual for applying various evaluation methodologies. The author of this thesis found the evaluation examples in (Jordanous, 2012b) quite helpful to follow. In the following, the symbols for the Ritchie's criteria formulations and the choices of parameters for the second survey are presented:

- R: stands for the result set or outputs of the system.
- I: stands for the inspiration set and refers to the items, which were presented originally to the system. For instance, the items in I set are the items in the training set. The inspiring set in the case of this project are the Persian music Dastgāh dataset that were formerly introduced in chapter 5. None of the members of R are members of the inspiring set. Therefore, we have $R - I = R$, and $R \cap I = R$.
- typ (Typicality): The audio pieces were evaluated according to their typicality (whether they are examples of musical improvisation). The weights in the likert are normalized between 0.2 and 1. The Ritchie's suggestion for the evaluation weights were in the range of 0-1. This leaves the weights as: strongly disagree (0.2), disagree (0.36), neutral (0.52), moderate (0.68), agree (0.84), and strongly agree (1).

- val (Value): The value ratings were obtained for each of the audio samples in the survey. The weights in the likert were selected similar to the case of typ.

- $T_{\alpha,\beta}(R)$ stands for the subset of R which meets the suitability levels of typicality. Suitable ranges of typicality were given to items that were rated as moderate, agree, and strongly agree. The items that were rated higher than neutral were considered to have reached acceptable levels of typicality. The neutral weight in this case is 0.52.

  $T_{\alpha,1}(R)$,$\alpha = 0.52$ the mean typicality of audio samples are compared with α. If the mean typicality of an audio sample is greater than α, it is said that it has reached the ranges of suitable typicality. The audio samples that meet this specification are tagged with T label in the table.

  $T_{0,\beta}(R)$,$\beta = 0.52$ the mean typicality of audio samples were compared with β. If the mean typicality of an audio sample is less than β, it is said that it reached the ranges of suitable atypicality. The audio samples that meet this specification are tagged with *A* label in table J-1.

- $V_{\alpha,\beta}(R)$ stands for the subset of R which meets the suitable levels of value. Acceptable ranges of value were assigned to items that were rated as moderate, agree, and strongly agree. The items that were rated higher than neutral are considered to have reached acceptable levels of value. The neutral weight in this case is 0.52.

  $V_{\gamma,1}(R)$,$\gamma = 0.52$ the mean value of audio samples are compared with γ. If the mean value of an audio sample is greater than γ, it is said that it has reached the ranges of acceptable value. The audio samples that meet this specification are tagged with *V* label in table J-1.

- $AV(F, A)$: This notion obtains the mean value of function F over the set defined in A. (Please note this is different from Atypical label which is abbreviated as A)

- $ratio(A, B)$ The number of items in set A were divided by the number of items in set B.

- θ : is the threshold value for each of the criteria. In this application, a unique θ value is applied for all the evaluation criteria. θ was selected to be 0.5 in this application.

Table J-1. The *typ* and *val* measurements for the seven audio pieces in the second public survey. The typicality ratings indicate whether the audio samples are examples of musical improvisation. The value rating shows whether an audio sample is a good quality musical improvisation. The T symbol in the table stands for the 'typical' cases where the typ meaurement satisfies $T_{0.52,1}$. The V symbol stands for 'valuable' cases in which the val measurement satisfies $V_{0.52,1}$. The number of typical cases are seven and the number of valueable cases are two. There are no atypical cases in this table. The atypical cases are those which the typ measurement satisfies $T_{0,0.52}$.

Ratings	Piece 1	Piece 2	Piece 3	Piece 4	Piece 5	Piece 6	Piece 7
Is the audio an example of a musical improvisation? (Typicality)							
Mean Typicality	0.64 T	0.61 T	0.65 T	0.58 T	0.60 T	0.58 T	0.58 T
Is the audio a good musical improvisation? (Value)							
Mean Value	0.64 V	0.51	0.54 V	0.49	0.51	0.47	0.49

The mean typicality, and mean value are obtained in table J-1. The implementation of Ritchie's criteria for the second survey is in the following:

(1) $Average(typ, R) > \theta \rightarrow$

$$\frac{0.64 + 0.61 + 0.65 + 0.58 + 0.6 + 0.58 + 0.58}{7} = 0.61$$

$$> 0.5 \xrightarrow{yields} Criteria\ 1\ is\ TRUE$$

(2) $Ratio\big(T_{\alpha,1}(R), R\big) > \theta \rightarrow \frac{7}{7} = 1 > 0.5 \xrightarrow{yields} Criteria\ 2\ is\ TRUE$

(3) $Average(val, R) > \theta \rightarrow$

$$\frac{0.64+0.51+0.54+0.49+0.51+0.47+0.49}{7} = 0.5214 > 0.5 \xrightarrow{yields} Criteria\ 3\ is\ TRUE$$

(4) $Ratio\left(V_{\gamma,1}(R), R\right) > \theta \to \frac{2}{7} = 0.29 < 0.5 \xrightarrow{yields} Criteria\ 4\ is\ FALSE$

(5) $Ratio\left(V_{\gamma,1}(R) \cap T_{\alpha,1}(R), T_{\alpha,1}(R)\right) > \theta \to$

$\frac{2}{7} = 0.29 < 0.5 \xrightarrow{yields} Criteria\ 5\ is\ FALSE$

(6) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), R\right) > \theta \to$

$\frac{0}{7} = 0 < 0.5 \xrightarrow{yields} Criteria\ 6\ is\ FALSE$

(7) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), T_{0,\beta}(R)\right) > \theta \to$

$\frac{0}{0} = undifined \xrightarrow{yields} Criteria\ 7\ not\ applicable$

(8) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), V_{\gamma,1}(R)\right) > \theta \to$

$\frac{0}{2} = 0 < 0.5 \xrightarrow{yields} Criteria\ 8\ is\ FALSE$

(9) $Ratio(I \cap R, I) > \theta \to$

$\frac{0}{0} = undifined \xrightarrow{yields} Criteria\ 9\ not\ applicable$

(10) $\left(1 - Ratio(I \cap R, R)\right) > \theta \to$

$1 - \frac{0}{7} = 1 > 0.5 \xrightarrow{yields} Criteria\ 10\ is\ TRUE$

(11) $Average\left(typ, (R - I)\right) > \theta \to$

$equivalent\ to\ Criteria\ 1 \xrightarrow{yields} Criteria\ 11\ is\ TRUE$

(12) $Average\left(val, (R - I)\right) > \theta \to$

$equivalent\ to\ Criteria\ 3 \xrightarrow{yields} Criteria\ 12\ is\ TRUE$

(13) $Ratio\left(T_{\alpha,1}(R - I), R\right) > \theta \to$

$$equivalent\ to\ Criteria\ 2 \xrightarrow{yields} Criteria\ 13\ is\ TRUE$$

(14) $Ratio\big(V_{\gamma,1}(R-I),R\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4 \xrightarrow{yields} Criteria\ 14\ is\ FALSE$$

(15) $Ratio\big(T_{\alpha,1}(R-I),R-I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 2{,}13 \xrightarrow{yields} Criteria\ 15\ is\ TRUE$$

(16) $Ratio\big(V_{\gamma,1}(R-I),R-I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4{,}14 \xrightarrow{yields} Criteria\ 16\ is\ TRUE$$

(17) $Ratio(V_{\gamma,1}(R-I) \cap T_{\alpha,1}(R-I),(R-I)) > \theta \rightarrow$

$$\frac{2}{7} = 0.29\ < 0.5 \xrightarrow{yields} Criteria\ 17\ is\ FALSE$$

(18) $Ratio(V_{\gamma,1}(R-I) \cap T_{0,\beta}(R-I),(R-I)) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 6 \xrightarrow{yields} Criteria\ 18\ is\ FALSE$$

Appendix K. Evaluation of the Ritchie's Criteria for the Third Survey

K.1 Evaluating Ritchie's Criteria: Musical Improvisations

The first set of typicality, and value questions in the third survey are:

- Is the audio an example of a musical improvisation?
- Is the audio a good musical improvisation?

The symbols for the Ritchie's criteria formulations and the choices of parameters for the third survey are similar to the second survey discussed in the previous appendix.

Table K-1. The *typ* and *val* measurements for the seven audio pieces in the survey for professionals.

The typicality ratings indicate whether the audio samples are examples of musical improvisation. The value rating shows whether an audio sample is a good quality musical improvisation. The A symbol in the table stands for the 'atypical' cases where the typ meaurement satisfy $T_{0,0.52}$. The V symbol stands for 'valuable' cases in which the val measurement satisfies $V_{0.52,1}$. The number of atypical cases are seven and there are no valuable or typical cases in this table.

Ratings	Piece 1	Piece 2	Piece 3	Piece 4	Piece 5	Piece 6	Piece 7	Piece 8
Would you agree that this is an example of musical improvisation? (Typicality)								
Mean typicality	0.48 *A*	0.50 *A*	0.43 *A*	0.34 *A*	0.32 *A*	0.37 *A*	0.37 *A*	0.32 *A*
Would you agree that this is a good musical improvisation? (Value)								
Mean value	0.41	0.46	0.46	0.46	0.41	0.43	0.41	0.41

The mean typicality, and mean value are obtained in table Table K-1. The *typ* and *val* measurements for the seven audio pieces in the survey for professionals. The implementation of Ritchie's criteria for the third survey is presented in the following:

(1) $Average(typ, R) > \theta \rightarrow$

$$\frac{0.48 + 0.5 + 0.43 + 0.34 + 0.32 + 0.37 + 0.37 + 0.32}{8} = 0.39$$

$$< 0.5 \xrightarrow{yields} Criteria\ 1\ is\ FALSE$$

(2) $Ratio\big(T_{\alpha,1}(R), R\big) > \theta \rightarrow \frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 2\ is\ FALSE$

(3) $Average(val, R) > \theta \rightarrow$

$$\frac{0.41+0.46+0.46+0.46+0.41+0.43+0.41+0.41}{8} = 0.43 < 0.5 \xrightarrow{yields} Criteria\ 3\ is\ FALSE$$

(4) $Ratio\big(V_{\gamma,1}(R), R\big) > \theta \rightarrow \frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 4\ is\ FALSE$

(5) $Ratio\left(V_{\gamma,1}(R) \cap T_{\alpha,1}(R), T_{\alpha,1}(R)\right) > \theta \rightarrow$

$$\frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 5\ is\ FALSE$$

(6) $Ratio\big(V_{\gamma,1}(R) \cap T_{0,\beta}(R), R\big) > \theta \rightarrow$

$$\frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 6\ is\ FALSE$$

(7) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), T_{0,\beta}(R)\right) > \theta \rightarrow$

$$\frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 7\ is\ FALSE$$

(8) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), V_{\gamma,1}(R)\right) > \theta \rightarrow$

$$\frac{0}{0} = undifined \xrightarrow{yields} Criteria\ 8\ is\ unapplicable$$

(9) $Ratio(I \cap R, I) > \theta \rightarrow$

$$\frac{0}{0} = undifined \xrightarrow{yields} Criteria\ 9\ not\ applicable$$

(10) $\big(1 - Ratio(I \cap R, R)\big) > \theta \rightarrow$

$$1 - \frac{0}{8} = 1 > 0.5 \xrightarrow{yields} Criteria\ 10\ is\ TRUE$$

(11) $Average\big(typ, (R - I)\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 1 \xrightarrow{yields} Criteria\ 11\ is\ FALSE$$

(12) $Average\big(val, (R - I)\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 3 \xrightarrow{yields} Criteria\ 12\ is\ FALSE$$

(13) $Ratio\big(T_{\alpha,1}(R - I), R\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 2 \xrightarrow{yields} Criteria\ 13\ is\ FALSE$$

(14) $Ratio\big(V_{\gamma,1}(R - I), R\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4 \xrightarrow{yields} Criteria\ 14\ is\ FALSE$$

(15) $Ratio\big(T_{\alpha,1}(R - I), R - I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 2{,}13 \xrightarrow{yields} Criteria\ 15\ is\ FALSE$$

(16) $Ratio\big(V_{\gamma,1}(R - I), R - I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4{,}14 \xrightarrow{yields} Criteria\ 16\ is\ FALSE$$

(17) $Ratio(V_{\gamma,1}(R - I) \cap T_{\alpha,1}(R - I), (R - I)) > \theta \rightarrow$

$$\frac{0}{8} = 0\ < 0.5 \xrightarrow{yields} Criteria\ 17\ is\ FALSE$$

(18) $Ratio(V_{\gamma,1}(R - I) \cap T_{0,\beta}(R - I), (R - I)) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 6 \xrightarrow{yields} Criteria\ 18\ is\ FALSE$$

K.2 Evaluating Ritchie's Criteria: Dastgāh-likeness

The second set of typicality, and value questions in the third survey are:

- To what extent is the generated audio an example of the Dastgāh Persian music genre?
- To what extent is the produced audio a high quality example of Dastgāh?

The symbols for the Ritchie's criteria formulations and the choices of parameters for the third survey are similar to the second survey discussed in the previous appendix.

Table K-2. The *typ* and *val* measurements for the seven audio pieces in the survey for professionals.

The typicality ratings indicate whether the audio samples are examples of Dastgāh music. The value rating shows whether an audio sample is a quality example of Dastgāh music. The A symbol in the table stands for the 'atypical' cases where the typ meaurement does not satisfiy $T_{0.52,1}$. The V symbol stands for 'valuable' cases in which the val measurement satisfies $V_{0.52,1}$. The number of atypical cases are seven and there is only one valuable case. There are no typical cases in this table.

Ratings	**Piece 1**	**Piece 2**	**Piece 3**	**Piece 4**	**Piece 5**	**Piece 6**	**Piece 7**	**Piece 8**
To what extent is the generated audio an example of the Dastgāh Persian music genre? (Typicality)								
Mean typicality	0.34 *A*	0.34 *A*	0.39 *A*	0.34 *A*	0.30 *A*	0.32 *A*	0.32 *A*	0.30 *A*
To what extent is the produced audio a high quality example of Dastgāh? (Value)								
Mean value	0.55 *V*	0.48	0.50	0.50	0.46	0.46	0.48	0.48

The mean typicality, and mean value are obtained in table K-2. The implementation of Ritchie's criteria for the third survey is presented in the following:

(1) $Average(typ, R) > \theta \rightarrow$

$$\frac{0.34 + 0.34 + 0.39 + 0.34 + 0.3 + 0.3 + 0.32 + 0.32 + 0.3}{8} = 0.37$$

$$< 0.5 \xrightarrow{yields} Criteria\ 1\ is\ FALSE$$

(2) $Ratio\left(T_{\alpha,1}(R), R\right) > \theta \rightarrow \frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 2\ is\ FALSE$

(3) $Average(val, R) > \theta \rightarrow$

$$\frac{0.55+0.48+0.5+0.5=0.46+0.46+0.48+0.48}{8} = 0.49 < 0.5 \xrightarrow{yields} Criteria\ 3\ is\ FALSE$$

(4) $Ratio\left(V_{\gamma,1}(R), R\right) > \theta \rightarrow \frac{1}{8} = 0.125 < 0.5 \xrightarrow{yields} Criteria\ 4\ is\ FALSE$

(5) $Ratio\left(V_{\gamma,1}(R) \cap T_{\alpha,1}(R), T_{\alpha,1}(R)\right) > \theta \rightarrow$

$$\frac{0}{0} = undedined \xrightarrow{yields} Criteria\ 5\ is\ not\ applicable$$

(6) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), R\right) > \theta \rightarrow$

$$\frac{1}{8} = 0.125 < 0.5 \xrightarrow{yields} Criteria\ 6\ is\ FALSE$$

(7) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), T_{0,\beta}(R)\right) > \theta \rightarrow$

$$\frac{1}{8} = 0.125 < 0.5 \xrightarrow{yields} Criteria\ 7\ is\ FALSE$$

(8) $Ratio\left(V_{\gamma,1}(R) \cap T_{0,\beta}(R), V_{\gamma,1}(R)\right) > \theta \rightarrow$

$$\frac{1}{1} = 1 > 0.5 \xrightarrow{yields} Criteria\ 8\ is\ TRUE$$

(9) $Ratio(I \cap R, I) > \theta \rightarrow$

$$\frac{0}{0} = undifined \xrightarrow{yields} Criteria\ 9\ not\ applicable$$

(10) $\left(1 - Ratio(I \cap R, R)\right) > \theta \rightarrow$

$$1 - \frac{0}{8} = 1 > 0.5 \xrightarrow{yields} Criteria\ 10\ is\ TRUE$$

(11) $Average\left(typ, (R - I)\right) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 1 \xrightarrow{yields} Criteria\ 11\ is\ FALSE$$

(12) $Average\left(val, (R - I)\right) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 3 \xrightarrow{yields} Criteria\ 12\ is\ FALSE$$

(13) $Ratio\big(T_{\alpha,1}(R-I), R\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 2 \xrightarrow{yields} Criteria\ 13\ is\ FALSE$$

(14) $Ratio\big(V_{\gamma,1}(R-I), R\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4 \xrightarrow{yields} Criteria\ 14\ is\ FALSE$$

(15) $Ratio\big(T_{\alpha,1}(R-I), R-I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 2{,}13 \xrightarrow{yields} Criteria\ 15\ is\ FALSE$$

(16) $Ratio\big(V_{\gamma,1}(R-I), R-I\big) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 4{,}14 \xrightarrow{yields} Criteria\ 16\ is\ FALSE$$

(17) $Ratio(V_{\gamma,1}(R-I) \cap T_{\alpha,1}(R-I), (R-I)) > \theta \rightarrow$

$$\frac{0}{8} = 0 < 0.5 \xrightarrow{yields} Criteria\ 17\ is\ FALSE$$

(18) $Ratio(V_{\gamma,1}(R-I) \cap T_{0,\beta}(R-I), (R-I)) > \theta \rightarrow$

$$equivalent\ to\ Criteria\ 6 \xrightarrow{yields} Criteria\ 18\ is\ FALSE$$

Appendix L. Liquid Persian Music Manual

LPM is the latest version of LBM released. It explores the concept of artificial life systems to control the generation and synthesis of audio. In this version, sounds are produced by synthesis toolkit stringed instruments class and OpenAL. Pattern-matching rules classify output from the CA and update the parameters of the synthesizer. Users can manually alter the pattern-matching rules to control the way the synthesizers' parameters change.

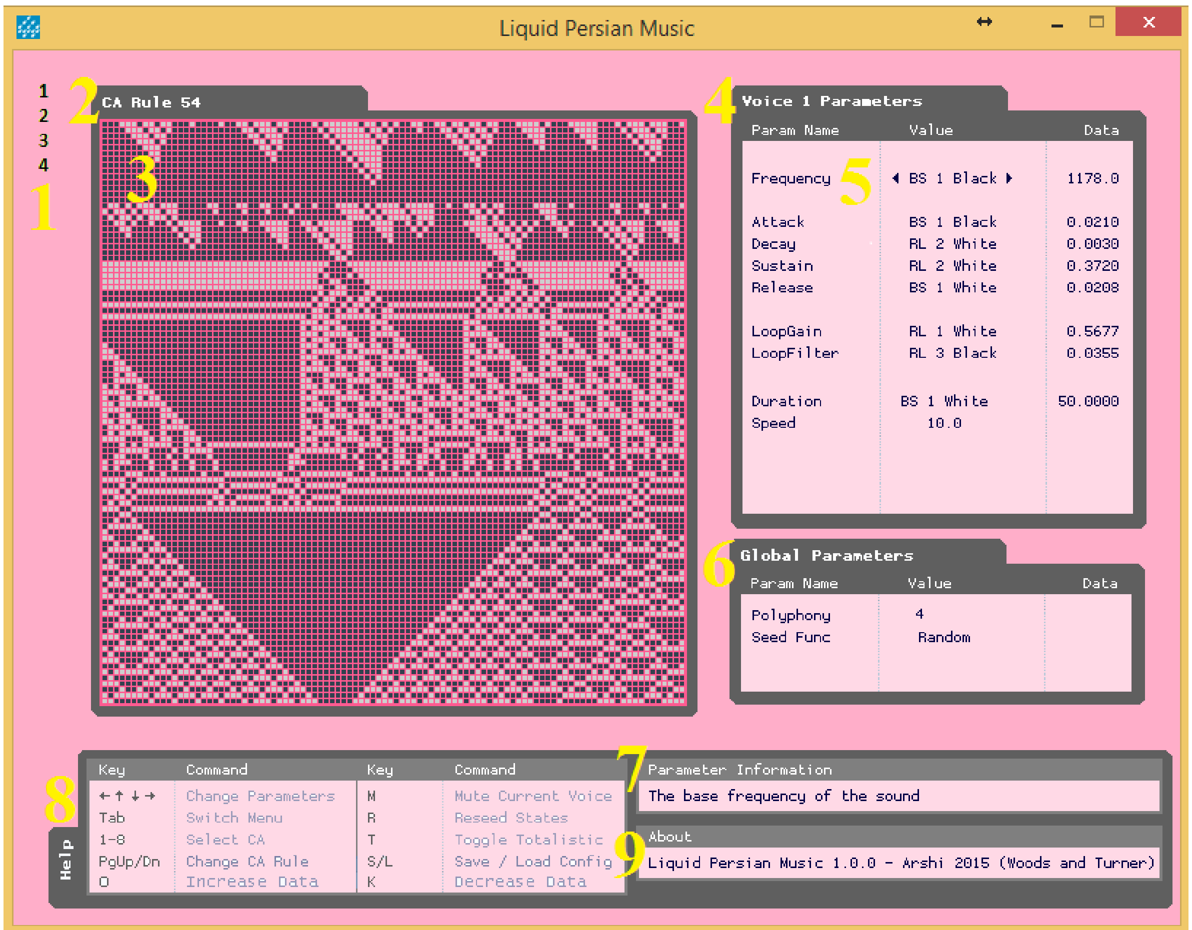


Figure L-1. LPM user interface.

L.1 Software Specifications

- ❖ Individual voice parameter controllers.

- Four concurrent instrument voices.
- Sound synthesizer parameter controllers including: Note frequency, ADSR envelopes, Loop gain.
- Twenty pattern-matching rules to control synthesizer parameters arbitrarily.
- Global parameters controllers as well as: number of polyphony, and the seeding.

L.2 Software Parts Description

1) Active voices depict the number of voices that are currently working
2) Cellular automata rule number for the selected voice.
3) Time space evolution of CA for the current voice number.
4) Local parameter display menu lists the selected parameters of the current selected voice and their selected pattern-matching rule relevant values,
5) Global parameter display menu: a controller for changing the polyphony and seed values. The user can determine the number of concurrent voices by changing the number of polyphony; each of the voices can have different controllable sets of parameters.
6) Parameter -Information panel: gives a brief explanation about the selected parameter in the controller menu.
7) About panel: indicates the developers, version and release year of the software.

L.3 User's Manual

- Navigating between local and global menu items: Up and Down cursor keys enable the users to navigate between menu items.
- Changing the pattern-matching rules for a parameter: right and Left cursor keys gives the ability to navigate between different pattern-matching rules. If the

manual menu is selected from the pattern-matching rules, then the user can adjust the data parameters by “O”/ “K” keys for increasing/decreasing the parameter values.

- Changing the CA rules: To change the CA rule e for the current voice, the “Page Up” /“Page Down” keys are applied. This range varies from 0 to 256 for a one dimensional CA with two possible states.
- Reseeding CA: pressing “R” key generates a row of random states in CA which will be evolved in the next iterations.
- Voice selection: to navigate and adjust the parameter of different voices the “1-8” keys are used.
- Muting a voice: “M” key, alternates the state of a voice between muted/ unmuted.

Save/Load keys: The “S” key saves all the voices parameters to a database. The Load menu item gives the user the opportunity to first select the saved item from the database by choosing the row number and then load it by pressing the “L” bottom.