2012

Emergent Rhythmic Structures as Cultural Phenomena Driven by Social Pressure in a Society of Artificial Agents

Magalhaes Martins, Joao Pedro

http://hdl.handle.net/10026.1/1185

http://dx.doi.org/10.24382/4631
University of Plymouth

Emergent Rhythmic Structures as Cultural Phenomena Driven by Social Pressure in a Society of Artificial Agents

João Pedro Magalhães Martins

Faculty of Arts

University of Plymouth

A thesis submitted to the University of Plymouth in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy

August 2012

Abstract

Emergent Rhythmic Structures as Cultural Phenomena Driven by Social Pressure in a Society of Artificial Agents

João Pedro Magalhães Martins

This thesis studies rhythm from an evolutionary computation perspective. Rhythm is the most fundamental dimension of music and can be used as a ground to describe the evolution of music. More specifically, the main goal of the thesis is to investigate how complex rhythmic structures evolve, subject to the cultural transmission between individuals in a society. The study is developed by means of computer modelling and simulations informed by evolutionary computation and artificial life (A-Life). In this process, self-organisation plays a fundamental role. The evolutionary process is steered by the evaluation of rhythmic complexity and by the exposure to rhythmic material.

In this thesis, composers and musicologists will find the description of a system named A-Rhythm, which explores the emerged behaviours in a community of artificial autonomous agents that interact in a virtual environment. The interaction between the agents takes the form of imitation games.

A set of necessary criteria was established for the construction of a compositional system in which cultural transmission is observed. These criteria allowed the comparison with related work in the field of evolutionary computation and music.

In the development of the system, rhythmic representation is discussed. The proposed representation enabled the development of complexity and similarity based measures, and the recombination of rhythms in a creative manner. A-Rhythm produced results in the form of simulation data which were evaluated in terms of the coherence of repertoires of the agents. The data shows how rhythmic sequences are changed and sustained in the population, displaying synchronic and diachronic diversity. Finally, this tool was used as a generative mechanism for composition and several examples are presented.

Contents

List of Figures

Acknowledgements

Many people and institutions provided the support that made this work possible.

First, I would like to thank the Leverhulme Trust for financially supporting this research.

To my supervisors Eduardo Reck Miranda and Guido Bugmann for their constant guidance and support.

To all my colleagues at the ICCMR, specially Adolfo Maia Jr., Alexis Kirke, Alicja Knast, Andrew Brouse, Asad Malik, Christian Dimpker, David Bessell, Dimitri Papadimitriou, Etienne Drouet, Hilary Mullaney, Jaime Serquera, Jean-Jacques Maurice, John Matthias, John Murray, Martijn Bosma, Nigel Morgan, Nikolas Valsamakis, Noris Norowi, Peter Beyls, Qijun Zhang, Sigalit Hed, Simon Durrant, Simon Ible, Torsten Anders and Vadim Tikhanoff, for all the interesting discussions and companionship that have shaped me as a researcher and as a person.

To the University of Plymouth and its staff, specially Susan Matheron, Sarah Carne, Carole Watson, Sue Kendall and Lucy Cheetham.

To Peter Todd, who invited me for a visit to the Max Planck Institute in Berlin and to Simon Kirby, who hosted me at the University of Edinburgh. These two visits, along with the great discussions I had with the members of both research groups, opened my mind to new approaches in the fields of artificial intelligence and music.

To Oliver Bown, Olivia Ladinig and Darrell Conklin. With them I had valuable discussions about the research topics which are more closely related to my work.

To all the reviewers of the publications included, and to the external and internal examiners of this thesis, for taking their time to assess my research and provide comments that enrich my work.

To Paulo Martins, who went through my thesis with the critical lens of someone experienced in another research field.

To my friends Leandro Costalonga and Marcelo Gimenes, also members of the ICCMR. Together we formed the locally famous *chorinho* trio “The weeping crocodiles”, although nobody ever weeped at our performances.

To the members of my fado group “Verdes Anos”, António Dinis, Gonçalo Mendes, Luís Barroso, Miguel Drago and Rui Seoane and to my guitar trio partners Joana Sampaio and Tiago Correia. With these people I shared many hours of music making and with whom I have an enormous debt of friendship.

To all the people that visited me in Plymouth from far away, and specially to my brother Paulo, to Rui Namora and Miguel Drago who did it more than once.

To Anna Zacharioudaki and Frederic Verret - two good friends who saved me a couple of times when I needed the most - and to the friends from the Alien house.

To my mentors in earlier stages of my learning process José Raposo, Graciano Pinto, and the late Matti Karjalainen, without whom I would never have been so fortunate in being able to combine science and music. This thesis is also dedicated to them, for their commitment to teaching and for being my constant source of inspiration.

To my students, who have given me plenty of motives for feeling honoured and happy with my position.

To Eduardo Coutinho and Patrício da Silva who provided invaluable advice, help and friendship in different stages of this path. Also, a big hug to the family Urbano Coutinho. You guys kept the beat going.

To my family who have made me what I am and have been constantly there for me. You have provided me with all the possible and sometimes with the impossible.

To Helena Dürbaum, with love.

Authors declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

Relevant scientific seminars and conferences were regularly attended at which work was often presented.

This study was financed with the aid of a studentship from the Leverhulme Trust.

A record of activities can be found in Appendix A.

Signed: ______________________

Date: ______________________

Word count for the main body of this thesis: 52629

Abbreviations

AE	Artificial Environment
AI	Artificial Intelligence
A-Life	Artificial Life
BEAST	Birmingham ElectroAcoustic Sound Theatre
BP	Bol Processor
BPM	Beats per minute
EC	Evolutionary Computation
EMI	Experiments in Musical Intelligence
FSA	Finite State Automaton
GA	Genetic Algorithm
GP	Genetic Programing
GTTM	Generative Theory of Tonal Music
ICCMR	Interdisciplinary Centre for Computer Music Research (Plymouth)
IGA	Interactive Genetic Algorithm
IOI	Inter-onset interval
IRCAM	Institut de Recherche et Coordenation Acoustique/Musique (Paris)
MWFR	Metrical well-formedness rules
MPR	Metrical preference rules
NN	Neural Networks
OMME	Ontomemetical Model of Music Evolution
PS-Measure	Povel-Shmulevich measure of complexity
RU	Rhythmic Unit
SO	Self-organisation
TN	Transition Networks
WFC	Well-Formedness Constraints

Chapter 1

Introduction

In an age where music is being exchanged rapidly all over the globe, we can hardly remember the amazing fact that music existed through thousands of years, before the appearance of recording devices, in the heads of the people who used this form of art in all sorts of human activities. Whether it is the mother that tries to sooth her crying baby by singing a melody, or the fishermen who sing rhythmically to synchronise their movement when pulling the net from the sea, or the community of pygmies who sing independent rhythmic voices to create a single complex texture of sound , or the group of youngsters who get together in a garage to compose tunes, we can find displays of music in all cultures and in a multitude of forms and social contexts.

Most of these tunes were not created by the persons who interpret them. Some are folk tunes residing in the collective memory without identification of their author. Yet, they fulfil an important functions to both the individual and the society.

The emergence of a standardized music notation system in 10th century was a form of communicating pieces to distant musicians, and to leave a relatively faithful record of the music to future generations.

The appearance of the printing press in the 15th century brought changes to the lives of composers. John Dowland in the 16th century, was one of the first composers whose works were printed, and as a result he became one of the wealthiest composers of his time. Tunes from famous composers were being sought for all over Europe, and in the 19th century musicians in Europe developed a special interest in the music of composers from the past. The royalties gained by composers with the selling of their music was one of the reasons why, during the romantic period, some were able to become independent from patrons of the nobility or the church, fulfilling the ideal of the romantic artist.

Today the internet has changed the way how professional musicians and ordinary people publish their music. It is probably soon to understand the the whole phenomenon but the consequences are already visible. All the changes in music technology have had an impact into the music content that was produced, but many musicians, if not most, still learn music the way our ancestors did - by imitation. This process not only takes place in the initial learning stages but also later in life of a musician. Even highly acclaimed artists recognise that much of their works is subject to influences from the surrounding environment. Salvador Dali, has put it in rather radical terms by saying that "Those who do not want to imitate anything, produce nothing" (Dali 1970 p.173).

The process of imitation can be rather detailed reproduction of the original source, or it can focus more on some features that define a particular music style. Music depends so much on innovation as on imitation. Total imitation plays a central role in rituals, work songs and lullabies - in every music with a functional character - but without novelty, no diversity would be possible. What is considered to be novel or strange today, will eventually evolve into becoming the rule of tomorrow.

I chose rhythm as a research subject as it seems to me the most basic feature in music, and also because it was the feature with which I struggled the most since the beginning of my music studies. Another reason is the fact that rhythm is a strong identifying trait of a music culture. Latin american music differentiates its styles by looking into the pattern of accented beats in its metrics (bossa-nova, salsa, rumba, etc.). This can also be observed in balkan music with its complex patterns of beats, or in the andalusian music of flamenco, where genealogical trees of styles are identified on the basis of beat patterns (bulería, soleá, fandango, rumba, etc.) (Toussaint 2002; Díaz-Bañez et al. 2004).

Instead of looking into data from existing music I chose to investigate the dynamics of cultural transmission with the help of simulations and evolutionary computation. For me it is a personal goal to use these processes in composition and at the same time to learn the emergent properties of such a complex system. I believe this study contributes to the field of evolutionary musicology, and in the future this field will play an important role in understanding how the internet and the current changes in technology will shape the creation and appreciation of music.

1.1 Research questions

The questions that motivated my research are as follows:

- Q1. Is it possible to use multi-agent systems to model the evolution of rhythms? What would be the criteria for the design of such a system?
- Q2. How can rhythms be represented and the product of this representation be recombined in a creative manner?
- Q3. What could such multi-agent systems models contribute to our understanding of the motivations for change and evolution of music?
- Q4. What could such systems contribute to our understanding of how the dynamics of musical culture? For instance, how can such volatile things as unwritten pieces of music hold for several generations, or spread to large parts of the population, and what are the features in oral transmission that enable these processes to take place?

- Q5. Would such multi-agent systems be useful for composers? For instance, would it be possible to use the results of simulations in actual musical compositions?

The goal of this thesis is to study the evolution of rhythms in an artificial environment, where these are produced and selected by artificial intelligence agents in a cultural transmission setting. The selection process of the rhythms is guided in a bottom-up manner using feedback from neighbouring agents based on evaluations of complexity and exposure.

1.1.1 Rational

In engineering, multi-agent systems are used to solve a problem where distributed processing can provide better results than analytical or traditional sequential approaches. There is yet another application of multi-agent systems, which is to simulate living systems, by trying to model the behaviour of individuals in a society and looking at the results of the collective behaviour. The choice of multi-agent systems derives directly from the goal of looking into the cultural aspects of music. The framework of multi-agent systems, used in the context of this thesis, enables the treatment of autonomous entities processing musical information and interacting in a virtual environment. In terms of composition, these artificial agents can be seen as metaphors for different entities such as music players, listeners, musical instruments, voices, etc.

Rhythm is a large field of research within musicology, and it is the most basic structural element in music from all cultures. The adoption of rhythm pertains to the fact that rhythmic structures can be evaluated, and have been done so by psychological studies, for both information and performance complexity. The notion of the evolution of rhythmic complexity as a result of the interaction between the agents is one of the central issues developed throughout the thesis. Namely, how complexity evolves as a result of exposure to rhythms.

The debate nature vs. nurture or the establishment of boundaries between behaviours which are genetically determined or culturally transmitted is a long running debate with difficult progression due to the up to now scarcity of knowledge in biogenetics. This situation is rapidly changing with the new discoveries of genes and gene complexes, and their influence in human behaviour. One of the ways in which these questions can be addressed is through simulations where genetic vs. environmental factors can be compared.

The topics covered by this thesis are of relevance for scientists and composers investigating cultural transmission. On one side, with the proposed system it is possible to investigate the evolved behaviours within the multi-agent system, and on the other side it is possible to use the artifacts, meaning the rhythms themselves, created by the system for composition.

1.1.2 Methodology

In order to shed light into the problems faced by the posed questions, the work presented in this thesis involved the creation of computational tools for exploring the notion of socially constructed rhythmic sequences in a process of self-organisation. Specifically, I created a new multi-agent system, named A-Rhythm, where agents interact in a virtual environment, to study different scenarios of interaction.

The multi-agent system comprises a set of interactive autonomous agents, endowed with rhythm perception/production capabilities which are placed in a virtual 2D environment and exchange rhythmic structures. The system A-Rhythm has two implementations: the first is aimed at studying different modes of interaction and how these affect the emergence of new rhythms; the second contains a grammar and the agents evolve syntactic rhythmic structures guided by additional perceptual constraints.

In developing both implementations of the system, it was important to look for a symbolic representation of rhythmic memes[1] and find a mathematical methodology to construct a grammar that originates the rhythmic structures. The chosen representation takes the form of sequences of inter-onset intervals which characterise the rhythmic units. The chosen grammar took the form of a Markov process.

Furthermore, measures of complexity and similarity were developed to evaluate and drive the evolutionary process.

The analysis of the system was conducted by plotting the evolution of different values for the individual agents and for the society as whole. Examples of this variables are the size of the repertoires, the complexity of the rhythms, the lifetime of a sequence, number of agents sharing a rhythm, or the hedonic evaluations of the agents.

Using the similarity measurement between repertoires of the agents, further statistical analysis of the society was carried out. This involved hierarchical clustering of the agents displayed in dendrograms, and principle component analysis (PCA) of the similarity tables.

The output of the simulations, meaning the rhythmic structures generated by the agents, were used in several compositions, of which two examples are shown in this thesis.

1.1.3 Evaluation

The evaluation of the system takes place at two different stages in the thesis, namely, when characterising the features of the architecture in comparison to other systems and when observing the produced behaviours. Regarding the architecture of the system, a set of criteria is defined to establish the basis for evaluation, and at the same time to describe the properties of existing systems in the literature. These criteria enable the

[1] A meme is a basic unit of cultural transmission in the same way that a genes, in biology, are units of genetic information. This term was first introduced by Richard Dawkins in the *Selfish Gene* (Dawkins 1989).

critical reasoning of existing works and delimit the scope of research. Further in the thesis the system is analysed in terms of the behaviour of both the individual agents and the considered communities of agents. Specifically, this analysis focuses on the questions of diversity and complexity of repertoires and the emergence of grammars of the agents, depending on the contextual organisation of agents' societies.

The generated rhythms have also been used to write music pieces that mirror the process of artificial cultural evolution developed in A-Rhythm.

1.2 Plan of the thesis

The thesis starts by providing a theoretical overview of the field, dedicating two chapters to the subject of rhythm and computational methods for music analysis and composition (Chapters 2 and 3). Chapter 4 presents an overview of related research with systems that focus on musicological and compositional goals. Chapters 5 and 6 show the development of the system, divided in the preliminary stages and the final grammar model, and presents the results from the simulations. Chapter 7 discusses the compositional applications of the systems developed and present two compositions based in data collected from the simulations. The final chapter of the thesis provides the discussion about the system and contributions of the thesis to the field of Computer Music, in particular the field of Evolutionary Computer Music. The appendices contain list of activities and conferences attended, the scores for the pieces discussed in chapter 7 and a glossary of terms.

Someone reading this thesis with an engineering perspective, might want to go deeper on the technical discussions presented in chapter 5 and chapter 6. Someone reading the thesis with a musical perspective, and a stronger focus in computer science, might prefer to skip the the technical discussions related to computational models and go deeper on chapter 2, on Rhythm, and chapter 7, on Composition.

1.3 Context

The next sections present the contexts in which A-Rhythm was developed.

1.3.1 Rhythm in the context of computer music

Rhythm is a key element in music. Not only organises all musical information in time, as well exercises the capability of entrainment. Entrainment, or the property by which human beings are able to synchronise to an external pulse, is the process that enables people to play in groups in an inherently coherent manner.

Studies in time estimation and synchronisation go back to the 19th century (von Vierordt 1868; Stevens 1886), where it was found that people tend to underestimate long intervals and overestimate short ones. Much research continues to be made on the subject, and quite recently, experiments confirmed that beat perception is present in newly born babies (Honing et al. 2009). For the purpose of this thesis, the works of Povel and Essens (1985), Shmulevich and Povel (2000), Longuet-Higgins and Lee (1984) and other works concerning rhythmic complexity, play an important role. The reason for this is twofold: rhythmic complexity is considered to be an important factor in music appreciation and current evolutionary models lack the use of complexity as a structural feature of both design and analysis. In chapter 2, research works in the field of rhythm will be reviewed and chapters 5 and 6 explain how rhythmic complexity and similarity are developed in context of A-Rhythm.

1.3.2 Music in culture

Alan Merriam, the famous ethnomusicologist has said that “Music sound cannot be produced except by people for other people” (Merriam 1964 p.6). This quote is brought up when trying to define the field of ethnomusicology as “music in culture”.

This quote stresses the fact that music is inherently a social activity. Merriam also alludes to the fact that musicologists have focused much more on music structure, ignoring the contexts that have originated the music pieces (Merriam 1964 p.29). Since 1964 the situation has changed, and the effect of the surrounding environment has been taken into account more seriously by ethnomusicologists and anthropologists. This has helped in creating a more complete panorama of music made in specific contexts but still little is known about how the culture directly shapes the aesthetical behaviour of individuals or conversely, how culture is shaped by the action of individuals.

Due to the complexity of human behaviour, this problem is probably one of the most difficult to be grasped by the scientific method.

In 1866 the French Academy of Sciences issued a ban on publications on the topic of language evolution, deeming it to be an unsolvable problem. Due to the appearance of new analytical methods, more than a century later, researchers have been able to come back to this problem with a refreshed interest (Cangelosi and Parisi 2002; Christiansen and Kirby 2003; Cangelosi et al. 2006). Similarly, the study of music evolution has faced similar problems in the past and only quite recently this topic has been addressed in multidisciplinary terms (Wallin et al. 2001).

In the end of the 20th century, Steven Pinker compared music to “auditory cheesecake”, meaning that from an evolutionary perspective it was simply a spandrel, a by-product of language (Pinker 1997). This provocative statement has raised responses from various researchers who believe otherwise, but also from researchers who are willing to take this hypothesis seriously still considering the inescapable fact that music is extremely important for most human beings being an important part of their daily lives (Honing 2011).

One thing we are sure of is that all cultures have music, but only a small part of all possible music becomes part of a particular culture.

If we think of music in terms of information processed by the brain, there are potentially infinite pieces of music that can be created and appreciated. The constant expansion of the boundaries of the concept of music has created new dimensions that need to be addressed by research. Some music is created with acoustic instruments, like all music known before the 19th century, but also an enormous amount today is created using synthesised or electrically transformed sounds. Some music is harmonic, in the sense of the western tradition, but a vast number of musicians do not regard harmony as a primary structuring element. These two examples are drawn here only to illustrate the evident notion that culture is shaped both by time and space.

More specifically, and pointing to the subject of this thesis, some music is immediately recognised by a significant part of the population, and other music is only meaningful (and very importantly so) to a small number of people. Some music lives in the head of many generations, whereas some pieces only exist for a short period in the head of their creators.

1.3.3 Music preference

To the question "why do we like certain tunes?" Minsky (1981) attempts to provide with two possible answers:

- *Because they have certain structural features*;
- *Because they resemble other tunes we like*.

The first answer is related to the structural features in a tune that contribute to make it pleasant. This features could be, in analogy to language, described by a set of grammatical rules which would define if a stream of sounds is syntactically admissible, or should contain elements that "make them sensible or even pleasant to the ear" (Minsky 1981 p.32). The investigation of these features can be observed in the grammatical theory explained in the Generative Theory of Tonal Music by Lerdahl and Jackendoff (1983).

The second answer, relates to past experiences contributing to the appreciation of new music. Minsky argues that music is by no means only structure and other factors contribute to the internal processing of new pieces, involving the attribution of meaning to music. Therefore, he points out that there are some shortcomings in describing music with a set of syntactic rules. For this, one must make an effort to understand how music is processed and memorised.

In this thesis, there are some factors not being considered, which undoubtedly contribute to music appreciation. The first is the capacity for music to affect emotional states. Music is often referred to as the "language of emotions", and there is a whole body of research on how music is used as a tool in expressing or inducing emotional states (Meyer 1956; Juslin and Sloboda 2010). Out of the scope of this thesis is the issue of identification of authorship in music. Surely, the association of a music piece to its author plays an important role in the appreciation of music, but it is also very common for traditional music not to have a reference to its author. Also left out were social factors extrinsic to music context, such as empathy towards a particular performer, or appreciation due to other social contexts apart from the controlled conditions of the simulations.

1.3.4 Dynamics of change

Music underwent changes across time as new elements were introduced by composers in their pieces. Western music has seen radical changes in style when Beethoven incorporated the notions of the romantic ideals and changed the way how dynamics was used in a music piece. Later, the presupposed rules of harmony started cracking with Liszt, Debussy and Wagner and were totally dissolved by Schönberg and the followers of second Vienna school. Russolo, in the 20th century introduced the concept of noise as musical object and the appearance of electronics gave rise to a whole new set of sounds that expanded our culturally accepted notion of music.

These changes have had an identifiable origin which was followed by many. Nevertheless, these artists were under the influence of many other artists which have influenced them and ultimately gave some contribution to the more or less abrupt transitions in style.

In this thesis, transformation of rhythm is considered on the level of the basic rhythmic elements. The rhythms are cumulatively transformed from existing material in the memory of artificial agents.

1.3.5 Self-organisation

Self-organisation is a property playing an important role by the field of Artificial-Life (A-Life). It concerns the development of patterns without a central control or externally delineated plan. These patterns are not observed at the local level, but emerge in the macro level when special conditions arise in the environment. Examples of self-organisation can be found in crystallisation of water, patterns in the skin of animals, ferromagnetism, etc. (Oudeyer 2006).

Self-organisation is also a property of systems of interacting parts, which show complex patterns or behaviours as a result of relatively simple modes of interaction. Self-organisation can be observed at different scales, such as the atomic level, where physical properties enable the formation of molecules; at the molecular level when molecules assemble into stable entities giving rise to proteins and other basic components of living organisms; at the level of the individual, where basic building blocks interact to achieve a task, like limbs in the locomotion process; or at the social level, where the individual behaviour of elements in a society influences the collective emergent properties in the absence of a centralised control, such as food foraging in ant colonies or swarming behaviour of birds.

Language can also be seen as self-organising system, as the emergence of communication systems does not obey a central plan. The evolution of language can be studied with the tools provided by Artifical-Life (A-Life). These tools, such as artificial neural networks and genetic algorithms, have been used to describe the evolution of vowel systems (Boer 1999), the evolution of a speech system (Oudeyer 2006), the biological, neural, and adaptive mechanisms that lead to the evolution of language (Cangelosi and Parisi 2002) and the evolution of grammatical structures via the iterated learning model (Kirby 2002).

Self-organisation is observed in multiple aspects of music. In the process of music performance we can easily see how the processes of synchronisation, or more obviously, singing in tune, do not actually need a conductor to take place. At the level of the society, the music scales that a particular music culture uses, or a particular rhythm that becomes the basis for a traditional dance, are also emergent properties not guided by anyone except by repeated action of the participants. The same can be argued for the popularity of a particular song, or the level of music complexity present in the repertoire.

The system implemented in this thesis involves many agents interacting in a virtual environment and the development of their repertoires as well as internal process are subject to self-organisation.

1.3.6 Computer simulations

In the field of A-Life and music, the reasons behind the usage of computer simulations are mainly two (Todd 1999):

- Simulations work as a proof of concept and serve to show that certain behaviours exist evolving from an initial state through a set of cumulative stages;

- Simulations are intrinsically one of the best tools for studying evolutionary processes.

Simulations must be simple in order to be able to run for a significative amount of iterations in replicating the evolutionary process.

There are obvious limitations to simplification of such a complex process. Todd compares it to the use of a *cheap electronic synthesiser to replay an orchestral symphony* (Todd 1999 p.362). McElreath and Boyd (2007) use a different analogy for using simple mathematical models to describe complex processes: as if one is looking at a map with a much smaller scale than the actual represented objects. In both cases the conclusion is the same: valuable insights can be gained from a process which is otherwise impossible to be understood.

In music, computational methods have been extensively used ever since the early days of the computer. These methods were applied mainly to composition, but in the last twenty years also musicology is using computers to study music structure and music evolution. A-Rhythm is a computational system designed to provide simulations of musical behaviour, that looks into the phenomena of music from both a musicological and compositional perspective.

1.3.7 Composition

Apart from some of the performers playing free improvisation, most composers and players in the history of music use some kind of self devised, or socially constructed, set of rules during the composition process. Normally these rules are not sufficient to construct interesting compositions or performances, and some form of conscious or unconscious decisions are made to produce the final musical piece. In the case of music algorithmically generated, usually some sort of manipulation of the data is required, in order to have an interesting musical experience.

In evolutionary computation, the algorithm itself provides a life-like quality characterised by a dynamical balance between the organisation and unpredictability. Many *a posteriori* decisions regarding mapping, play a significative role in the aesthetic experience, and despite the fact that the composer possesses many ways of intervening in the algorithm by changing the parameters, the evolutionary computation paradigm becomes a source of exploration for composers dealing with complex phenomena.

In the case of the system developed in this thesis, one of its goals is to provide with diverse musical material for composers to use in their musical pieces.

1.3.8 Sonification

The amount of manipulation a music piece requires is also a measure on how the algorithm becomes transparent by listening to its output. Some pieces, or sonic objects, are not meant to be aesthetically interesting, but to convey information about the patterns in the data. Sonification is the use of non-speech audio to convey information or perceptualize data (Kramer 1993). Visualisation of data is a more prevalent method in science than sonification. But for some years there have been many cases where sound actually gives a stronger impression of the patterns in the collected data. Examples of sonification devices are the Geiger counter and the sonar. Sonification of data has played an important role in composition as well. As an example, the piece "Navegar é preciso, viver nao é preciso" by Alberto de Campo and Christian Dayé (de Campo and Dayé 2006) is inspired in the circumnavigation trip by Magellan and uses statistical data from 15 countries nearest to the route, which is then mapped into sounds to create an electroacoustic music composition.

In this thesis the objective is to combine the both approaches. The system was obviously designed to function as a generator of potentially interesting musical material delivered by an evolutionary process. On the other hand there is a secondary goal which is to help the listener to focus on the simulation and provide a less abstract way of looking into the process of artificial evolution.

1.4 Summary of the chapters

The organisation of the chapters is as follows:

Chapter 2 presents the definition of rhythm, the basic elements of rhythm, aspects of rhythm perception, models of perception of rhythmic sequences and rhythm evolution;

Chapter 3 deals with present research in artificial creativity and using artificial life models applied to the scope of this thesis;

Chapter 4 shows an overview of the existing systems on the evolution of music;

Chapter 5 presents the stages of the development of A-Rhythm, including representation choice, rhythmic transformations, investigations on interaction games, the artificial environment, simple complexity evaluations and similarity;

Chapter 6 presents the second version of A-Rhythm, including the grammar and simulations based on complexity and exposure preferences;

Chapter 7 presents two compositions which illustrate the process of cultural evolution for the proposed systems;

Chapter 8 shows a critical analysis of the results obtained, the contribution of the thesis to knowledge, and perspectives for future research in the area.

Chapter 2

Rhythm Phenomenon

When Stravinsky *premiered* the “Rite of Spring” on 29 May 1913, he was not only redefining the extent of the concept of music to the western world, but he was also creating a shift in the arts. After achieving public notoriety with some earlier ballets coloured with the atmosphere of Russian folk music, the composer, who obtained 4 nationalities, created a piece full of innovative rhythmic patterns and metre changes, today thought to be the most influential piece of the 20th century in the western music repertoire. “The highly complex music and unusual choreography led to arguments breaking out between audience members about whether what they were experiencing was truly *art*, and these arguments soon escalated into full-scale fist-fights” (North and Hargreaves 2008 p.33).

2.1 Definitions and rhythmic behaviour

The meanings of the word rhythm are multiple and pervasive both in art and everyday life. Musical rhythm, work rhythm, circadian rhythms, rhythm of a painting - all have different meanings. The focus of this work is solely on musical rhythm, nonetheless the definition in this narrower is still hard to pin down. Ladinig (2009) gives an account of the difficulty that researchers have faced in trying to define musical rhythm.

The Oxford English Dictionary defines Rhythm as “a movement marked by the regulated succession of strong or weak elements” (London 2001). Windsor and Desain

(2000) define rhythm as a sequence “where the events occur with some repetitive structure”.

Rhythm is frequently associated with living beings. Different rhythms are present in locomotion and animal gait such as walking, running, trotting, galloping or swimming. Breading also has a rhythmic quality and in humans, sucking in newborns and rocking are also displays of rhythmic behaviour.

These are common to many other mammals, but synchronised movement with the goal of enjoyment is thought to be present only in humans.

2.2 Perception and production of rhythm

A number of researchers have developed theories on the human abilities to process rhythmic information (Cooper and Meyer 1963; Steedman 1977; Lerdahl and Jackendoff 1983; Povel and Essens 1985; Drake and Gérard 1989; Sundberg and Lindblom 1992; Parncutt 1994; Drake 1998; Large and Jones 1999).

Starting from the definition, a movement marked by the regulated succession of strong or weak elements, it then becomes important to define the difference between strong and week events. A strong event is one that possesses an accent. According to Cooper and Meyer (Cooper and Meyer 1963 p.8) accented events are “marked for consciousness in some way”. Lerdahl and Jackendoff (1983) describe three types of accents: the phenomenal accent, the metric accent and the structural accent. The differences between these accents are explained bellow (Sec. 2.5 on p.32).

The phenomenal accents are dependent or different perceptual cues such as inter-onset intervals (IOIs), intensity, duration, timbre, melodic contour or attack profiles. Research on perceptual relevance of the cues to perceived accents determines that the inter-onset interval is the most determinant cue influencing the subjective quality of the accent (Handel 1989; Parncutt 1994).

2.3 Internal clock and synchronisation

In assessing time, humans and particularly musicians, have to rely on the existence of an internal clock. People rely on an inner train of pulses in order to estimate time intervals and develop rhythmic structures with relation to it. This pulse is often called the *tactus* or *tempo* and the frequency measured in beats per minute (bpm). The importance of this notion for music theory has been stressed by numerous researchers (Steedman 1977; Handel and Oshinsky 1981; Handel and Lawson 1983; Povel and Essens 1985).

As mentioned in the introduction, the works of Povel and Essens (1985), Shmulevich and Povel (2000), and Longuet-Higgins and Lee (1984) play an important role in this thesis. The goal of these studies is to develop models of processing of rhythmic events with relation to an internal clock. In the case of Povel and Essens , their goal was to verify the existence of an internal clock and describe an efficient code for rhythms with relation to the induced internal clock. Longuet-Higgins and Lee (1984) had previously found that the metric structure plays an extremely important role in music listening. In their study, a measurement of rhythmic syncopation was created, and it was hinted that complexity and syncopation were closely related.

Later, Shmulevich and Povel (2000) extended the previous study of Povel and Essens to derive a measurement of complexity. This measurement of complexity will be used in the system implemented in the context of this thesis. The reason for this is twofold: rhythmic complexity is considered to be an important factor in music appreciation and current evolutionary models lack the use of complexity as a structural feature of both design and analysis.

In the next following sections these three studies are reviewed in more detail.

2.3.1 Investigation on the nature of the internal clock (Povel and Essens 1985)

The Povel and Essens study (Povel and Essens 1985) investigates the nature of the internal clock looking into different hypothesis:

- The absolute clock;
- The clock with a time unit derived from the sequence;
- The hierarchical clock.

The first hypothesis assumes that humans perceive an absolute value of a minimal time interval, say 1 ms, and encode the longer values in memory in terms of this value. This model fails to predict the filled duration illusion, where filled intervals are perceived longer than empty ones (Thomas and Brown 1974; Wearden et al. 2007; Repp 2009) and is also not able to explain why the same pattern played with a different *tempo* will be recognised as structurally identical. This model would also predict that patterns with identical number of intervals would be equally well perceived and reproduced, and this has been disproved by experimental results (Fraisse 1956; Povel 1981; Sternberg and Knoll 1982).

The second hypothesis assumes that the listener will be able to select a time unit which corresponds to the shortest time interval of the sequence, this being encoded in terms of multiples of this unit. This model would predict that sequences with multiple intervals of the unit of [1 1 2] and [1 2 2] should be easy to conceptualize and equally well reproduced which is not the case. For further examples please view Povel's paper (Povel 1981).

The third hypothesis of the "hierarchical clock" is more plausible and has been subject to extensive research in music (Lerdahl and Jackendoff 1983; London 2004; Temperley

2004). In the words of Povel and Essens (1985 p.404) "In simple music at least, there is an equally spaced pulse which might very well determine the unit of an internal clock. Interestingly, the pulsed intervals are not composed of very small durations, but rather of a medium duration which can either be subdivided or concatenated." This means that there is a regular pulse that depends on the sequences considered, working as perceptual mechanism or a "mode of attendance" (Cooper and Meyer 1963) to a rhythmic surface. This mode is regular hierarchical in nature as it works in several time scales corresponding to the division or concatenation of the pulses.

The provisionally proposed model considers only two time scales, defined by the unit and the subdivision of the unit. Hierarchies with more levels are certainly important in rhythm perception, both in speech and music (Martin 1972; Jones 1976), but the model considers only the two more salient levels.

In this study, the notion of accent is central to the question of internal clock induction. Although accents can be perceived with other cues, the most relevant cue is determined by the IOIs and the grouping structure. Following on a study of Povel and Okkerman (1981) on temporal sequences composed of identical tones, three rules have been devised for perceptually marked (accented) events. Events are perceptually marked if:

- They are relatively isolated;
- They are the second tone of a cluster of two tones;
- They are the first and last events in a cluster of three or more tones.

The accent rule applied to a sequence of rhythmic events is depicted in Fig. 2.1.

Povel and Essens (1985) then proceed to create a computer program of a clock induction. For all possible clocks, the valid ones being all those which have a period less or equal to half the length of the sequence, they calculate the score C of the induced

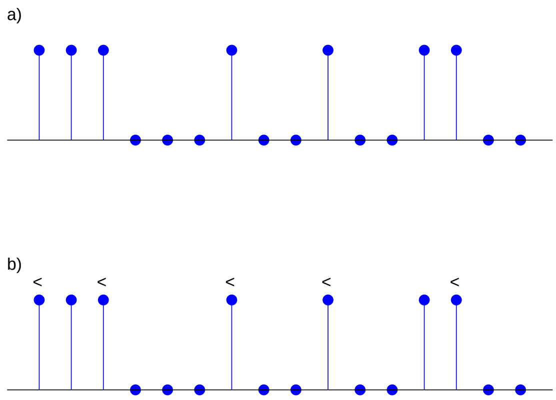


Figure 2.1: Accent rule applied to a sequence of events: (a) the temporal pattern as presented to the subjects; (b) the resulting accents determined by the rules (Povel and Essens 1985).

clock. This score depends on the number of clock ticks coinciding with an accented event ($+ev$), on the number of ticks coinciding with an unaccented event ($0ev$) or on the ones coinciding with silence ($-ev$). The score C results in the the following formula

$$C = (W * -ev) + (1 * 0ev) \tag{2.1}$$

where W is a parameter concerning the relationship between $-ev$ and $0ev$ and since it is assumed that these values reflect a relative negative effect to clock induction, W should be larger than 1.

As the sequences are assumed to be cyclical, only clocks with a frequency that results in ticks being spaced equally are admissible. This means that the clock unit (or period) has to be a divisor of the length of the sequence (div). Fig. 2.2 shows the clock induction process for a sequence with the IOIs [1 2 2 1 1 2 3] and $W = 4$.

The smallest value of C will provide the best clock. It is pointed out in the paper that this model is not a processing model. It does not work left to right in looking for the best

<table>
<tr><td>Input sequence</td><td>1 2 2 1 1 2 3</td></tr>
<tr><td>Transform into time-scale notation</td><td>1 1 0 1 0 1 1 1 0 1 0 0</td></tr>
<tr><td>Add accents</td><td>1 2 0 2 0 2 1 2 0 2 0 0</td></tr>
<tr><td>Generate all clocks
(unit < ½ period)
Apply weights and determine
divisor</td><td>
<table>
<tr><th>unit</th><th>loc</th><th>+ev</th><th>0ev</th><th>-ev</th><th>div</th><th>strength</th></tr>
<tr><td>1</td><td>1</td><td>5</td><td>2</td><td>5</td><td>1</td><td>22</td></tr>
<tr><td>2</td><td>1</td><td>0</td><td>2</td><td>4</td><td>1</td><td>18</td></tr>
<tr><td>2</td><td>2</td><td>5</td><td>0</td><td>1</td><td>1</td><td>4</td></tr>
<tr><td>3</td><td>1</td><td>2</td><td>2</td><td>0</td><td>1</td><td>2</td></tr>
<tr><td>3</td><td>2</td><td>2</td><td>0</td><td>2</td><td>1</td><td>3</td></tr>
<tr><td>3</td><td>3</td><td>1</td><td>0</td><td>3</td><td>1</td><td>12</td></tr>
<tr><td>4</td><td>1</td><td>0</td><td>1</td><td>2</td><td>1</td><td>9</td></tr>
<tr><td>4</td><td>2</td><td>3</td><td>0</td><td>0</td><td>1</td><td>0</td></tr>
<tr><td>4</td><td>3</td><td>0</td><td>1</td><td>2</td><td>1</td><td>9</td></tr>
<tr><td>4</td><td>4</td><td>2</td><td>0</td><td>1</td><td>1</td><td>4</td></tr>
<tr><td>5</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>5</td></tr>
<tr><td>5</td><td>2</td><td>1</td><td>1</td><td>1</td><td>0</td><td>5</td></tr>
<tr><td>5</td><td>3</td><td>1</td><td>1</td><td>1</td><td>0</td><td>5</td></tr>
<tr><td>5</td><td>4</td><td>2</td><td>0</td><td>1</td><td>0</td><td>4</td></tr>
<tr><td>5</td><td>5</td><td>1</td><td>0</td><td>2</td><td>0</td><td>8</td></tr>
</table>
</td></tr>
<tr><td>Select ‘best’ clock</td><td>
<table>
<tr><td>4</td><td>2</td><td>3</td><td>0</td><td>0</td><td>1</td><td>0</td></tr>
</table>
</td></tr>
</table>

Figure 2.2: Flowchart of the model of Povel and Essens (1985)

hypothetical clock (Povel and Essens 1985 p.419), whereas other models (Longuet-Higgins and Lee 1984) attempt to do this. Nevertheless, the model is corroborated by experimental data and makes psychologically relevant distinctions Grube and Griffiths (2009).

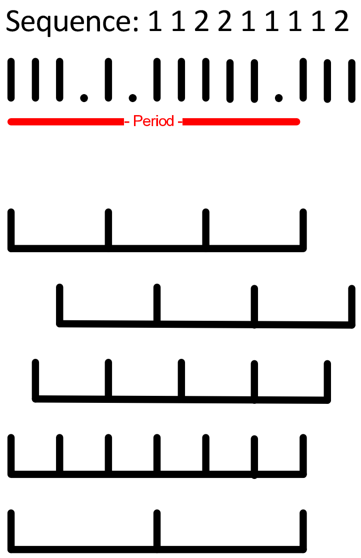
Analysing six different sets of IOIs, and the corresponding circular permutations[1], in a total of 605 stimuli, the authors were able to divide the sequences into 9 different categories of clock induction strength:

- Category 1. Best clock is induced by accented elements only;
- Categories 2, 3 and 4. Best clock is induced by accented and unaccented elements (1, 2 and 3, respectively);
- Categories 5, 6 and 7. The ticks coincide once with a silence ($-ev$) and depending on the category with zero, one or two unaccented events;
- Categorie 8. The ticks coincide once with a silence ($-ev$) and three unaccented events. No sequences in the present stimuli were found in this category;
- Category 9. The best clock for these sequences has two ticks coinciding with a silence ($-ev$).

The set of permutations of [1 1 1 1 2 3 3 4] has sequences in all the categories except for number 8.

The theoretical framework of the model is extended by relating the induction process to the need for encoding temporal information in a time sequence. By superimposing a clock, the subject is creating a hierarchical structure that can be better represented by

[1]Circular permutations are all the permutations of events that generate sequences which are not repeated by shifting a particular sequence in time.



Clock	Unit	Loc.	Code
1	4	1	1-1-2, 2-1-1, 1-1-2
2	4	3	S_2, S_4 2-1-1
3	3	2	1-2, 2-1, S_3, 2-1
4	2	1	S_2, E, E, S_2, S_2, E
5	6	1	1-1-2-2, 1,1,1,2

Figure 2.3: The coding of pattern [1 1 2 1 1 1 1 2] in terms of different clocks presented in Povel and Essens (1985).

the brain. The efficiency of the coding is inverse to the number of symbols represented in a clock unit. Three coding rules apply (Fig. 2.3):

- If a clock unit is subdivided into equal intervals, this is described in the code by means of the symbol S and a subscript indicating the number of intervals in that unit;
- If a clock unit is empty, this is indicated with the symbol E;
- If a clock unit is subdivided in intervals of unequal length, no reduction is possible and the subintervals are described by indicating which proportion of the clock unit the take.

This coding scheme is central to the idea of complexity of a rhythmic sequence which is further developed on Shmulevich and Povel (2000) and explained in Sec. 2.3.3.

Povel and Essens (1985) go on to devise three experimental setups to validate this theory. In the first experiment, 35 stimuli were selected from permutations of the intervals

[1 1 1 1 1 2 2 3 4], corresponding to five sequences from each of the first seven previously defined categories. Subjects were asked to reproduce the sequences four times after listening to them for as long as they wanted. The results, supported by the variance analysis, show that the number of times the sequences need to be presented to the listeners, before they are able to reproduce them, is dependent on the categories. Analysis of mean deviations (ms) from accurate reproduction of the sequences in this experiment, shows similar evidence for increased deviation according to the category.

A second experiment was designed to test the clock induction, by adding a low pitched clock track to 20 of the previous sequences. Only sequences from categories 1 to 4 were used, to avoid confusion when a clock track event coincided with a silence in the sequence. This time subjects found it even simpler to reproduce the sequences and category significance was also observed.

Finally, the third experiment tested for the hypothesis of the perceptual coding scheme. This consisted on presenting the subjects with identical sequences where different clocks tracks were added. The sequences had a length of 12 and were presented twice: one with a clock track with isochronous ticks of length 3 and another with a clock with length 4. The question was to whether the subjects would be able to recognise the high pitched sequences as similar. In 9 out of 10 times the answer was negative, meaning that the induced clock actively changes the perception of temporal sequence and the hypothesis of perceptual coding according to the internal clock receives positive support. For a more recent study in the effects of temporal encoding see also Grube and Griffiths (2009).

This study was highly influential to line of research of this thesis as it pertains to aspects of rhythm perception and production, directly related to complexity of rhythm. In the next section we will take a look at another study that explored the idea of measuring complexity.

2.3.2 Syncopation according to (Longuet-Higgins and Lee 1984)

This model is in many ways related to the Povel and Essens (1985) and has been influential to many music researchers (Smith and Honing 2006; Fitch and Rosenfeld 2007; Thul and Toussaint 2008). The model specifically attempts to quantify syncopation as a relation between the rhythmic surface and an underlying metric grid.

In Longuet-Higgins and Lee (1984 p.425) the authors propose a perceptual model based on the analysis of relative durations, stating that “even when there are no words, when the notes are of indefinite pitch, and when the performance is devoid of accent, phrasing or rubato: even in such impoverished condition the listener may still arrive at a rhythmic interpretation of the passage based solely on the relative duration of the notes” and continue “In this article we [...] consider the criteria that might lead a listener to favor a particular rhythmic interpretation of given a sequence of notes”.

This distinction between the metrical structure and the surface rhythm arises from the fact that there is an underlying ambiguity present in a rhythmic sequence, with multiple interpretations depending on the listener perception of the metric structure, in the same line of approach as Povel and Essens (1985) and Lerdahl and Jackendoff (1983).

The article follows on explaining how subjects generate a particular interpretation of a rhythmic sequence by minimizing the amount of syncopation induced by the sequences using a generative approach. This is similar to the work of Lindblom and Sundberg (1972) using folk songs from Sweden, although the present approach focuses solely on the relation between the rhythmic surface and metre.

This model generates tree structures, as in Chomsky’s syntactic structures (Chomsky 1959). The root node corresponds to the bar level, and the nonterminal nodes correspond to lower levels of the metrical hierarchy.

In order to understand the concept of syncopation it is necessary to define the “weight” of a note, or rest, with respect for the metrical hierarchy: "The weight of a given note or rest is the level of the highest metrical unit that initiates" (Longuet-Higgins and Lee 1984). Arbitrarily it is defined that the weight of the corresponding highest level is 0 and any other level is $n-1$, being n the level of the "parent" unit.

Syncopation, and syncopation strength, is thus defined by Longuet-Higgins and Lee (1984) as follows:

- If R is a rest or a tied note, and N is the next sounded note before R, and the weight of N is no greater than the weight of R, then the pair (N,R) is said to constitute a syncopation. The "strength" of the syncopation is the weight of R minus the weight of N.

Finally, if a parsing algorithm exists that successfully delivers an interpretation of the musical passage this interpretation is called a "regular passage", or a sequence of bars where the least syncopated version of the metric is used.

2.3.3 Measuring complexity (Shmulevich and Povel 2000)

Povel and Essens (1985) provided a way of analysing the rhythmic elements of a musical work in terms of the cognitive effort necessary for perception and production of those elements. Their work on the internal clock induction strength of a rhythmic sequence, related the cognitive complexity of the sequences with the complexity of an encoding scheme. Later works have tried to address the complexity of rhythmic sequences in formal terms (Tanguiane 1993; Pressing 1998; Manaris et al. 2005).

Shmulevich and Povel (2000) have proposed a rhythmic complexity measure based on the earlier work of Povel and Essens (1985) named the Povel-Shmulevich measure (PS-Measure).

In Shmulevich and Povel (2000), the PS-Measure is compared against two different

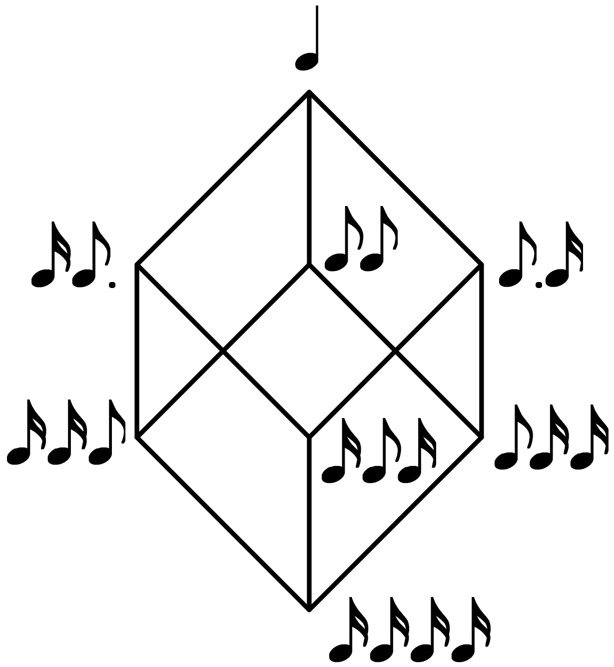


Figure 2.4: Elaboration of a quarter note according to Tanguiane (1993).

models of rhythm complexity, namely the Tanguiane measure presented in the book of Tanguiane (1993) and the Lempel-Ziv measure (Lempel and Ziv 1976), used for the evaluation of computational complexity of sequences.

The Tanguiane measure is hierarchical in nature and uses the idea that a rhythmic pattern can be described in terms of more simple patterns, simultaneously at different levels. The simpler patterns are based in the concept of *elaboration* defined in Mont-Reynaud and Goldstein (1985). The elaboration of a quarter note can be seen on Fig. 2.4.

The quarter note is elaborated, or transformed, into two note patterns on the immediate lower level, which in turn are further elaborated into 3 note patterns on next level and to a four note pattern on the lowest level. The patterns are interconnected by a line drawn between the levels. The patterns which do not correspond to elaborations of any other patterns are called *root patterns*. The root patterns are used to determine the complexity of the sequences.

x... x	Empty(E)
x.x. x	Equally subdivided (S_2)
x..x x	Unequally subdivided (U)
.x.. x	Starting with silence (N)

Table 2.1: Four types of segments as in Shmulevich and Povel (2000).

Another measure used for comparison in the Shmulevich and Povel (2000) study, is the string complexity measure developed by Lempel and Ziv (1976). This algorithm proceeds from left to right in the sequence and analyses new substrings, increasing by 1 the complexity value each time a new substring is found.

Although the measure captures hierarchical information, and redundancy, which are important features of rhythm, it does not regard any perceptual coding mechanism associated to it. This means that some important perceptual features might not be captured by the measure and some information embedded in the sequence might not be important to the human listener.

The PS-Measure "should be a combination of the induction strength of the best clock on one hand", provided by the C-score from Eq. 2.1, "and the efficiency of coding the rhythm on the other" (Shmulevich and Povel 2000 p.64). The existence of a hierarchical coding scheme presumably complies with *Gestalt* rules, such as the simplicity principle, in which a sensory input is encoded in the simplest possible way.

To accomplish this objective the authors propose a new equation for complexity adding a new term D which reflects the efficiency of the code:

$$D = \sum_{i=1}^{n} c_i + m \cdot d_5 \tag{2.2}$$

The coding complexity is computed by associating different weights to each type of segment using $d_1, \ldots, d_4$ to correspond to E, S_2, U, N (Tab. 2.1) and d_5 for repetitions

λ	W	d_1	d_2	d_3	d_4	d_5
0.2223	1.1695	0.0235	1.2722	1.2955	0.0736	0.7931

Table 2.2: Estimated parameters for the PS-Measure.

of segments.

Finally the complexity value P is given by the weighted sum of the C and D values:

$$P = \lambda \cdot C + (1 - \lambda) \cdot D \tag{2.3}$$

The free parameters of this measure were determined according to the study of Essens (1995). In this study 20 subjects were asked to judge the complexity of 24 rhythmic patterns on a scale of 1 to 5. The parameters of the PS-Measure were optimized to increase the correlation of the PS-Measure with the judged complexity leading to the parameters in Tab. 2.2.

2.4 Limits of rhythm perception

The human ability of synchronisation and entrainment is one of the most fundamental facts in music perception and music evolution studies (Fitch 2006; Patel 2007). There are some animals that display rhythmic behaviour, such as birds, gorillas, chimpanzees, whales and seals (Fitch 2005), but none is capable of entrainment (synchronisation) to a rhythmic pulse.

Humans are able to do so within certain boundaries. The shortest interval that we are capable of hearing or performing is about 100 ms (IOI) whereas the upper limit, set by the capability to hierarchically integrate successive events, lies around 5 to 6 seconds (London 2004). The capability of estimating an interval without subdividing, is set at around 2 seconds (or 30 bpm)(Handel 1989). The human ability to group isochronous elements in two (duples) or thee (triples), also known as subjective rhythmization, requires intervals between pulse to lay between the values of 115ms and 1580ms (Bolton 1894). For an overview of these limits please refer to London (2004).

2.4.1 Tempo

Tempo, also referred as *Tactus*, is a musical concept describing the pace of the internal clock. Tempo pertains to the number of isochronous pulses fitting in one minute thus being measured in beats per minute (bpm). Pieces perceived to be fast usually have a *tempo* with high bpm value and slow pieces usually have a low bpm value. But as the clock is hierarchical in nature, this might not always be the case. *Tempo* can be variable to accommodate for different musical effects within the musical piece.

2.5 Metre

Metre is defined as a repeating pattern of accented and unaccented beats. Metre in music is associated with the notion of cycle. We also know that metre is fundamental in dance, enabling many dancers to coordinate their steps.

Gjerdingen (1989) considers “meter as a mode of attending”, meaning that the psychological superposition of a metric grid enables the listener to actively engage in the perception of time sequences. This view is also shared by London (2004). The metrical grid represents points in time which possess an accent - a metrical accent. This approach most specifically established by Lerdahl and Jackendoff (1983 p.17-18), who have made a distinction between metrical accents - “any beat that is relatively strong in its metrical context”, structural accents - “caused by melodic/harmonic point of gravity in a phrase or section”, phenomenal accents - “give emphasis or stress to a moment in the musical flow, such as *sforzandi*, sudden changes in dynamics or timbre, long notes, leaps ... and so forth.”

2.5.1 Metric well-formedness

Musicologists have since long tried to theorise about the admissibility of some metric units. In the nineteen century, music theories were arguing whether a five beat measure was admissible. These accounts, more than describing human behaviour or musical ability fulfilled a prescriptive role. London cites several examples such as Hauptmann who claimed that measures with five or seven beats were simply inconceivable (London 2004 p.69,70).

With the publication of the "Generative Theory of Tonal Music" (GTTM), Lerdahl and Jackendoff made a clear distinction between the metric structure of a piece of music and its musical surface divided into groups. They postulate the existence of metrical well-formedness rules (MWFR) (Lerdahl and Jackendoff 1983 p.69-72).

- MWFR 1: Every attack point [on the musical surface] must be associated with a beat at the malles metric level present at that point in the piece;
- MWFR 2: Every beat at a given level of music must also be a beat a all smaller levels present at that point in the piece;
- MWFR 3: At each metrical level, strong beats are spaced either two or three beats apart;
- MWFR 4: The tactus and immediately larger metrical levels must consist of beats equally spaced throughout the piece. At subtactus metrical levels, weak beats must be equally spaced between the surrounding strong beats.

Along with the MWFR, there are also metrical preference rules (MPR), which are aimed at sorting the ambiguity of parsing rhythmic groups with relation to metrical structure in the analysis of the musical piece:

- MPR 1 (Paralelism): Where two or more groups or parts of groups can be construed as parallel, they preferably receive parallel metrical structure;
- MPR 2 (Strong beat early): Weakly prefer a metrical structure in which the strongest beat in a group appears early in that group.

MWFR 1 and 2 are considered by Lerdahl and Jackendoff to be universal, whereas MWFR 3 and 4 are style specifics, although these last two tend to be valid for most of western classical music, which is the object of the GTTM.

London (2004) proposes a set of universal metrical constraints which are meant to describe both Western and non-Western musical practice. The graphic representation of the metric grid is a circle containing N equidistant points corresponding to the basic beats, as well as geometric figures inside the circle connecting the points of higher levels of accentuation. This graphic representation is drawn according to the following set of well-formedness constraints (WFCs) stated in London (2004 p.72):

- WFC 1: The IOIs between the time point on the N-cycle must be categorically equivalent. That is, they must be nominally isochronous and must be at least $\approx$100 ms;
- WFC 2: Each cycle - the N-cycle and all subcycles - must be continuous, that is, they must form a closed loop;
- WFC 3: The N-cycle and all subcycles must begin and end at the same temporal location, they must all be in phase;
- WFC 4: The N-cycle and all subcycles must all span the same amount of time, that is, all cumulative periods must be equivalent. The maximum span for any cycle may not be greater than $\approx$ 5 seconds;

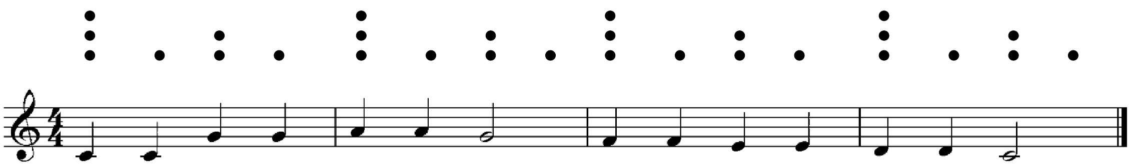


Figure 2.5: Twinkle twinkle little star.

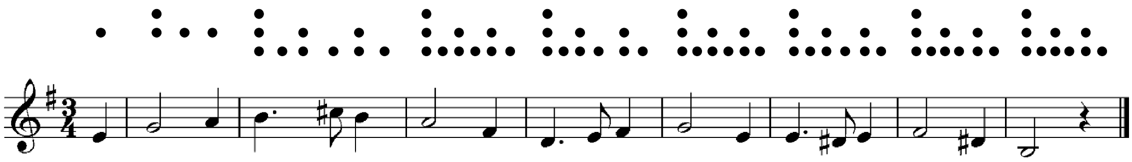


Figure 2.6: The first phrase of the tune *Greensleeves*.

- WFC 5:Each subcycle must connect nonadjacent time points on the next lowest cycle. For example, each successive segment of the the beat cycle must skip over at least one time-point on the N-cycle.

2.5.2 Hierarchical nature

Although the theory of the hierarchical nature of metrical accents was firmly established by Lerdahl and Jackendoff (1983), this notion was already acknowledged by Cooper and Meyer (1963) and was initially developed by Komar (1971) and Yeston (1976).

Inspired by the notation of poetic metres Lerdahl and Jackendoff (1983) have created a special notation that focuses on time instants (Figs 2.5 and 2.6).

2.5.3 Regular metres

A metre is said to be regular when the metric hierarchical structure of a piece presents a constant number of subdivisions of each beat at on every level of the hierarchy. The tune *Greensleeves* is built upon a regular 3 beat pattern (Fig. 2.7).

2.5.4 Irregular meters

When metrical accents, or beats, in the intermediate level of the metrical hierarchical are not evenly spaced, we are in the presence of what is called an irregular metre (Fig.

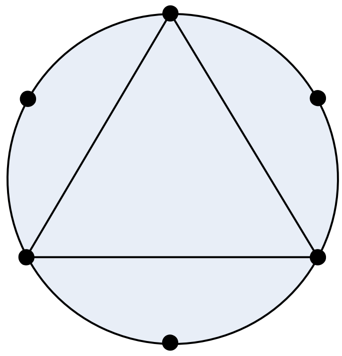


Figure 2.7: Regular 3 beat pattern used to describe the metre contained in the tune *Greensleeves* (Fig. 2.6).

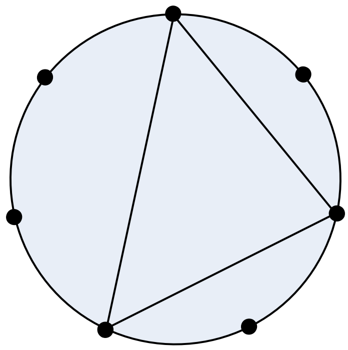


Figure 2.8: Irregular 3 beat pattern contained in a 7-cycle is found in traditional music from Greece, Turkestan, Bulgaria, and Northern Sudan (Arom 2004).

2.8).

Although the great majority of western music is built upon regular rhythms, one cannot say that irregular forms are uncommon, even in the western culture. Examples in 5/4 metre are easily found in both classical, jazz and pop music, such as the second movement of the 6th Symphony "Patheíique" of Tchaikovsky, or "Take Five" by David Brubeck, "Money" by the Pink Floyd, or the song "15 Step" by the band *Radiohead*.

These metres are quite widespread within traditional music of Central Europe. to the point that many cultures make them the distinctive character of their music. In traditional music from Greece, Macedonia, Serbia, Croatia, Bosnia and Romenia, irregular metres are ubiquitous, often supporting dance. In Bulgaria, metres are often so irregular up to the point of having cycles of 34 beats with 15 accents (Bulgarian Necklace) (Demaine et al. 2009)

When referring to the Stravinsky's "The Rite of Spring" as well as to some of its own compositions, Bela Bartok Bartok (1938 p.538) has said: "It is astonishing how helpless orchestral musicians were, not so long ago, when presented with such rhythms. They had become so accustomed to hand-organ-[hurdy-gurdy]-like symmetrical rhythms that they could not grasp these rhythms at all, which were so unfamiliar to them, yet so very natural."

Most of these influences may come from middle-eastern music, as some of these countries were part of the Ottoman empire until the 19th century. Turkish, Arabic and Persian music have strong percussive components and irregular metre is widely present.

Also indian music contains irregular modes of metric accent. The *Tālas* are musical rhythms which have a corresponding phonetic notation and are so important that are normally named in the heading of a piece, along with the *Rāga*, its melodic counterpart. Sachs (1953) states that "A *tālas* like *tīn*, which has 8/8 [...], is not a square product of two halves or four quarters, but rather the sum of 4+2+2. Or 8/8 can be organized in the ubiquitous patter of the Grecian dochmiac: 3+3+2."

2.5.5 Polyrhythms

In Sub-Saharan African countries rhythm plays a central role in music and this music is often connected to social functions (Chernoff 1979). Much of the music found is polyrhythmic, this meaning that several monodic lines with apparently different metric organisation are superimposed to create a interweaving pattern of beats. Arom (2004) describes the polyrhythmic nature of traditional music of the central african countries, and creates a taxonomy of the musical structure based in 4 different arrangements of structural elements:

- strict polyrhythmics - the superposition of two or more rhythmic figures, each of which is so articulated that its constituent elements (accents tone colour, and attacks) are interspersed among those of the others so as to create an interwoven effect;

- polyphony produced by hocket - is based on the interweaving, interlocking and overlapping of several rhythm figures which are tiered on different pitch heights in a fully defined scalar system. In Central Africa this is accomplished by using wind instruments;

- polyphony produced by melodic instruments - melodic instruments played two-handed to produce melodically and rhythmically different parts simultaneously;

- vocal polyphony - superposition of two or more melodically divergent lines with different rhythmic articulations.

2.6 Complexity

Complexity is in itself a term that escapes a simple definition. Longuet-Higgins and Lee (1984) state in their study that a human listener will tend to parse rhythms in order to minimise syncopation. Other more recent studies in music perception have found evidence that rhythmic complexity is associated to syncopation (Fitch and Rosenfeld 2007; Ladinig 2009).

In the study conducted in this thesis, complexity of the rhythmic units is rated objectively with an explicit function of the musical representation, either by experimentation or based on existing studies in the literature. In chapter 5 a complexity estimate is developed that attempts to capture quantitatively the amount of operations of division and splicing that a rhythm is subjected to in the system. In the study presented in chapter 6, the system used the psychologically more relevant measurement of complexity provided in the study by (Shmulevich and Povel 2000).

Machado and Cardoso (1998) propose that aesthetic judgements results from the balance between two forms of complexity. The first type is related to the difficulty of processing information by the brain, which causes aesthetic rating to decrease with complexity. The second type concerns the intrinsic complexity of a work of art in terms of information theory, which increases its aesthetic value. As an example, it is pointed out that fractal images present interesting aesthetic qualities because they are easy to process due to the self-similarity properties, and at the same time visually complex. Later, these ideas were also applied to music, relating the aesthetic qualities of a piece to the distribution of its musical features according to the Zipf-Mandelbrot law (Manaris et al. 2005).

There is yet another way of looking at complexity, insofar as it can be seen as subjective quality dependent on previous exposure to music material. Heyduk (1975 p.84) has pointed out that "By employing the additional assumption that experience with an event reduces its psychological complexity, predictions may also be made about the nature of preference changes with continued exposure to a musical selection". Based on the theory of optimal complexity by Walker (1970), Heyduk (1975 p.89) proposes that "behaviour is a product of a relationship between a situational factor (psychological complexity) and a parameter of the individual (optimal complexity level)". This means that complexity would be determined not only by fixed perception of complexity but also would decrease with repeated exposure. I am inclined to think this to be the case, as repeated exposure might fine tune the mental representation of information, therefore decreasing the level of perceived complexity. For the purpose of this study, I have considered complexity to be an inherent property of the stimulus, according to literature review, and I prefer to tackle the problem of exposure from an independent perspective.

2.7 Rhythm similarity

Computers find it simple to discriminate if something is equal or different, but the problem rises when there is the need to evaluate if something is similar (Minsky 1988). The necessity of similarity measures concerns many areas of music research, specially rhythm perception and production (Gabrielsson 1973), music information retrieval systems (Hewlett and Selfridge-Field 2005), automatic rhythm transcription of human-performed music to MIDI protocol (Takeda et al. 2003), evaluation of copyright issues, and evolutionary music (Miranda 2004).

On the side of the abstract models, interesting results were achieved using the *Levenshtein* distance, also called *edit distance*. This is a popular method for measuring similarity between strings of text of arbitrary length. The algorithm counts the number of insertions, deletions and substitutions necessary to change one string into another, being this number the measure of similarity between the sequences. Orpen and Huron (1992) have applied this distance to measure melodic, rhythmic and harmonic similarity in Bach chorales. Mongeau and Sankoff (1990) provided a method which can be seen as an extension of the previous. Instead of considering that each transformation to the sequence contributes with the value of one to the distance, each transformation contributes with a weighted value sensitive to the kind of musical differences which are to be measured.

In A-Rhythm, the system developed in the context of this research, the measurement of similarity plays an important role in the analysis of the generated rhythms. Similarity between repertoires of rhythms defines cultural proximity and enables the composition of the elements provided by the generative system. Sec. 5.1.2 presents a rhythm similarity measure that was devised in the context of this research.

2.8 Rhythmic ability

Music practice is a fundamental element of the daily activity of a music performer. Practice is defined in Cayne (1990) as “repeated performance or systematic exercise for the purpose of learning or acquiring proficiency” and Barry and Hallam (2002) expand on the goals of practice as “to acquire, develop and maintain aspects of technique, learn new music, memorize music for performance, develop interpretation, and prepare for performance. A key purpose of practice is to enable complex physical, cognitive, and musical skills to be performed fluently with relatively little conscious control, freeing cognitive processing capacity for higher order processing (eg. communicating interpretation)”.

Research shows, that in order to perform technically difficult pieces, musicians have to segment music into smaller passages and the more complex the music is, the smaller are the sections (Barry and Hallam 2002).

In the system presented in this thesis practice time influences the ability to play complex rhythmic sequences. This notion will be further elaborated in the description of the system in chapter 6.

2.9 Summary

This chapter presented a review on the field of research of rhythm. Several models of rhythm perception were explained with special focus on the works of Povel and Essens (1985), Shmulevich and Povel (2000) and Longuet-Higgins and Lee (1984). The last two models have a special impact in the notion of rhythmic complexity, which is fundamental to the topic of this thesis. Furthermore, an overview of the notion of metre was presented with special attention to rules of metric well-formedness, metric hierarchy, regular and irregular metres, and polyrhythms. Finally, the notion of complexity was summarised in Sec. 2.6 and a brief overview of concept of music ability was presented.

Chapter 3

Methods in computer-assisted composition

Chapter 2 reviewed aspects of rhythm perception and production. This chapter introduces computational methods used in algorithmic music composition and computational modelling of musical behaviour. These methodologies have played an important role in the study of computational creativity.

3.1 Algorithmic music composition

Algorithmic music composition comes to play much earlier than modern computers. Around 1670, Samuel Pepy composed *Musarhithmica Mirafica* and in 1791 Mozart composed the *Musikalisches Wuerfelspiel* (Schwanauer and Levitt 1993), explained above. These pieces follow an algorithm but they were not made with the help of a machine.

In 1957 Lejaren Hiller composed the first music piece, the Illiac Suite for string quartet, where the computer played an actual role in the composition process (Hiller and Isaacson 1959). Hiller commented on the reverse of the album: “I observed that if we could program a computer to simulate a ’walk’ through, say, ordinary space, we could also simulate a ’walk’ through a grid defined to represent musical elements such as pitch, rhythmic durations, and timbre choices” (Hiller et al. 1967). The Illiac Suite is a four movement suite for string quartet exploring a broad range of composition techniques, ranging from sixteen century counterpoint to twelve note composition.

A detailed discussion on algorithmic composition falls outside the scope of this thesis. The reader is referred to text by Roads (1996), Miranda (2001) and Nierhaus (2009).

Algorithmic music composition has been one of the areas of focus in the study of computational creativity. Computational creativity is an artificial intelligence field concerning the study of creativity with computers. As Saunders and Gero put it, "artificial creativity is a computational approach to studying creative behaviour using closed-world simulations of social creative systems. In a similar way to Artificial Life (A-Life), the aim of artificial creativity is to provide insights into the nature of *creativity-as-it-is* by studying *creativity-as-it-could-be*" (Saunders and Gero 2002 p.80). The production, communication and recording of creative ideas and artifacts is hence studied with the help of computational models of creativity.

The next section, presents the technologies involved in musical style modelling and briefly discusses how the computer can be used as means of composing new music.

3.2 Computational models for the study of composition in music

This section focuses on artificial intelligence methods that were used in studying the rules and processes derived from examples of the musical historical record. Once the parameters are set these methods also allow to compose music. Markov processes were used to build the grammar in the implementation of A-Rhythm (chapter 6), and therefore this methodology is explored in more detail.

3.2.1 Knowledge based approaches

The history of music and composition builds its knowledge on analysing pieces of music from a given period and establishing the rules and sets of constrains pertaining to that particular style.

One of the early examples of the use of computers to formalise rules of compositions is automatic species counterpoint program by Schottstaedt (1989). Based on the work *Gradus as Parnassum* of J. J. Fux, which presents a set of constraints easily coded in terms of IF..THEN statements, Schottstaedt developed a programm to find acceptable renditions of the species. In the words of Schottstaedt: “These attributes can easily be defined in such manner that a computer program can use them to find acceptable solutions to species counterpoint problems.”.

Knowledge based approaches, including constraint programming and case based reasoning, have been successfully applied to harmonisation. CHORAL (Ebcioglu 1992) uses a rule based approach to produce choral harmonisations in the style of J.S Bach. Tsang and Aitken (1991) use “constraint logic programming” and Pachet and Roy (1998) use a system of constraint satisfaction also for harmonisation. More recently, Strasheela (Anders 2009*b*,*a*), a constraint based harmonisation system, renders harmonisations in the style of different composers. For surveys on the subject of constraint based harmonisations, the reader is referred to Pachet and Roy (2001), Machado (2006) and Anders and Miranda (2011).

Case based reasoning was applied to musical composition in the SICOM system (Pereira et al. 1997). This system uses a database of musical works and corresponding analysis under a tree structure, to generate new pieces of music according to a distance measure established in the work of Macedo et al. (1996).

3.2.2 Markov models

“The most elementary grammars which, with a finite amount of apparatus, will generate an infinite number of sentences, are those based on a familiar conception of language as a particularly simple type of information source, namely, a finite-state Markov process”

(Chomsky 1956).

Andrey Markov, a russian mathematician, studied the sequence of 20,000 characters in Pushkin's poem *Eugeny Onegin*. From this analysis he discovered that the stationary vowel probability in the poem is $p = 0.432$, that the probability of a vowel following a vowel is $p1 = 0.128$, and that the probability of a vowel following a consonant is $p2 = 0.663$ (Basharin et al. (2003) cited in Nierhaus (2009 p.67)).

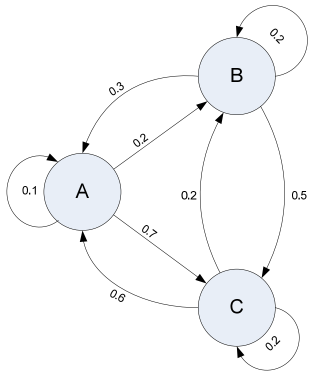
A Markov chain is a mathematical description of a stochastic process, in which a system transits between a finite set of states with probability values associated to the transition between any two states. Associated to Markov chains is the Markov property, namely the absence of memory in the stochastic process. This property asserts that, in a Markov process, the probability that a state transits to another, depends only on the current state and not on a past one.

Markov chains have been very useful in describing processes where a stochastic process depends on a time variable such as in information theory, telecommunications, robotics, and even the page rank system in internet search engines (Page 2001). It has also been one of the most popular strategies for algorithmic composition (Roads 1996 p.878). A Markov chain can be described by a *state-transition diagram* (Fig. 3.1a) or by a *state-transition matrix* (Fig. 3.1b).

The sum of the probabilities of the arcs that leave a particular state in the *state-transition diagram* is equal to 1, as well as the corresponding sum of probabilities in each line of the *state-transition matrix*.

A Markov chain whose matrix representation has non-zero entries immediately on either side of the main diagonal, and zeros everywhere else, constitutes a random walk process (Miranda 2001).

In order to capture long-term dependencies in a time sequence, it is possible to increase



	A	B	C
A	0.1	0.2	0.7
B	0.3	0.2	0.5
C	0.6	0.2	0.2

Figure 3.1: a) Markov *state-transition diagram*: the probability of transitions discriminated in the arcs. b) Markov *state-transition matrix*: the transition arcs going out from each state correspond to the lines of the matrix and the transition arcs going into each state correspond to the columns of the matrix.

the *order* of the Markov chain. In this case the process is not memoryless and the Markov property is not observed. The *order* of the chain indicates the number of prior states that are taken into consideration in computing the probabilities of transition to the next state of the system. Two special cases are the *zero-order* chain, which has fixed probabilities of transition, independent of the current state, and a *first-order* chain which corresponds to the process depicted in Fig. 3.1 (Roads 1996). Increasing the order of the chain can give, to some degree, coherency and increasing structure in a piece, with several layers emerging from the process. Problems remain as to which order to use, or the lack of control in the process, leading to crude splicing of sequences (Roads 1996 p.879).

Initially this method was used in algorithmic composition to create melodies, using pitches for each of the states. One way to expand its capabilities is to use other parameters, such as rhythm, dynamics, articulation, or even timbre.

Another possibility is to use short excerpts of music as states. Long before the formalisation of the theory behind Markov chains, in 1791, W. A. Mozart composed a musical piece named *Musikalisches Wuerfelspiel*, or musical dice-game (Schwanauer and Levitt

1993), corresponding to a Markov process. In Mozart's musical piece, he provides a set of short lines of music that can be randomly interconnected by throwing a dice and finding the proper connection in a table. This produces possibly infinite renditions of waltzes.

Markov models have long been used in musical applications as they provide a computationally inexpensive way of encoding sequential information. Music systems that have used Markov stochastic processes are the *Cybernetic composer* (Ames and Domino 1992) and the *Jam Factory* (Zicarelli 1987). Conklin and Witten (1995) use a technique to extract knowledge from examples and then analysing it using information theory (Shannon 1950).

3.2.3 Grammars

Grammars have their origins in linguistics but have been widely used in computer science and music.

With the publication of *Syntactic Structures* by Chomsky (1957) a new way of looking into linguistics was born. With this approach a difference between deep structure and surface structure in language was developed. Surface structure is the structure revealed in the phrase in the way it is produced, and deep structure corresponds to the unique semantic relations a phrase is able to express. The way a surface structure is converted into the deep structure, is by means of transformational rules. For a description of formal grammars in linguistics, see Santorini and Kroch (2007).

It is important to understand the notion of a formal grammar. A formal grammar is a way of representing hierarchical relationships (Roads 1996). It is a collection of descriptive or prescriptive rules for analysing or generating sequences of symbols (Miranda 2001).

Grammars can be divided in 4 types (Fig. 3.2): in terms of their restrictions, formal

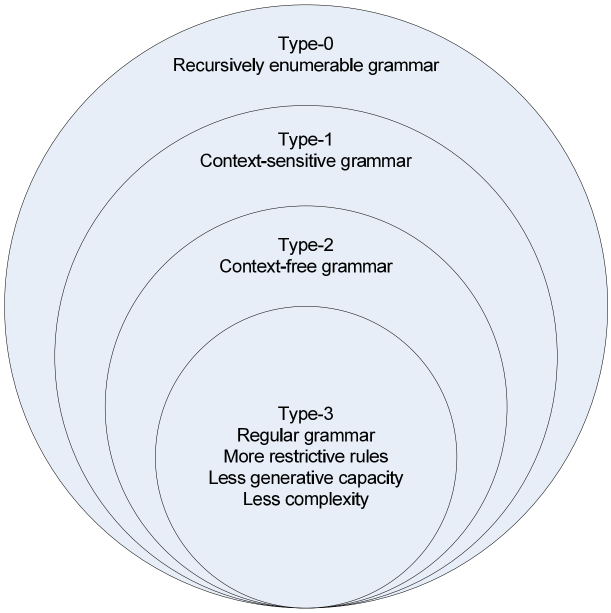


Figure 3.2: Grammar hierarchy

language produced, type of machine that generates it, generative capacity and complexity (Nierhaus 2009). A Type-0 grammar is called a *recursively enumerable grammar* or *partially decidable* and can be generated by a non-deterministic Turing machine. It has no restrictions on both sides of the generative rules giving it a very high generative capability. The complexity of the grammar, related to the time for computation of the final state, may go up to infinite.

Conversely a Type-3 grammar is called a *regular grammar*, it has the most restrictive rules for production, its language is generated by a *finite state automaton* (FSA), it admits less sequences than the lower order grammars, and its complexity is linear. The expressiveness of a Markov model is equal to one of a Type-3 grammar (Nierhaus 2009 p.91) and a Type-3 grammar can also be represented by a Markov model (Nierhaus 2009 p.90) .

A grammar can be described by a four element structure, or 4-tuple (N, T, P, S), where:

- N is a set of non-terminal nodes;

- T is a set of terminal nodes;
- P is a set of production rules;
- S is a starting symbol.

The generative rules, or production rules, are notated with a substitution relationship between nodes, of the form $\alpha \rightarrow \beta$, for the general case. The rule $\alpha \rightarrow \beta$ reads "α is rewritten as β". The symbols α and β can be replaced by sequences of letters, in which a capitalized letter means non-terminal node and a non-capitalized letter means a terminal node.

Roads (1996 p.893) points out that there are some problems associated to the use of grammars: music has innumerable non-hierarchical ways of being parsed. Another problem pointed by Minsky (1981) is that grammars describe only the structural nature of a language, and not its meaning, making it not suitable for music. In this respect Miranda has pointed out that "meaning in music is a much harder issue to deal with than meaning in language, but most musicians would surely agree that it is preferable to leave this issue unsolved anyway" (Miranda 2001 p.76).

3.2.4 Transition networks

Augmented transition networks (ATN), used in the works of David Cope, present a related formalism to *generative grammars*. The software "Experiments in Music Intelligence" (EMI) by Cope 1991; 1992*a*; 1992*c*; 1992*b*; 1996; 1997 analyses musical works by composers such as Bach, Mozart, Chopin or Rachmaninoff and produces new musical works by recombining elements of the music style embedded in the original pieces.

This system produced extremely compelling results, leading music students into thinking that they were in the presence of works from the actual composers when they were confronted with their works.

The process can be simply summarised in three stages:

- Deconstruction (analyse and separate into parts);
- Signatures (commonality - retain the elements which signifies style);
- Compatibility (recombinancy - recombine into new works).

The signatures are musical blocks extracted using "pattern matching" which reveal the style of the composer. The signatures are then recombined using Transition Networks (TN) which group the blocks into a coherent result, via a hierarchical process resembling a grammar, performed in a "top down approach". Later on, EMI uses a dictionary of local gestures to give explicit meaning to the symbols created by the hierarchy.

3.2.5 Multi-agent systems

Multi-agent systems are a class of Artificial Intelligence algorithms that aim to solve problems where distributed solutions are better suitable than centralised ones. Agents can display several properties such as autonomy and interactivity and are normally, but not always, set in a virtual environment. Multi-agent systems generate simulations that run through discrete time steps and the states of the agents are generally updated simultaneously.

Multi-agent systems were used successfully in several music areas, such as composition (Gimenes 2009; Kirke et al. 2011) and modelling of expressive music performance (Miranda et al. 2010; Kirke 2011).

Chapter 4 presents an overview of systems where the multi-agent framework was used.

3.2.6 Evolutionary computer music

Evolutionary computation (EC) is able to explore large variable spaces using biological and social processes. This is specially interesting for music, as music itself touches in multiple layers of human perception and evolves as a result of the interplay between individuals. The result is a complex adaptive system where EC has had good grounds for research (Kim and Cho 2006; Horner and Goldberg 1991; Miranda 2003; Burton and Vladimirova 1999; Bilotta et al. 2002)

In a paper by Miranda (Miranda 2004) evolutionary music is defined subdivided in three different approaches:

- **Engineering approach** - concerns the use of evolutionary methods for sound design. The most widely used techniques in this approach are genetic algorithms, genetic programming and cellular automata. The analysis of this this approach is out of the scope of this thesis;
- **Creative approach** - concerns the use of evolutionary methods for producing new compositions;
- **Musicological approach** - concerns the study of existing music compositions or the process of the evolution of music using evolutionary computation.

Creative approach

Traditional AI methods for the generation of computer music are very good at extracting regularities in music composition and imitating style, but they are somehow more problematic when it comes to create new pieces of music. Some algorithms use abstract rules and mathematical formula which, without human intervention, create music where the listener rapidly looses interest for being too predictable or too random. The EC paradigm gave a new perspective to computer aided composition. The relation to nature and the structural elements involved in life like forms seems to resonate on the listener with a stronger potential than purely abstract models. The key factors seems to be the processes of self-organisations and emergence of behaviour. From the EC area, successful music systems were developed using genetic algorithms (Gartland-Jones and Copley 2003; Biles 2007), genetic programming (Johanson and Poli 1998; Burton and Vladimirova 1999), neural networks (Mozer 1994), cellular automata (Miranda 2001; Brown 2005), swarm music (Blackwell 2007) and multi-agent systems (Kirke 2011; Murray-Rust 2008). Although there are still some problems, related to the difficulty in specifying fitness functions, or the need for human supervision in the evaluation, these processes have already proved to be useful in helping composers developing their ideas.

Musicological approach

This particular area in computer music is relatively new when compared to the applications of EC to algorithmic music composition. EC is today being applied more and more to the study of biological evolution and to models of cultural evolution in many complex systems such as language. Music is not strange to this reality and some interesting results have been achieved. Typically, this enquiry involves the creation of simulations in a multi-agent system framework, where aspects of human or animal music behaviour is tested with the help of the computers. Existing creative and musicological models are expanded in chapter 4. For general resources in music and evolutionary computation see Miranda and Biles (2007) and Romero and Machado (2007).

The next section deals with the study of models of music preference that will be useful for the system developed in this thesis.

3.2.7 Computational models of music preference

To the field of AI, the research works of Wundt and, years later, of Berlyne, have set the ground for the development of artificially generated art. Saunders (2002) has created the notion of “curious agents” mainly using the relationship between liking and the property of novelty to endow agents with a rating system that helps navigating in the space variables of artistic artifacts. In the case of scientific enquiry on the question of the evolution of music, Bown (2008) has shown that it is possible to evolve a system of hedonic values using music as mechanism to mediate social interaction between individuals.

3.3 Summary

In this chapter, several technical methodologies pertaining to computer-assisted composition were analysed. Special focus was given to knowledge based approaches, Markov models, grammars and transition networks. Further, two approaches for evolutionary methods in music composition and musicology were characterised. The system developed in this thesis employs evolutionary computation algorithms, namely the multi-agent framework and uses Markov processes as part of the internal representation of the agents. In the next chapter, several multi-agent systems pertaining to the scope of this thesis are reviewed, which were designed with either creative or musicological goals.

Chapter 4

Composition and musicological computer music systems

Chap. 3 introduced a variety of computational methodologies traditionally used in algorithmic music composition and computational modelling of musical behaviour. This chapter focuses on particular models that directly contributed to the outline of this thesis.

Due to the chosen framework for this study, special attention is paid to multi-agent systems. Although two of the sections have no relation to multi-agent systems - they are devoted to rhythmic generative systems (Sec. 4.1.8 and 4.1.9) - they are explained here in more detail, as they give some valuable insights to the models used by A-Rhythm.

The reasons for using multi-agent systems are enumerated as follows:

- Music is inherently social activity and multi-agent systems help to conceptualise ideas in terms of social interactions;
- A bottom up approach to the generation of music potentially generates interesting and novel behaviours that will emerge from the observation of clusters of agents;
- The systems are scalable, meaning that one can focus on the level of the small interactions or trends in a population;

- Autonomous interactive agents are a good framework for distributed knowledge, and this helps to understand the generative process in a modular way;
- Although multi-agent systems have been used in composition for long, their musicological applications are relatively new and present new avenues for research which cannot be explored otherwise.

As a composer I am interested in understanding the dynamics of change in society, and how this reflected in the complexity of the music created. In designing a artificial system to investigate this goal, there are some criteria which may contribute to answer the possible questions involved. These criteria help to define the specifications of such a system.

First, one has to understand the distinction between a system in which the computer will be generating material for compositions or if it will be generating examples to elucidate a particular musicological theory. There are several systems that are used for computer assisted composition, some other are designed to perform in real time with musicians, and yet other, fulfill the goal of providing musicological explanations or proofs in the field of the evolution of music. In most cases, the systems clearly lean either towards a more compositional goal or to a musicological one. In some cases this is not so obvious, for some systems with a musicological goal have sufficiently complex generative processes to make them interesting to generate compositions. Therefore, this distinction is reflected in two criteria for systems that fulfill the compositional or the musicological goals.

The next criterium concerns systems that were specifically designed to deal with rhythms. Most of the systems reviewed in this thesis, have rhythms as a property of the music material, but rhythm is not the object of any evolutionary algorithm or the characteristics of the rhythmic structure do not contribute to the evolutionary process. If rhythmic structure is used in the evolutionary process then this criterium is satisfied.

The study of the evolution of complexity in rhythm is an important topic of this research. Within the systems that deal with rhythm, some address the topic of rhythmic complexity and some consider rhythm as an important part of the musical material without bringing up the subject of complexity. The complexity criterium was satisfied whenever rhythmic complexity is discussed and/or measured.

In considering the study of music in culture, exposure to music material has a direct influence into the persistence of musical memes. Exposure is loosely defined as the amount of times a particular musical material is used in a given context. Although there is research in the area of music exposure (North and Hargreaves 2008), none of the existing systems try to analyse the musical material according to exposure. Yet I have considered an exposure criterium to be satisfied whenever the music material has some sort of reinforcement measure during the iterative process.

The design of multi-agent systems includes the choice of the process by which the agents interact. The occurrence of interactions between agents in a multi-agent system may be governed by some arbitrary algorithm, by a network of interaction, or, in the case of A-Life systems, by some form of spatial distribution of the agents. The virtual environment solves many problems regarding the choice of the interaction mode and provides an intuitive setting to the distribution of agents in societies. Therefore, a spatial criterium is included, which is satisfied whenever the agents are placed in an environment and are allowed to move in different directions.

Included in the compositional and musicological goals of this thesis is the extraction of rules evolved by an artificial system, through the process of self-organisation. The grammar criterium is satisfied when the system describes/evolves compositional rules.

In the studies of cultural origins (Lippo et al. 2006), special focus is given to the different forms of cultural transmission. In this thesis two types of transmission were analysed: horizontal transmission and vertical transmission. Horizontal transmission (HT) happens when music material is exchanged between agents of one generation. If one of the compared systems performs the analysis of this process, the corresponding criterium is satisfied. Vertical transmission (VT) happens when music material is exchanged between agents of different generations. I have considered that to be the case when the transmission occurs via social transmission, or when offspring genetically inherit the tunes of their parents, as it is the case with the systems used to analyse bird songs. In both cases, social vs. genetical, some sort of population renewal mechanism has to be in place. If one of the compared systems performs the analysis of this process, the vertical transmission criterium is satisfied.

Finally a criterium for genetic evolution is considered. This criterium is satisfied when the traits of the agents or the evolved material suffers changes via some form of genetic algorithm.

The set of ten criteria are summarised as follows:

Composition: If the system is designed for composition (yes/no);

Musicology: If the system is has musicological goals (yes/no);

Rhythm: If the system is focused on rhythm (yes/no);

Complexity: If it uses perceptual models of complexity (yes/no);

Exposure: If it considers exposure of sequences (yes/no);

Spatial: If the agents have a representation in space (yes/no);

Grammar: If it considers the emergence of grammar (yes/no);

Horizontal transmission (HT): If agents exchange music material within a generation (yes/no);

Vertical transmission (VT): If music material is exchanged across generations (yes/no);

Genetic evolution: If variation on the genetical level is observed (yes/no);

The next sections describe 14 systems related to this research and will be analysed according to the criteria chosen in the process of designing A-Rhythm.

4.1 Composition and performance systems

The systems described in this section were designed mainly for computer assisted composition or performance. The Bol processor and the RGeme are systems that have considered musicological aspects in the modelling but have evolved mainly to compositional tools. Most of the systems have features related to evolutionary computation, but will focus on particular elements of music, or use specific technology.

4.1.1 GenJam

Gen Jam is a system for jazz music improvisation with evolutionary algorithms first developed by Al Biles in 1993 (Biles 1994) and in the words of its creator "has evolved from a proof-of-concept demonstration to a viable improvization agent that maintains a regular performance schedule as a soloist in the author's virtual quintet" (Biles 2007). The principle involved is of an Interactive Genetic Algorithm (IGA) which is a type of genetic algorithm where the fitness function is not easily chosen or is not known a priori. In the case of IGAs the selection of individuals in a population is accomplished by human intervention.

Although IGAs have been applied in many aesthetic contexts (Haggerty 1991; Sims 1993; Dawkins 1996), in music, the length of musical excerpts creates additional problems. As Biles (1994) points out, the human intervention constitutes a "very narrow bottleneck and is, in fact, the limiting factor on population sizes, number of generations, and size of any generation gap". He coined the term "fitness bottleneck" to refer to this problem.

GenJam's architecture contains many built in modules to support a musical performance. Namely a script for the performance named *Choruses*, a rhythmic section provided by *Band in a Box*, a chord progression module, a *Head Sequence* module which is a MIDI file scripting composed parts of the piece, and another module with MIDI parameters.

The evolutionary nature of GenJam is observed in both the *Measure* and *Phrase* modules. These use a pool of musical measures and phrases which are genetically mutated and crossed-over to provide improvisation material to be played in a traditional jazz setting. The system has a supervised learning stage where the training is made by the performer by listening to the measures and pressing 'g' when a measure is perceived to be good or 'b' when the measure is perceived to be bad. Each new tune corresponds to a new evolved generation or soloist.

For performance the system is able to engage in different modes such as:

- trading fours or eights;
- performing collective improvisation;
- interbreeding human measures from the head and the human's solo chorus with measures in the measure population.

4.1.2 Swarm music: SwarmM

One of the focus of artificial life is collective animal behaviour. Examples of collective animal behaviour are fish schools, bird flocks, nest building ants, ungulate herds, and other groups of animals moving together as a coordinating unity. This kind of behaviour arises in nature by self-organisation, without central control, and its form of intelligence transcends the abilities of the individuals in the group (Camazine et al. 2001). This swarming coordinated behaviour is believed to be used to strengthen group cohesion and avoid predators.

Artificial swarms are created by programming autonomous agents with rules derived from theoretical biological models. These swarms are then visualised at more abstract level in spatio-temporal model.

Swarming models can be divided in bio-swarms, simulation swarms, or social swarms (Blackwell 2007). The former, are as accurate as possible and aim at scientific goals of hypothesis development and testing, the second kind are for visualisation purposes in real-time aesthetic applications, and the later use information networks for optimisation purposes.

The simulation swarms, developed by computer graphics specialists for aesthetic purposes, have been used in the film and computer games industries. (Reynolds 1987; Burton 1992; Allers and Minkoff 1994)

The first actual computational model of a swarm was *Boids* designed by Reynolds (1987). This model is based on three simple steering behaviours describing how an individual boid moves based on the positions and velocities of its nearby flockmates:

Cohesion: Steer to move toward the average position of local flockmates;

Separation: Steer to avoid crowding local flockmates;

Alignment: Steer towards the average heading of local flockmates.

The boids only reacted to flockmates within a certain neighbourhood constituting this a model of limited perception.

Looking at a musical composition from a dynamic systems point of view involves considering the different musical notes - with pitch, loudness, duration, onset time - as individuals that wander in a musical dimensional space moving closer to each other or drifting apart. This process can be extended to other levels of musical description such as melody, rhythm, articulation, sound spectrum or orchestration, depending on the mapping of the space, or even considering a whole object that combines all different elements into a unity such as an artificial player. The individuals may converge, diverge, merge to create a new individual, or separate into more elemental particles, depending on the attracting forces involved.

Self-organisation will be present in the sense that the different elements combine to form a whole: different melodic lines are combined to produce harmonies; different rhythmic streams produce polyrhythmic woven pieces; sound spectrum may evolve to more complex textures; or different artificial players may create a fully fledged piece of music.

Tim Blackwell (Blackwell 2007) has developed several interactive systems using swarms and self-organisation, to be used in live music settings.

Blackwell suggests that musicians, when improvising, are letting self-organisation govern the process of creation. This means that musical structure is built as a bottom up process rather than a pre-planed top-down approach. The author has proposed "a model of interaction based on stigmergy leading to the design and implementation of swarm music systems that can interact with people in an improvised setting".

Stigmergy is a biological process that involves animals communicating with others over long time scales by modifying the environment. Ants use the pheromone trace from previous elements of the colony to find their way between home and a food source. People leaving sticky notes for others to read or signs left by mountaineers can be considered a form of human stigmergy.

An autonomous music system capable of human-compatible performance is called a live algorithm (Blackwell and Young 2005). All live algorithms have a model based on three build blocks. An external sound is processed by a module P that converts the acoustic information to internal parameters treated as an image p. Then there is an ideas engine f, operating an internal space H, that is guided, but not determined, by the inputs p. A final block of synthesis Q, re-interprets the internal state as a new external sound. In the particular case of the systems of this chapter, f is a swarm that provides the spatio-temporal pattern defining H.

In the next part of this section, I will review the concepts involved in swarm music with the focus on the systems *Swarm Music*, *Swarm Granulator* and *Swarm Techtiles*. These systems respond to human improvisers and generate new musical material arising from patterns in an artificial swarm (Blackwell 2007, 2008). According to Xenakis (1989) and Roads (2001) , music structure can be observed in different perceptual time scales, from which Blackwell selects 4 of them, possibly overlapping, which relate to the scope of the aforementioned systems:

Macro: This level encompasses the musical piece, measured in minutes, hours, or even in extreme cases, days;

Meso: Division of the music piece into movements, sections, phrases or groups;

Mini or Sound Object: This level corresponds to the note level, or a texture, with sounds ranging from few hundreds of milliseconds to several seconds;

Micro: This is the level which reaches the threshold of auditory perception with sound particles going to the tenths of milliseconds in duration.

The different systems developed by Blackwell (2007) can be categorised by the different time scales in which the parameters influence the music. *Swarm Music* works on the mini and meso levels, *Swarm Granulator* on the micro level and *Swarm Techtiles* operates at the sample and micro levels.

The movement of the swarm through the internal space H of musical notes in *Swarm Music* corresponds to the arrangement of notes into melodies.

Swarm Granulator, developed by Blackwell together with Michael Young (Blackwell and Young 2004) uses granular synthesis parameterized in the space H resulting in sound textures synthesised by Q.

Swarm Techtiles (Blackwell and Jefferies 2005) combines elements from the social swarm models with the swarm granulation technology used in *Swarm Granulator*. The particles fly above a landscape, searching for optimal regions as quantified by evaluation of an objective function which measures local image texture. The texture is a 2D picture which is itself drawn by the audio inputs in a process known as “woven sound”.

The virtual swarms used by the aforementioned systems, communicate with the musicians in an analogue process to stigmergy and act as a system of “ideas generator” for the musicians involved.

While developing these systems, the authors were concerned with the principle of “transparency”. This means that the listener should be able to understand the evolution of the music according to the real-time visualisation of the Swarm.

Other systems have also used the combination of swarms and music or audio, by choosing different configurations for the three modules (P, H and Q) taking advantage of the dynamical nature of the swarms.

Spector and Klein (2002) developed *SwarmEvolveMusic* where the interaction between the individuals and the virtual environment triggers musical events. The music signals have influence in the feeding of the individuals and different species are assigned to differen timbres of instruments.

Wilson (2008) developed several *Supercollider*[1] classes to use granular swarms in an arbitrary setting of loudspeakers. This system was tested with the Birmingham ElectroAcoustic Sound Theatre (BEAST), a loudspeaker array featuring more than 100 channels.

Unemi and Bisig (2005) have developed a system of 3D swarming particles, which controls audio parameters. The user interacts with the swarms with gestures recorded by a camera, which influence the swarming behaviour.

4.1.3 Rhythms as emergent structures: Pachet

François Pachet (2000) developed a multi-agent system for rhythm performance in which rhythm is seen as a musical form, emerging from repeated interaction between several rhythmic agents. The agents are furnished with perception and production modules. The perception module parses the rhythms from the overall performance and extracts information regarding the beat structure and beat emphasis. The production module uses a set of transformation rules to create the next rhythms in performance.

The transformation rules are seen as applied genetic operators to create variations of the previous sequences. These rules can emphasise the beats, create syncopation, and add or remove random notes.

This is a very simple type of system and it can produce complex rhythms by using the defined transformations. Despite this fact, it differs significantly from A-Rhythm as there is no environment or interaction between agents' behaviours.

4.1.4 Kinetic Engine: KinEn

The Kinetic Engine, developed by Arne Eigenfeldt (2009), is an evolutionary system for composition and performance, used to create complex polyphonic rhythms. It has been used as an instrument in group performance, as an installation, as a controller

[1]SuperCollider is an open source environment and programming language for real time audio synthesis and algorithmic composition.

for a robot with 12 articulated arms and as a composition tool to generate percussion pieces. The system is multi-agent based and uses first order Markov chains evolving via genetic algorithms. Rhythms are individuals in a population and are generated in real time. The fitness is evaluated using musical examples provided by the user. In this way, it is possible to overcome the problems created by the *fitness bottleneck* (Biles 1994). The examples are initially loaded from MIDI files and the music material is analysed to extract the relevant features which will be used as a fitness function. During the evolutionary process the features of the offspring will be analysed and become the new parameters for the fitness. The genetic algorithm uses only mutation as a genetic operator, as the reproduction of the agents is asexual.

The polyphonic structure is controlled by a central agent named the Player which culls the elements of the population to be played. The Player is guided by syncopation and density parameters provided by the user and by similarity evaluations between individuals of the population. The repeated culling of the same individual exponentially decreases its possibility of being picked again to be played in the future.

4.1.5 Note sequences: NoteSq

Gong et al. (2005) have developed a composition system based on the notion of cultural evolution. It follows a similar approach to Miranda (2002*a*), who uses interactive games to build repertoires of cultural objects related to music. The music is constituted of melodic lines of 32 notes divided in 8 bars, taken from 22 notes from the C major scale. The time signature is 4/4, and there are no pauses. There are also no differences between timbres and durations.

The agents assume the roles of composers and critics. The composers generate random sequences at the beginning and then play the music in their memories to other agents during communication. The critics will evaluate the played sequences according to simple aesthetic functions which evaluate the “pleasantness” of the played sequence. The “pleasantness” is a measure based on 4 types of musical preference rules: the beat rule, which prefers particular kinds of accent patterns; the sequence rule, which prefer particular melodic contours; the jump rule, which prefers jumps of less than an octave between consecutive notes; and the interval rule that prefers major and minor thirds.

In order to avoid evolutionary process stagnation and to find a constant balance between continuity and novelty, two of the mentioned rules go through a fuzzy evaluation function, in which the agents will prefer high scores of “pleasantness” in only 3 to 5 bars, of a total of 8 bars that constitute a music.

The overall evaluation value of one piece of music is the weighted sum of 4 evaluation values, based on the 4 types of musical preference rules.

In case a tune scores low in the measure, the agent will modify the sequence and send it back to the player agent. The player evaluates the returned song and if it scores higher, the new song is retained and the old is discarded.

The agents keep a historical record of past interactions. Positive interactions between agents will reinforce the bonds between them and will increase the probability of interaction in the future. In case the first agent finds the received song not so pleasant as the original tune, then the bond is not reinforced.

Finally, the evolved songs scored higher in the preference of the agents than the initial random ones.

4.1.6 VirtuaLatin: VLatin

VirtuaLatin (Murray-Rust 2003; Murray-Rust et al. 2005) is an early development of *Musical Acts - Musical Agents* (MAMA) by David Murray-Rust which is based on speech act theory from Linguistics and Pragmatics (Murray-Rust et al. 2006; Murray-Rust 2008).

VirtuaLatin is a multi-agent system designed to play in a band of cuban Salsa music. The object of the work is primarily one intelligent agent, the *timbalero*, that develops improvisations supported by a set of pre-recorded rhythmic lines. The *timbalero* develops a structured representation of the rhythms and harmonies specific to latin music. This representation is inspired by the work of Lerdahl and Jackendoff (1983) on the General Theory of Tonal Music (GTTM), and incorporates specific domain knowledge.

The system was evaluated by human listeners in comparison to a human performer. The results show that the VirtuaLatin can trick a normal listener, although not the experts. MAMA (Murray-Rust 2008) has inherited the music representation devised for VirtuaLatin.

4.1.7 CinBalada: CinB

CinBalada is a multi-agent rhythmic lab for automatic creation of polyphonic rhythmic performances (Sampaio et al. 2008), which mixes elements from different music styles. The organisation of the agents follows a virtual circle of percussionists with horizontal negotiation of the rhythmic patterns to be played. The rhythmic material used by the system consisted of pre-recorded rhythmic patterns from different human cultures and categorised according to their original instruments.

There are four possible roles that the agents can take: *base*, *complementary base*, *solo* and *fill*. The roles depend on timbral characteristics of the instruments, so for instance a *Surdo*, which is an instrument quite rich in low frequencies, will be assigned the role of the *base*, but the roles can change during the performance. Before each bar is played, agents with the same role negotiate which rhythms will be played, according to an evaluation of the material that has been played before.

As *CinBalada* is dealing with polyphonic rhythms, the authors propose two measures, based on experts' advice, which assess the compatibility between the monophonic

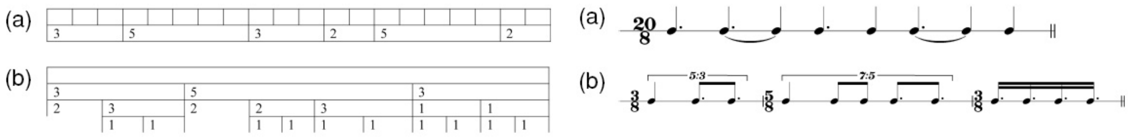


Figure 4.1: a) Additive rhythm; b) Divisive rhythm.

parts. The *measure of interplay* calculates the difference between the number of coinciding events (matches) and the number of events that do not coincide (mismatches) when the monophonic phrases are aligned in the metrical grid. The *measure of partial derivation* measures up to what extent one pattern derives from another by insertion of events.

The system was evaluated with listening tests where the users assessed the quality and diversity of the material when listening to 8 different instrumentation setups. In most of the experiments, *CinBalada* performed better than randomly arranged rhythms.

4.1.8 Additive and divisive rhythms: Nauert

Paul Nauert has developed computer assisted composition system to explore the rhythmic space (Nauert 2007). This system explores the difference between two different approaches in rhythmic generation, namely an *additive* approach vs. a *divisive* approach. The concern of this system was not "the ability of addition- or division-based models of rhythm to accurately describe existing musical traditions, but the value of these models as a basis for generating rhythms in an algorithmic-composition system".

An additive rhythm is a rhythmic sequence composed of several elements that are combined in an additive manner (Fig. 4.1a). A divisive rhythm is a rhythm that presupposes a hierarchical level of organisation within the rhythmic sequence, determined by the nature of the generative process (Fig. 4.1b).

OMTimePack is an OpenMusic library for creating and manipulating rhythms, structured as sequences of IOIs. The system contains a selection mechanism based in

Markov chain of order zero, one or two. The user specifies parameters for the Markov transition probabilities, the pool of possible durations and the target for the total duration of a sequence. A iterative generative process follows until the duration of the sequence is achieved. This generation process based in stochastically sequencing single events is purely an additive strategy.

Nauert (2007) describes a process of identifying the probabilities for the Markov matrix, consisting in analysing the relative frequencies of a given sequence. There is a random element introduced in the subsequent stochastic generative process, and new sequences may become much more complex to perform than the original sequence. Complexity is then discussed in terms of the quantization of the elements according to a chosen quantization grid. One way of reducing the complexity is to change the grid while trying to maintain distortion to a minimum. The amount of distortion that is allowed depends on the goals of the composer. According to the author, the main shortcoming of this process is not so much the generate and quantize process, but how to relate an additive generative process with a notation system where divisive principles play a significant role.

4.1.9 Bol Processor: BolProc

The Bol Processor (BP) was developed by Bernard Bel and applied to the research for the tabla music by Jim Kippen (Bel and Kippen 1992; Kippen and Bell 1992). Its rule sets are very similar to the formal grammars that are used in computer science to define machine-readable languages (Sec. 3). More recently, the Bol Processor 2 (BP2) was developed to be used as a program for music composition and improvisation with real-time MIDI and Csound output (Bel 2008).

The work has its origins in a grammatical description of the music of the tabla, a North Indian percussion instrument, provided by Jim Kippen (Kippen 1988). In the oral tradition, there are system of verbal symbols called *bols* used for transmission and occasionally for performance. Bols are onomatopoeic mnemonics that allow for drummers to remember and even practice the rhythmic sequence. *Dha*, *ti*, *ge*, *na*, *tirakita (trkt)*, *dhee*, *tee*, *ta*, *ke* are all examples of *bols*.

The original goal of the Bol processor was to trace the development of a formal-language representation for tabla music and to implement it in a computer system. The object of research is a form called *qa'ida*, the "theme and variations" form par excellence.

The variations obey rules of construction, but these rules are rarely expressed formally by traditional musicians. Instead, an implicit model is transmitted by means of sequences of positive instances of the "language". Negative instances composed by their students during the course of training are rejected or corrected.

The "active" element of the BP is an inference engine that uses a stochastic process to generate sentences derived from the grammar, and a membership algorithm to check whether or not a sentence entered into the editor is consistent with the grammar.

In the synthesis process, the BP is able either to enumerate all sentences of the language or to generate one sentence randomly.

In the BP2 the composer may instruct the software to generate well formed variations and then select those to be kept.

4.1.10 Ontomemetics: RGeme

Marcelo Gimenes developed a framework known as Ontomemetical Model of Music Evolution (OMME) which uses interacting autonomous agents with goals of exploring how human beings perceive, represent and create music. More generally, this framework is made to design interactive musical systems to study the evolution of musical style and explore the transmission of musical influence between humans and machines

(Gimenes et al. 2005, 2007; Gimenes 2009). This framework led to the creation of two systems for music composition and performance - RGeme (Rhythmic Meme Generator), which is an earlier version of the system, focusing on rhythms, and the interactive system iMe (Interactive Musical Environments).

Central to this approach is the notion of *meme*, developed by Dawkins, which is a basic unit of cultural transmission in the same way that genes, in biology, are units of genetic information. “Examples of memes are tunes, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm and eggs, so memes propagate in the meme pool by leaping from brain to brain through a process which, in the broad sense, can be called imitation”(Dawkins 1989).

The system comprises a perception module, that analyses the psychoacoustic features of a musical piece or performance (rhythm, contour, melodic leaps, etc.), which will be encoded into genotypes. Then, these genotypes are segmented into *memes*, and finally there is a recombination of the material to be used in a creative way. The agents also have a list of tasks that guides their ontogenic behaviour.

The design of the agents draws on the concepts of short-term memory (STM) and long-term memory (LTM), both explained in the musical context by Snyder (2001). The agentŠs STM is the simplest of the two and stores the *n* memes that were most recently received into the memory. The LTM is a series of *FeatureTables* (FTs) in which all the genotypes are stored according to their category. There is a system of weights representing relative importance of the genotypes stored in the feature table. These weights are reinforced if the genotypes are recognised in new heard material, otherwise they are decremented.

4.2 Musicology and evolution of language

In the collection of works on biomusicology and evolutionary history of music, *The Origins of Music* (Wallin et al. 2001), the single contribution containing computational modelling was that of Peter Todd (Todd 1999). Computational models of emergent animal communication systems based on bird songs have had further development since then (Sasahara and Ikegami 2003, 2004; Vallejo and Taylor 2005).

In this paper, Todd argues that although we can understand how music instruments have evolved by looking into archeological records (fossilised bone flutes), behavioural changes in music production and perception abilities are not so easily understandable by just looking into those records. The proposed solution tries to replicate the evolution process in the computer in the form of *evolutionary computer simulations*, building artificial environments where individuals create and possibly perceive musical signals. Then, one can shape selective forces and behavioural endowments of the artificial creatures and evaluate the results in an iterative evolutionary process. This allows for the generation of new hypothesis and testing of existing theories.

The scientific purpose is to find out why some species (birds, whales, dolphins, and humans), show signs of creativity in their communication. If the purpose of these species were only to communicate some form of meaning, the communication form would be expected to have a stable code in order to minimise the amount of errors in the information, as it happens with other species (vervet monkeys, bees) with genetically determined communication codes. Furthermore, the scientific approach is more oriented into evolving behaviours and systems of communication, than into creating particular instantiations of the communication process, artifacts, or specific pieces of music.

As a starting point, we have looked at existing systems designed for algorithmic composition, as these are aimed at creative goals. In another paper, Todd and Werner (1999) provide a review of evolutionary systems for music composition, where they stress the problems with the nature of the methods and the poor results obtained by most existing systems at the time.

The evolutionary process can be described into three different steps: generate, test and repeat. The first involves the generation of musical pieces to be played, the test phase involves the attribution of a fitness values to generated material, and the repeat phase involves some selective and reproductive criteria.

Focusing on the test phase, these systems can be divided into four different ways of evaluating the individual fitness: human critics, automated rule-based critics, learning critics, and coevolving critics.

Systems based in human critics are constrained by the *fitness bottleneck*, first characterised by Biles (1994) and explained in Sec. 4.1.1. Automated rule-based critics can present interesting results in a given rule context. The problems arise with limitations are set in the specification of the rules. On one hand, there are always new rules that can be built into a system but on the other hand too many rules end up lacking surprise in the output.

Typical learning systems are neural networks. The development of learning critics presuppose a database of existing music for the system to learn from existing examples. With neural networks it possible to keep some generalisation ability, which the strictness of rule based systems is not able to provide. Still the results are not great, due to limitations on the size of the training set. Generally, evolutionary algorithms will, on the generative side, explore the weaknesses of a static critic, finding shortcuts to obtain better results on the fitness side.

The solution envisaged by Todd is to use coevolving critics. Coevolution, in biology, refers to the change of a biological object, triggered by the change of a related object.

This pattern of change can be observed in different scales such as the protein level (micro level) and also in the process of arms races between different species (macro level) (Todd 1999).

This framework has two main advantages according to Todd: first it can produce diversity within a population at any one time - *synchronic diversity*; and second, it can generate diversity across time - *diachronic diversity*.

4.2.1 Evolving melodic birds: BirdSong

In order to study these phenomena, Werner and Todd (1997) designed a simulation framework to study the evolution of diversity in the context of birdsong. An artificial evolutionary system was designed, in which male birds generate songs, that are subject to the evaluation of coevolving females birds. Each male has genes that directly encode the notes of one song, with 32 notes selected from 24 pitches. Females possess a transition matrix of 24x24 cells which encodes the expectancy of the transitions between heard notes. There are three algorithms for the preferences of the females: *local transition preference*, *global transition preference* and *surprising preference*. According to each studied case, the females attribute a score to a particular song. In the first method, the score is computed by cumulative sum of the females' expectancy values, corresponding to the transitions in a song generated by a male. In the second method, the score is determined by computing a matrix of relative occurrences of transitions in a male song. This matrix is then compared with the expectancy matrix of the female. In the third method, females prefer songs that surprise them, and the score is computed by calculating a sum of the expected-minus-actual transition probability value for all the notes.

Each female listens to a courting choir (fixed number of males), and has one child in each generation, resulting from genetic mutation and crossover with the selected male.

This brings the population at about 50% over the carrying capacity (target population size), which is then reduced by one third to bring it to a constant size through the simulation. The male songs were initialised at random and the female preferences were initialised with transition matrices calculated from a folk song database. Several conditions were studied: the three preference modes (local, global and surprise), static vs. evolving preferences, and variable courting sizes.

When females had non coevolving preferences, evolution stagnated. Under coevolution, the *local transition preference* was shown to produce very little diversity under all conditions. Male songs usually converged to either a single pitch or alternation between two pitches. As expected, the *surprising preference* mode produced a greater diachronic diversity among generations. The *global transition preference* mode led to a rapid variation in the initial generations, which was expected for initial random songs, but eventually produced poor diachronic diversity. The choir sizes influenced the synchronic diversity with smaller male choir sizes producing more diversity of repertoires at a given time.

Some problems with this model were pointed by its authors, such as the absence of changes within one generation: in human music, and even in bird species with learning capabilities, preferences evolve within a generation (no horizontal transmission).

4.2.2 Handicap principle: Handicap

Given the prevalence of rhythm in communication between species, one may ask whether the rhythmical nature of these signals has some adaptive function. The hypothesis investigated by Eva van den Broek and Todd (van den Broek and Todd 2009, 2003) is that "rhythm may be used as a signal of an individual's underlying traits, and in particular, may indicate factors that are important in mate choice". The theoretical framework for this research relates to the work in sexual selection (Darwin 1871; Miller 1999) and the

		Expected	
		0	1
Heard	0	X1	X2
	1	X3	X4

Figure 4.2: Female preference table for judging male songs

handicap principle (Zahavi and Zahavi 1997).

The handicap principle refers to a selection mechanism dependent on signals received. A signal is not simply the product of observation by one individual, but instead is an information received from another individual, portraying some aspect of its traits which might be potentially important for both.

The handicap principle suggests that reliable signals must be costly to the signaler, costing the signaler in fitness, what could not be afforded by an individual with less of a particular trait. Famous examples of this costly displays are the stotting of the gazelle in the presence of a predator, and the peacock's long tail.

The hypothesis proposed by van den Broek and Todd is that rhythm in bird song may function as an indicator for the quality of the males in terms of neural activity and for the amount of noise at the neural level. In the framework created to address this hypothesis, simulations were devised where female birds are endowed with a perceptual mechanism to compare songs sung by male birds to a template they possess. The songs are encoded in strings of 10 bits.

The overall goal of the simulations is to see if the evolved templates of songs in the males and the template of expectations in the females will evolve into a higher rhythmic quality.

Females compare the songs according to a preference table for expectations (Fig. 4.2).

The quality of a male bird is modeled by a probability of error insertion at any given point in the singing template. This value is initialised randomly for each male agent.

In the first paper (van den Broek and Todd 2003), the authors have devised 20 psychologically plausible payoff matrices, and calculated the most-discriminating template. *Discriminability* is the gap between the best normalized score possible for each template (given a certain preference table), and the mean over all noisy performances of that template. This value is a measure of how well can a female discriminate, with this template, between low quality (one mistake on average) and high quality (no mistakes) males. Interestingly, for most of the preferences tables, the sequence that presents the most-discriminating sequence is the sequence "0101010101", which is rhythmic in character. This result suggests that rhythmic song templates appear to be the most useful type of signal for discriminating high quality males from low quality ones.

Then, it was studied whether the templates could be evolved for the preference table $X1X4 = \{1,0,0,1\}$. A simulation was run with a genetic algorithm with mutation and crossover, using a similar population dynamics as the previous models from Todd (Werner and Todd 1997; Todd 1999). In this case, each individual contains both a singing and a judging template, using only one of them depending on its sex. Crossover happens only between the two female- or male-associated templates, so that the templates used for judging cannot be mixed with those for singing. Females also have an error in judgement which is modeled by adding noise into the judgement process. The *rhythmicity* score is the average number of alternations between 1 and 0. The randomly expected mean value is 4.5, but the rhythmic sequences failed to evolve very far from this value, with a not so fast conversion to the mean, if the population was initialized with the highest value for the alternations.

In a more recent work, van den Broek and Todd (2009) decided to investigate the reasons why the evolutionary process was not able to overcome the drift in the earlier study. For this study, they used three different preference tables: symmetric $\{1,0,0,1\}$, asymmetric $\{0,-1,-1,1\}$ with a reward for expected notes, or arhythmic $\{0,-1,-0.5,1\}$, and divided the evolved templates into categories of rhythmicity. They also looked into songs with different lengths and changes in hereditary schemes.

The first conclusion is that only the symmetric scheme produced high values of rhythmicity. The asymetric scheme actually produced levels of rhythmicity under the mean, with templates looking more like *all zeros* or *all ones*. The remaining conditions suggest that "such regularly repetitive signals are also more likely to evolve in particular circumstances, including the use of simple symmetric preferences on the part of females, symmetric production mistakes (insertions and deletions of notes and pauses) on the part of males, and preferences and signals coded with the same genes".

4.2.3 Competition leads to cohesion: Bown

In the musicological domain, other avenues of research have been suggested to deal with the evolution of music. Bown (2008) explored the differences between competitive and cohesive approaches in the literature to construct a model that was implemented as a multi-agent system. Sexual selection is typically a competitive model (Miller 1999; Todd 1999; Zahavi and Zahavi 1997). Cohesive models of the evolution of music see music as means to foster group bonding (Dissanayake 1992), social interaction between parents and children (Cross 2003), rituals (Cross 2010), or synchronised choirs to attract mates (Merker 1999). One of the problems with the sexual selection is that sexually selected traits normally cause significative differences between males and females (ex. peacock's tail), and no evidence exists for such differences in humans. On the other hand the evolution of cohesion is difficult to initiate, as the individuals that adopt such behaviours would be outrun by others due to drift forces.

The thesis presented by Bown (2008) is that music is originally a maladaptation. It originally evolved as a competitive behaviour, helped by a runaway selection process, and eventually the process of social learning brought enough adaptive gains to be sustained

(Boyd and Richerson 1988). The concept of fitness determined by the environment would give space to a notion of social determination of fitness. Individuals listening to others would be under a form of *enchantment* and attribute status rewards to the players, if the style of the players scores high, according to a novelty measure. The style is a point in a multidimensional space.

The model proposed is exemplified by simulations, with a multi-agent system set in virtual ecological environment, to study the notion of social determination of fitness. The environment is said to be ecological in the sense that food grows with a growing rate and the availability of food constrains the life-cycle and reproductive ability of the agents. The agents pay status rewards to each other based on the evaluation of cultural artifacts produced by other agents. The status value is payed by the listener using a Wundt curve (Sec. 6.1) to determine its value, which rates the novelty of another agent's (player) style in relation to the his (listener) style. When a player is given a high status value it is allowed to feed from the environment.

The genetic variable evolved is the *susceptibility to enchantment* (SE) or the maximum amount of reward an agent will pay to another. This is the maximum value of the Wundt curve. When this value is 0 no reward will be given regardless of the novelty of the difference between the style of both agents.

It was not clear from the beginning whether SE would grow, insofar as when an agent pays status rewards to neighbouring agents this makes him less fit than its neighbours. This fact, makes it also plausible for the evolutionary process not to start, or to drift randomly.

Several other parameters are taken into account in order to study how the social determination of fitness could be influenced by other aspects, namely the learning rate of a style, the perception capabilities of the agents, and vertical vs. non-vertical transmission modes. The learning rate (L) influences how close the status are. At the end of each trial of status attribution, the agents change their style to be closer to the other agent's style. The perception capabilities (or limitations) of one agent is defined by the number of style parameters, from 1 to 10, an agent is able to perceive. The vertical vs. non-vertical transmission modes describe the situation where the children inherit style parameters or not.

The study of the simulations suggests that increasing SE, in the case of vertical transmission, is actually the result of a selection pressure favouring higher SE individuals and not the result of a process of random drift. Another hint from the simulation is that higher values of the learning rate actually hinder the growth of SE. The number of perceived parameters is directly affected by the learning rate.

4.2.4 Emergence of intonations: Miranda

Eduardo Miranda approaches the problem of the evolution of music from a social perspective (Miranda 2002*b*; Miranda et al. 2003; Miranda 2008). It differs from the musicological previously explained models in the sense that it does not concern evolution in genetical terms. The goal of the system is "to demonstrate that a small community of interactive distributed agents furnished with appropriate motor, auditory and cognitive skills can evolve a shared repertoire of melodies from scratch, after a period of spontaneous creation, adjustment and memory reinforcement"(Miranda 2002*a*; Miranda et al. 2003). The agents build repertoires of intonations and foster social bonds with other agents by trying to imitate each others vocalizations.

The agents are composed of a perceptual mechanism, a memory, a production mechanism and an enacting script. The imitation process comprises several processes in the brain of the agents, namely auditory analysis, an auditory-motor association process, encoding of the tunes in a motor control map, and audio synthesis via physical modeling of the vocal tract, vocal chords and lung pressure.

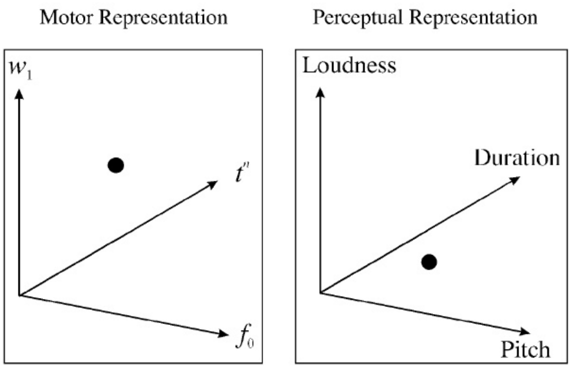


Figure 4.3: Motor and perceptual representation of one sound in the memory of one agent; w_1 stands for amplitude of voicing source, f_0 fundamental frequency in Hertz and t^n is the trigger time in seconds.

The tunes are represented in a dual form, one representation being the perceptual map resulting of the analysis, and another the motor map associated to the synthesis (Fig. 4.3). There are two reasons to use a dual representation. The first is that it accounts for the fact that in reality, there is no one-to-one mapping between perceptual sound features and their acoustic correlates (in this case, synthesis parameters). Second, it gives some flexibility in choosing different instruments for the synthesis of the perceptual representations.

Although the mapping of the parameters between the two perceptual maps is definitely more complex, the physical variables with a stronger contribution for the perceptual result were chosen and a simple one-to-one mapping was adopted.

There are two roles in the imitation game, the agent player (aPl) and the agent imitator (aIm). The enacting script guides the interaction between two agents:

- aPl produces tune p_1 randomly chosen;
- aIm analyses tune p_1, searches for similar tune (p_2) and sings it back to aPl;
- aPl analyses p_2 and looks for most similar p_n giving *satisfactory* feedback when p_2 is the closest existence tune to p_1;

- if feedback is *satisfactory* aIm will reinforce the tune p_2 and will change its motor control in order to sound a little bit more like p1;
- if feedback is *unsatisfactory*, aIm will either leave the tune p_2 unchanged, in case this tune has had strong reinforcement in the past, or change the motor controls to make it closer to p_1, in case the tune has had little reinforcement in the past. Then, aIm will generate a set of 10 random intonations and choose to keep the one that is closer to p_1;
- finally, both agents conduct 3 internal operations (*merge*, *forget*, and *create*) to clean the repertoire and add new intonations from randomly chosen parameters.

After the convergence of the repertoires, the perceptual representation should be identical, although the motor representation can be different from agent to agent.

In Miranda et al. (2003), the simulations are extended to the emergence of syntactic structures in connection to a semantic space of meanings. Combinations of sounds (riffs) are associated to emotions and combinations of emotions are associated to moods. This model is related to Iterated Learning Model (ILM) developed by Kirby (2002).

In a more recent paper (Miranda 2008), the system was implemented in robots and the emergence of the intonations was verified in a real environment.

4.3 Systems comparison

The systems described above can be categorised according to the criteria defined in the beginning of this chapter. Table 4.1 shows the categorisation of the described systems. Both the Bol processor and RGeme stand on the border between musicological and composition systems, and it is in the same way that the A-Rhythm should be looked at.

System	Composition	Musicology	Rhythm	Complexity	Exposure	Spatial	Grammar	HT	VT	Genetic
GenJam	x				x		x		x	x
SwarmM	x					x		x		
Pachet	x		x					x		x
KinEng	x		x	x	x		x	x	x	x
NoteSq	x				x			x		
VLatin	x		x				x			
CinB	x		x					x		
Nauert	x		x	x			x			
BolProc	x	x	x				x			
RGeme	x	x	x		x		x	x	x	x
BirdSong		x							x	x
Handicap		x	x						x	x
Bown		x			x	x		x	x	x
Miranda	x	x			x		x	x		

Table 4.1: Criteria satisfied by the evaluated systems.

4.4 Towards a system that fulfills the criteria

GenJam does not fit in a multi-agent system framework, but its development gave very important insights to the evolutionary music field due to the problem of evaluation.

Swarm music relates to the work of this thesis in the sense that contributes to the definition of a live algorithm. It is also one of the few systems where spatial representation plays a key role.

The system created by Pachet, differs significantly from A-Rhythm as there is no interaction between agents' behaviours.

The Kinetic Engine software has had multiple generations with very interesting results in performance and composition. It differs from A-Rhythm in the sense that the agents do not possess the same aesthetic judgement properties (A-Rhythm uses Wundt curves - and the evolution of the sequences proceeds in a different way. The Kinetic Engine does not consider the spatial location of the agents.

The system created by Gong et al. (2005) is interesting in the sense that applies the notion of games used in linguistics in order to generate musical material and study the formation of bonds between the agents. This approach could potentially be used with musicological goals. The system does not consider the evolution of the critics, being the evaluation functions essentially static. Also, neither rhythmic complexity nor location of the agents is taken into account.

VirtuaLatin is interesting as it is a clear application of the use of grammars in the context of rhythm. As only one agent interacts with musical content, it works well as a specific composition system, does not take advantage of the possibilities provided by the multi-agent systems approach.

CinBalada provides some interesting measures concerning polyphonic interaction of voices that guide the selection of rhythms from a database of musical rhythms for real-time improvisations. Rhythmic complexity is not addressed and no evolutionary process takes place in the system.

The algorithms from Nauert show the possibilities and potential problems of using divisive methods, which are important in the context of A-Rhythm, and using Markov processes in rhythmic generation. As it was mentioned before, it is neither a multi-agent system nor an evolutionary process.

The Bol processor was one of the first computational systems to have both musicological and compositional goals. The system is quite detailed in its expressive capabilities and it is supported by empirical research. It is not a multi-agents system and it does not contain any evolutionary algorithm.

RGeme and iMe were very influential in the development of A-Rhythm. The notion of cultural transmission is strongly present in the framework of this thesis. Contrary to A-Rhythm, these systems do not address specific issues related to rhythmic complexity, and the spatial location of the agents is not considered.

The system of Werner and Todd differs significantly from A-Rhythm in the sense that preferences evolve within a generation (no horizontal transmission) and it does not deal with rhythm nor it evolves a grammar.

The system from van den Broek and Todd deals with rhythm and addresses the most basic questions of the emergence of rhythmic behaviour. On the other hand, the system does not consider horizontal transmission and does not consider a grammar.

The simulations from Bown are not domain specific, so they do not address rhythm structure. They address the problem of social determination of fitness of individuals by evolving a system of style preference. Again this is a more basic level of interaction than A-Rhythm, and as departure level, A-Rhythm assumes the preexistence of a system of preferences which can evolve in terms of how novel and how complex the agents prefer their music material.

The mimetic work of Miranda, unlike A-Rhythm, does not use rhythmic structures and the intonations are only horizontally transmitted. One of the interesting features of this system is the fact that the agents can have quite different internal representations and still communicate effectively. Similarly to A-Rhythm this research also looks into the extraction of a musical grammar as a result of the process of interaction.

Although this thesis explores some of the avenues of research that were opened by the musicological approaches, A-Rhythm focuses more on the compositional possibilities inspired by research on evolutionary models for music.

4.5 Summary

This chapter revises 14 systems according to 10 established criteria to contextualize the development of A-Rhythm. The criteria were chosen based on the more salient features required for a compositional and musicological evolutionary system. Most of the features have been individually addressed in composition and musicological contexts. The approach pursued by A-Rhythm can be related to many other existing systems in the field of music and artificial intelligence, but the analysis of the systems in this chapter should be sufficient to put A-Rhythm into context.

Chapter 5

A-rhythm: imitation games

A-Rhythm is multi-agent system, designed to study the evolution of rhythms in an artificial environment. These rhythms are produced and selected by artificial intelligence agents in a cultural transmission setting. The first implementation of the system investigates different modes of transmission of the rhythmic units.

5.1 The agents

An autonomous agent is an intelligent agent (software entity) that implements a set of operations with some degree of independence or autonomy within an (artificial) environment. An agent is able to sense the status of that environment (including other agents), and acts on it, over time, in pursuit of its own goals. The environment affects agents' behaviour, and agents can also act on it in order to change it.

The agents are identical to each other and the number of agents in a group may vary. The agents move in a virtual 2D space and they normally interact in pairs (Figure 5.1). Essentially, the agents interact by playing rhythmic sequences to each other, with the objective of collectively developing repertoires of rhythms. At each round, all the agents that gathered in pairs are assigned with one of two different roles: the player and the listener. At each interaction, the agents may perform operations on the rhythms that they play to each other, depending on the interaction algorithm and on the status of the emerging repertoire. The agents are provided with a memory to store the emerging

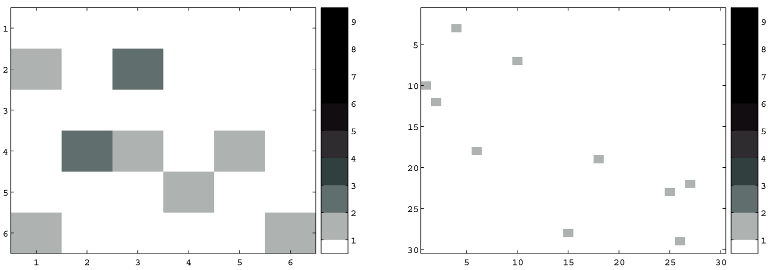


Figure 5.1: 2D virtual worlds with different sizes holding 10 agents. A darker colour indicates an interacting group of agents (this will be clarified in due course).

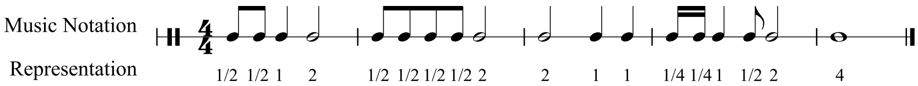


Figure 5.2: Standard music notation of a rhythmic sequence and its corresponding interonset representation.

rhythms and other associated information.

A fundamental characteristic of human beings is that we are able to perceive, and more importantly, to produce an isochronous pulse (Handel 1989). Moreover, humans show a preference for rhythms composed of integer ratios of the basic isochronous pulse (Drake and Bertrand 2001). Therefore, rhythms are represented here as interonset intervals in terms of small integer ratios of an isochronous pulse (Fig. 5.2).

5.1.1 Transformations of rhythms

At the core of the mechanism by which the agents develop rhythmic sequences is a set of basic transformation operations. These operations enable the agents to generate new rhythmic sequences and change the rhythmic sequences that they learn as the result of the interactions with other agents. The transformation operations are as follows:

- Divide a rhythmic figure by two (see Fig. 5.3a);
- Merge two rhythmic figures (see Fig. 5.3b);
- Add one element to the sequence (see Fig. 5.3c);

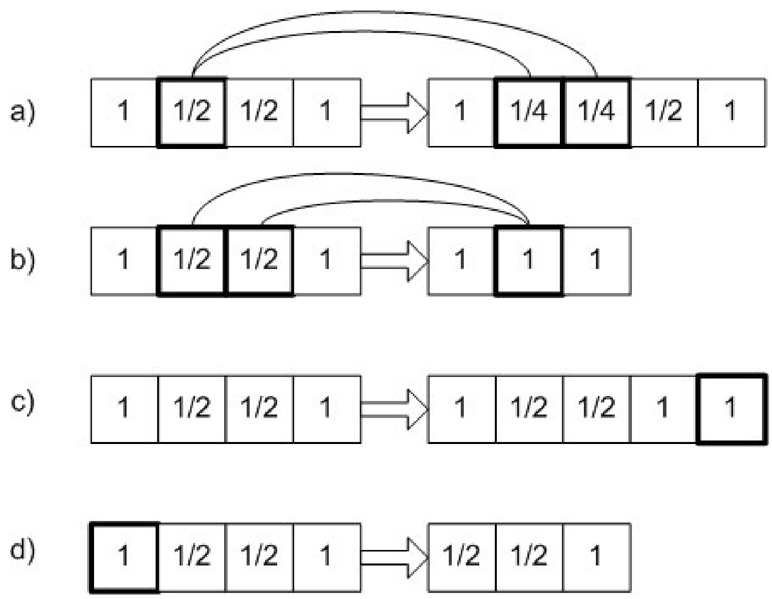


Figure 5.3: Examples of rhythmic transformations.

- Remove one element from the sequence (see Fig. 5.3d).

The definition of these transformations were inspired by the dynamical systems approach to study human bimanual coordination (Kelso 1984), and is based on the notion that two coupled oscillators will converge to stability points at frequencies related by integer ratios (Beek et al. 2000). Furthermore, common music notation facilitates these types of transformations. We have defined other transformations that divide a figure into three, five, and other prime numbers, but the impact of these additional transformations on the model is beyond the scope of this system. Addition and removal transformations were introduced to increase diversity in the pool of rhythms and to produce rhythms of different lengths.

5.1.2 Measurement of similarity of rhythms

The agents are furnished with an algorithm to measure the degree of similarity of two rhythmic sequences. This measurement is used when they need to measure the similarity of the rhythms that they play to each other. Also, this algorithm is used to measure the similarity between repertoires of rhythms from different agents.

In Martins et al. (2005), we introduced a method to measure the degree of similarity between two sequences of symbols by comparing various subsequences at various levels. The result is a vector, referred to as the *Similarity Coefficients Vector* (SCV), which contains the interim results of the comparisons between the subsequences. For the present work, we devised a version of the SCV method to deal with rhythmic sequences.

Let us define the block distance between two sequences containing the same number of elements as follows in Eq. 5.1:

$$d(\mathbf{v},\mathbf{w}) = \sum_{i=1}^{n} |v_i - w_i| \tag{5.1}$$

where $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_m)$ are the two sequences (vectors) being compared, v_i and w_i are the individual components of these vectors and m and n are the number of elements in each vector.

After obtaining the resulting evaluation of the block distances on a given level (length of a subsequence), we can write a matrix for the k-level, corresponding to the comparison of all the subsequences with length k between the two main sequences (Eq. 5.2):

$$\mathbf{D}^{(k)} = \begin{bmatrix} d_p(\mathbf{v}_1^{(k)}, \mathbf{w}_1^{(k)}) & \ldots & d_p(\mathbf{v}_1^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \\ d_p(\mathbf{v}_2^{(k)}, \mathbf{w}_1^{(k)}) & \ldots & d_p(\mathbf{v}_2^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \\ \vdots & \vdots & \vdots \\ d_p(\mathbf{v}_{(n-k+1)}^{(k)}, \mathbf{w}_1^{(k)}) & \ldots & d_p(\mathbf{v}_{(n-k+1)}^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \end{bmatrix} \tag{5.2}$$

Then, we can obtain a k-level Distance Matrix whose elements are non negative integers. Next, let us define the *k-level Similarity Coefficient* as follows (Eq. 5.3):

$$c^{(k)}(\mathbf{v},\mathbf{w}) = \frac{z(k)}{(n-k+1)(m-k+1)} \tag{5.3}$$

where $z(k)$ is the number of zeros in the matrix $\mathbf{D}^{(k)}$, n and m are the lengths of the vectors being compared, and the denominator in Eq. 5.3 corresponds to the number of cells in matrix $\mathbf{D}^{(k)}$. Roughly speaking, the similarity coefficient measures the *sparsity* of the matrix $\mathbf{D}^{(k)}$. The higher the coefficient $c^{(k)}$, the higher is the similarity between the subsequences of level k.

Next, we can collect all the k-levels coefficients in a vector referred to as *Similarity Coefficient Vector* (SCV). This is defined as follows (Eq. 5.4):

$$\mathbf{C} = \left[c^{(1)}, c^{(2)}, \ldots, c^{(min(m,n))}\right] \tag{5.4}$$

Fig. 5.4 shows an example of building a 3-level Distance Matrix, with a corresponding coefficient $SCV(3) = 0.125$, calculated by summing all the zeros in the matrix (1) and dividing by the number of cells in the matrix (8).

The values for the SCV corresponding to all the k-level distance matrices are $SCV = [0.4167 \quad 0.1333 \quad 0.1250 \quad 0]$.

From this vector, we can obtain a scalar value in to order establish a comparative analysis between larger sets of rhythms, such as the repertoires of two agents. We can take the rightmost nonzero value from the SCV, which corresponds to the higher level where two matching sequences can be found. We can either take a weighted sum of the SCV values or the average of all values, as follows (Eq. 5.5):

$$SCV_{av} = \frac{1}{min(m,n)} \sum_{j=1}^{min(m,n)} SCV(j) \tag{5.5}$$

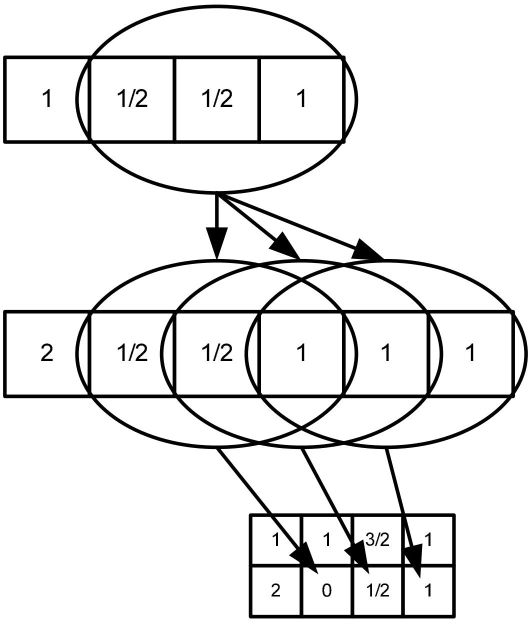


Figure 5.4: Example of building a 3-level Distances Matrix.

The next step is to compare the repertoire of the agents in order to observe the development of relationships amongst the agents in a group of agents. For instance, observing the agents form distinct sub-groupings.

The similarity of the repertoire of rhythms amongst the agents in a group is computed by creating a matrix of SCV_{av} values of the repertoires of all pairs of agents. Matrices with the columns and rows corresponding to the number of rhythms in the memory of each agent reveal how close their repertoires are to each other (Fig. 5.5).

By collapsing both the rows and the columns of the matrices, and taking the maximum values for each of them and an averaged sum, we obtain the scalar of similarity between repertoires, as follows (Eq. 5.6):

$$SimRep_{k,l} = \frac{1}{nR_{Ak}+nR_{Al}}\left[\sum_{i=1}^{nR_{Ak}} max(SCV_{av})_{rows} + \sum_{j=1}^{nR_{Al}} max(SCV_{av})_{cols}\right] \tag{5.6}$$

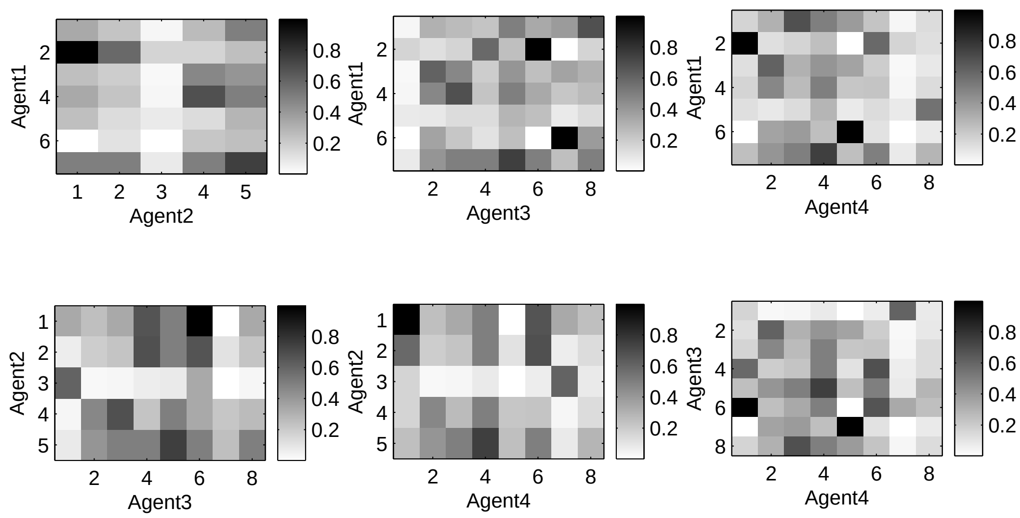


Figure 5.5: Similarity matrices between repertoires of 4 agents. Each axis of a graph represents the indexes of the rhythms from the repertoire of one agent. The darker the colour, the greater the similarity between two rhythms.

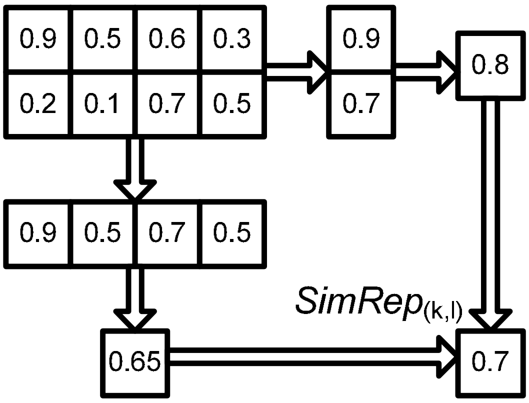


Figure 5.6: Scalar value of similarity between repertoires.

where nR_{Ak} and nR_{Al} are the number of rhythms in the repertoire of the compared agents. In the case shown in Fig. 5.6, this would be $SimRep_{k,l} = 0.7$ or, conversely, the distance between the repertoires of both agents, as defined in Eq. 5.7, would be $DistRep_{k,l} = 1 - Sim_{rep} = 0.3$. The values of 0.65 and 0.8 in Fig. 5.6 correspond to the similarity of the repertoires from the point of view of each agent, which is used to generate proximity matrices (Fig. 5.5) and graphs for monitoring the xbehaviour of the system (ex. Figs. 5.10 g),h) and f), Fig. 5.11 and Fig.5.21).

Finally, the development of the rhythm repertoires for the group of agents as a whole can be observed by conducting a hierarchical cluster analysis of all distance measures between the agents ($DistRep_{k,l}$). This cluster analysis produces a tree-like diagram (dendrogram), using a linkage method based on an unweighted average distance, also known as group average. The distance between two clusters A and B, D_{AB}, is given by the following (Eq. 5.7):

$$D_{AB} = \frac{1}{N_A.N_B} \sum_i d_i \tag{5.7}$$

where N_A and N_B are the number of elements in A and B, and d_i are pairwise distances between the elements of clusters A and B. The hierarchical cluster analysis produces a dendrogram (an example will be shown later in Fig. 5.11). The dendrogram is drawn through an iterative process until all the individuals or clusters are linked.

5.1.3 Measurement of complexity of rhythms

The theoretical background for complexity of rhythmic sequences is discussed in section 2.6. In the first implementation of A-Rhythm, the rhythms vary in number of elements and length. The rational behind the measurement developed in this section is that the representation of IOIs in fractions of the pulse can tell something about how much a pulse has been transformed. Complexity is therefore defined in this chapter, for the purpose of the first implementation of A-Rhythm, as a measure proportional to the number of rhythmic figures and to the values of the numerators of the fractions constituting the IOIs of a rhythm.

The complexity of a rhythmic sequence is defined as follows (Eq. 5.8):

$$Complexity = \frac{nF + \sum nN}{Duration} \tag{5.8}$$

where *nF* is the number of rhythmic figures contained in the sequence and *nN* is the sum of all numerators, considering that each rhythmic figure is a fraction of the pulse. This is a computationally cost effective method to measure the complexity of a rhythmic sequence. It is important to bear in mind that the implementation ensures that there are no reducible fractions included in the sequence, meaning that there always is a unique numerical representation for a given rhythm.

The rhythm $[1,2,1/2,1/2]$ has a complexity of $9/4$ as shown in equation 5.9.

$$Complexity([1,2,1/2,1/2]) = \frac{4 + \sum[1,2,1,1]}{\sum[1,2,1/2,1/2]} = 9/4 \tag{5.9}$$

Fig. 5.7 shows an example of a graph plotting the value of complexity of a sequence of interonset intervals $[1 \quad 1]$ after being subject to 30 successive transformations.

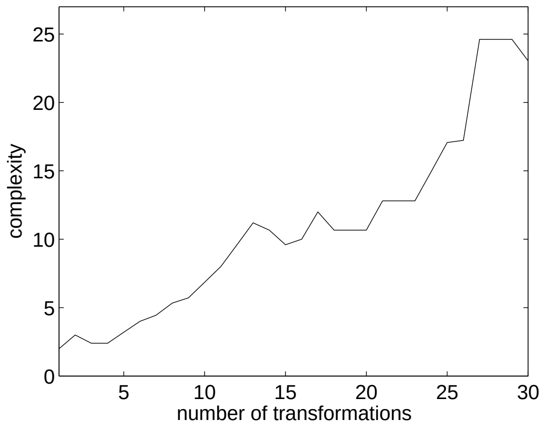


Figure 5.7: Complexity increases with the number of transformations.

5.2 Interaction algorithms and experiments

The interaction algorithms and the analysis methods that we have implemented in our system are introduced below. Each algorithm is introduced in the context of illustrative experiments aimed at studying the development of repertoires of rhythmic sequences from three different perspectives:

- From the perspective of an individual agent;
- From the perspective of a group of agents, referred to as the society;
- From the perspective of the developed rhythms.

From the perspective of an individual agent, the analysis focused on the study of the development of the size (number of rhythms) and on the complexity of the repertoire of individual agents. From the perspective of the society, the values of the corresponding individual measures from the agents were averaged. Furthermore, the similarity between agents was calculated and the agents were clustered in terms of the rhythms that they shared. Finally, from the perspective of the developed rhythms, the analysis measured their lifetime, the amount of rhythmic sequences that the society developed

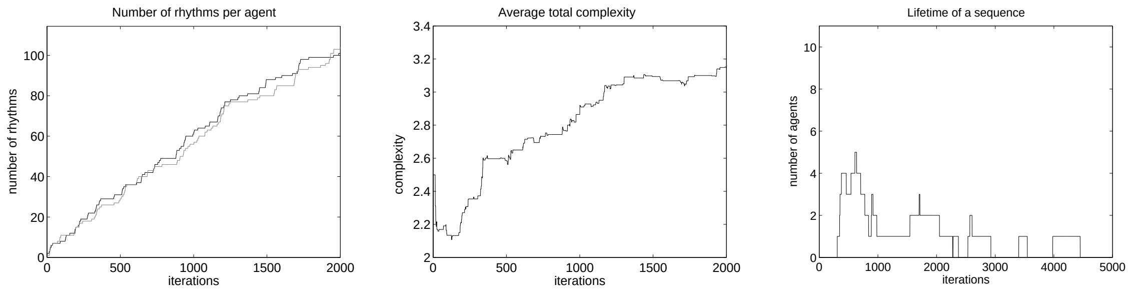


Figure 5.8: a) Development of the size of the repertoire for different agents; b) Complexity of the rhythms of the whole society; c) Number of agents sharing a particular rhythm.

and the degree to which the agents shared similar rhythms. The lifetime of a rhythmic sequence was traced by counting the number of agents that possessed this sequence at each iteration. Fig. 5.8 shows graphs illustrating these various types of analyses.

The experiments were run for 5000 iterations, with the objective of observing the agents' behaviour under different conditions. Experiments were run with societies of 3, 10 and 50 agents. Sometimes the algorithm considers the movement of the agents in the 2D space, which may or may not influence the nature of the interactions.

5.2.1 The popularity algorithm

Popularity is a numerical parameter that each agent attributes to a rhythm in its repertoire. The parameter is modified both by the listener and by the player during an interaction. If the listener recognises the rhythm (i.e., if it holds this rhythm in its repertoire), then it will increase the popularity index of this rhythm and will give a positive feedback to the player. A positive feedback is an acknowledgment signal, which will prompt the player to increase the popularity index of this rhythm in its repertoire as well. Conversely, if the listener does not recognise the rhythm, then it will add this rhythm to its repertoire and will give a negative feedback to the player, which will cause the player to decrease the popularity index of this rhythm. Furthermore, there is a memory loss mechanism whereby after each interaction all the rhythms have their popularity index

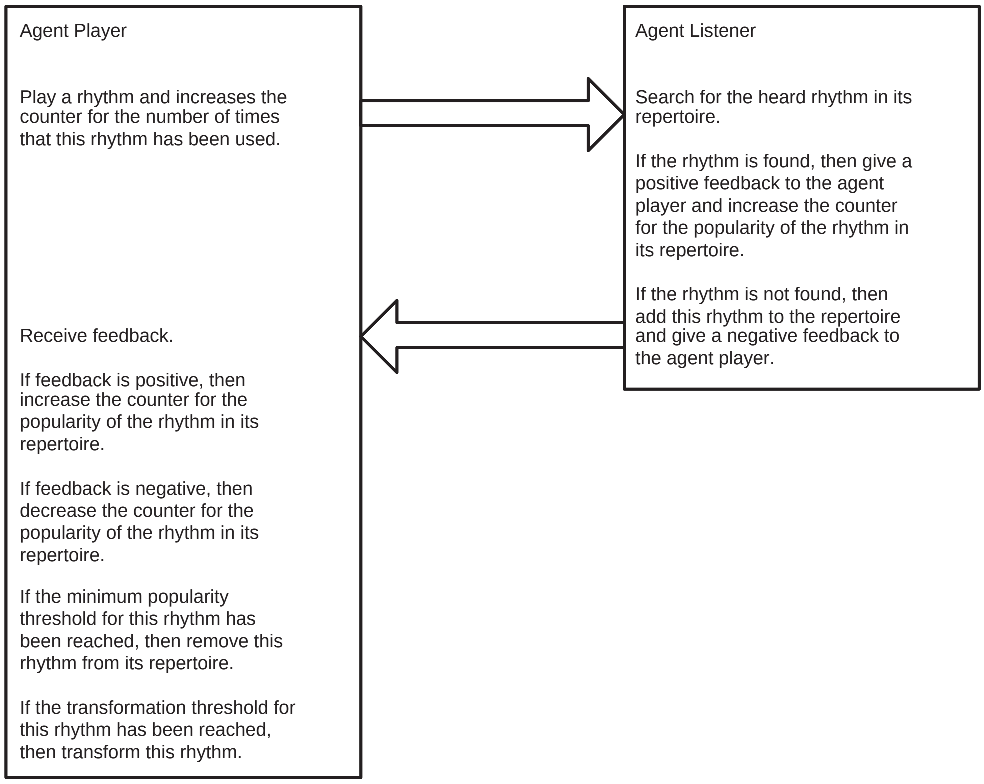


Figure 5.9: The popularity algorithm.

decreased by a small value of 0.05. This accounts for a natural drop in popularity due to ageing. The diagram of this interaction is displayed in Fig. 5.9.

Fig. 5.10 shows the results after 5000 iterations of the popularity algorithm without population renewal. Fig. 5.10a displays the development of the repertoire from the individual agents and the graph from Fig. 5.10b displays the corresponding average across the agents. Here, the repertoire of each agent grows monotonously during 500 iterations and subsequently oscillates around a stable point. Fig. 5.10c displays the development of the repertoire of the whole society, being a direct consequence of the lifetime of each rhythm. The average number agents sharing a rhythm (Fig. 5.10d) is calculated by summing the instant number of agents sharing a rhythm (Fig. 5.8c) for all rhythms, and dividing the result by the number of rhythms currently present in the society (Fig. 5.10c). This graph (Fig. 5.10d) provides the means to assess the global

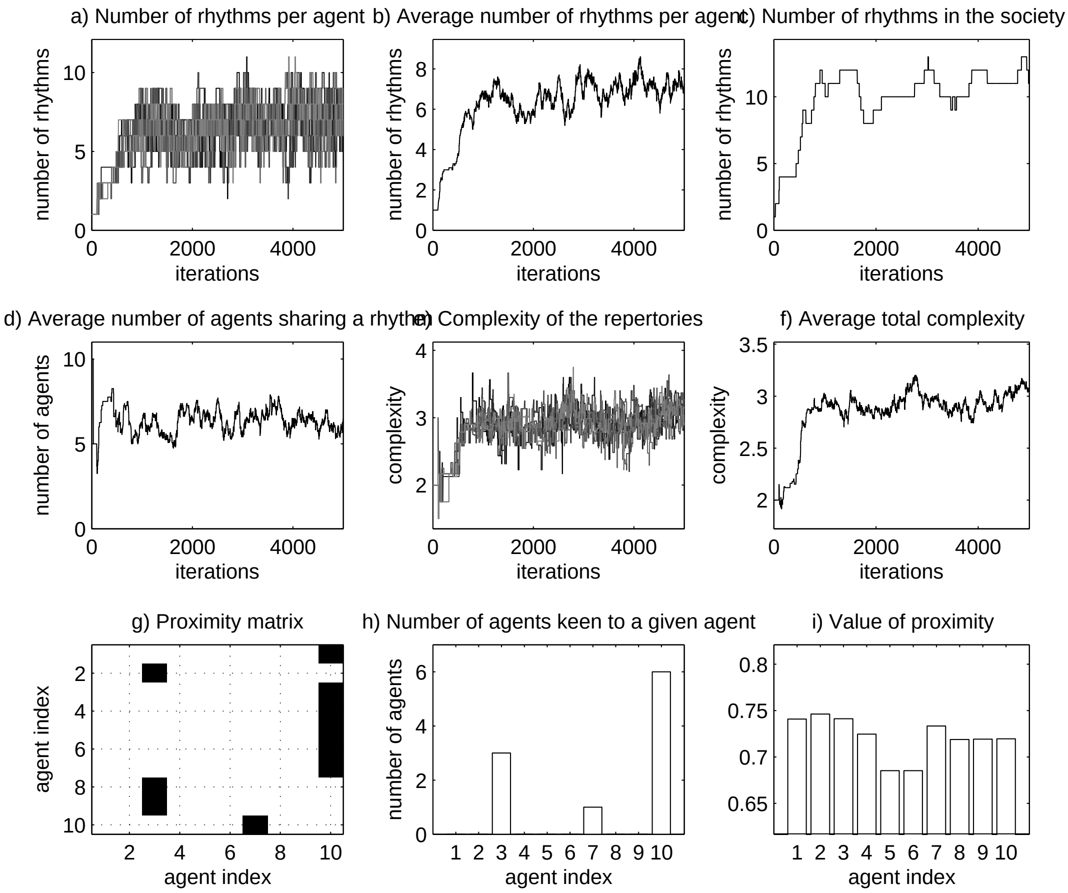


Figure 5.10: Results from a representative simulation using the popularity algorithm with 10 agents.

behaviour of the society; for instance, if it develops coherently in terms of popularity of existing rhythms. High values in Fig. 5.10d means that many rhythms are common between the agents.

Fig. 5.10e represents the development of complexity of the individual agents and Fig. 5.10f gives the corresponding average. Initially, the size and complexity of the repertoire of individual agents are very close to the average, but this trend is replaced quickly by repertoires of different sizes amongst the agents.

The last three graphs show the degree of similarity between the repertoires of the agents according to the measure defined in Sec. 5.1.2. Fig. 5.10g displays information about the identity of the agent with whom each agent relates most, i.e., has the highest similarity value. The graph in Fig. 5.10h shows the agents that are regarded by others as being most similar to them. This means that agent number 3 has three agents with similar repertoires (2,8 and 9), and agent 10 is the one that concentrates the highest number of keen agents, having 6 agents considering its repertoire to be more similar to theirs.

Hierarchical cluster analysis is conducted as described in Sec. 5.1.2 in order to observe groupings of agents according to the distances between them. Fig. 5.11 shows the dendrogram containing elements of three societies of 10 agents each: society 1 comprises agents 1 to 10, society 2 comprises agents 11 to 20 and society 3 the remaining 21 to 30. By comparing the three societies that were developed independently in three separate runs, with the same set of parameters, we can observe three clearly independent clusters. In addition to the previous observations, this suggests that the repertoires that emerged from the popularity algorithm display diversity because of the number of rhythms emerged, are stable in terms of size, and are coherent within their respective societies. We can also observe differences in the clusters within a given society.

We also investigated whether the interaction rules could influence the movement of the agents and whether this process would influence the development of their repertoires.

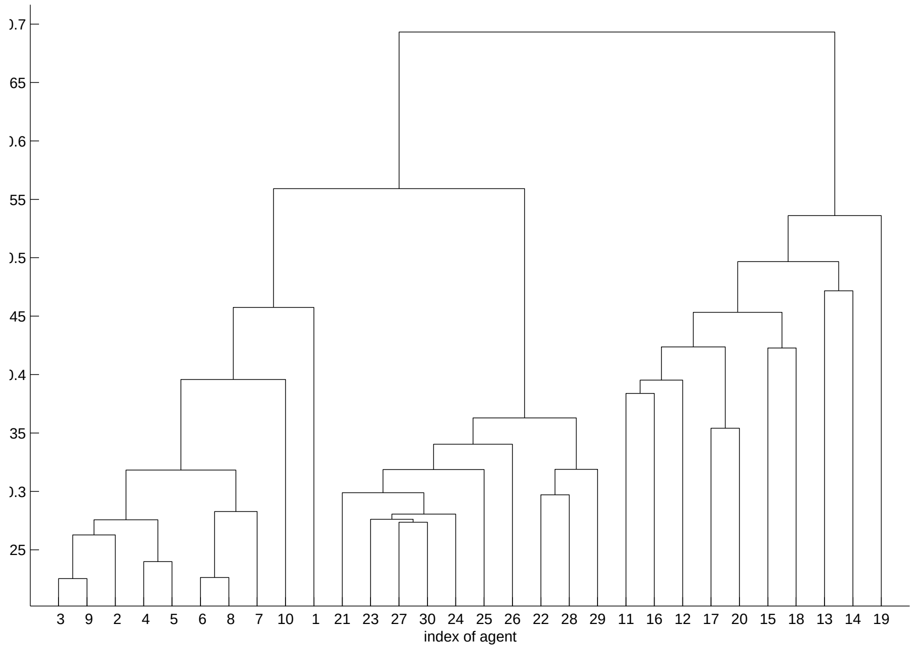


Figure 5.11: Dendrogram resulting from the hierarchical cluster analysis conducted in the context of the popularity algorithm, containing three independent societies with 10 agents each.

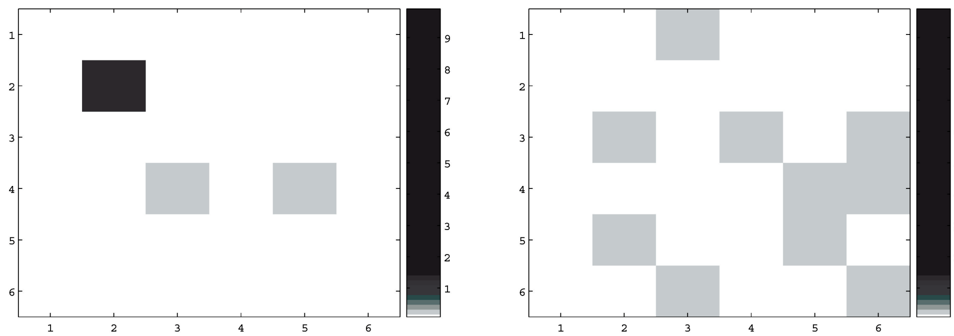


Figure 5.12: World visualisation of two steps of the iterative process where clustering is observed (figure on the left) and later broken (figure on the right). A cluster is indicated by a darker colour.

The type of movement used is based on a *random walk* algorithm. In this case, when a listening agent recognises a rhythm that has been played, originating also positive feedback, it will follow the player in the space in the next iteration.

Fig. 5.12 shows periodic clustering of one or more groups of agents that align future movement direction and keep interacting until the cluster is broken due to an unsuccessful interaction. In Fig. 5.13, we can observe two behaviours of the system that are typical of the popularity algorithm with movement taken into account. First, there are many more rhythms than in the case without movement affecting the interactions. This is due to the fact that every time a positive feedback occurs, an interaction between two or more agents from a group will also take place in the following iterative step. This rises the number of interactions and consequently the number of emerged rhythms. Second, there is an initial overshoot of the repertoire size before reaching a stable level, and this is possibly caused by the initial clustering of agents, when individual repertoires grow consistently among very closely related agents.

Fig. 5.14 shows the lifetime of sequences that emerged during typical runs of the popularity algorithm.

5.2.2 The transformation algorithm

The transformation algorithm is shown in Fig. 5.15. As its name suggest, the transformation algorithm applies transformations on a rhythm whenever it is communicated between agents. The motivation behind this algorithm is to foster novelty. We conducted experiments to evaluate the degree to which, transformations occurring during the interactions have an impact on the organisation of the emerging repertoire, as time progresses.

In Fig. 5.16 it is possible to observe that due to the rise of the amount of transformations, the repertoires are much larger than in the popularity algorithm. Looking into

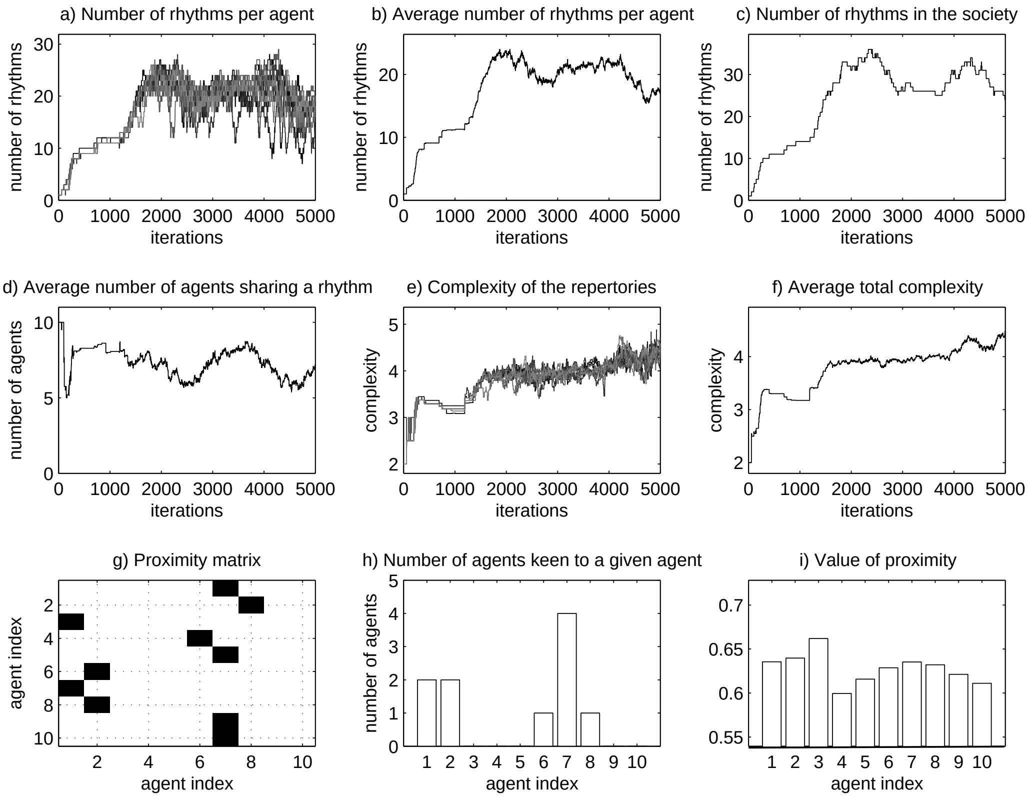


Figure 5.13: Results from a representative simulation using the popularity algorithm taking into account the movement of the agents as an influencing factor in the development of the repertoire.

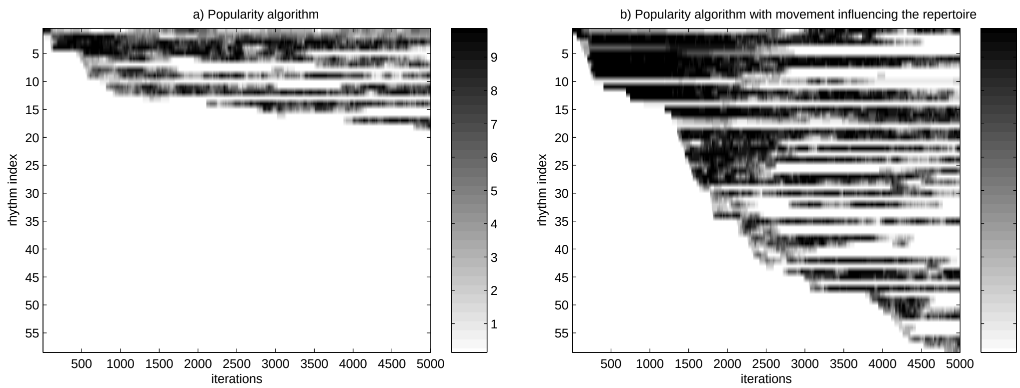


Figure 5.14: These graphs show the lifetime of all the rhythms that emerged in the society with the popularity algorithm in the cases where: a) movement does not influence the developments; b) when movement influences the developments. The number of agents that share a particular rhythm is represented by tones of gray (the darker the color, the higher the number of agents).

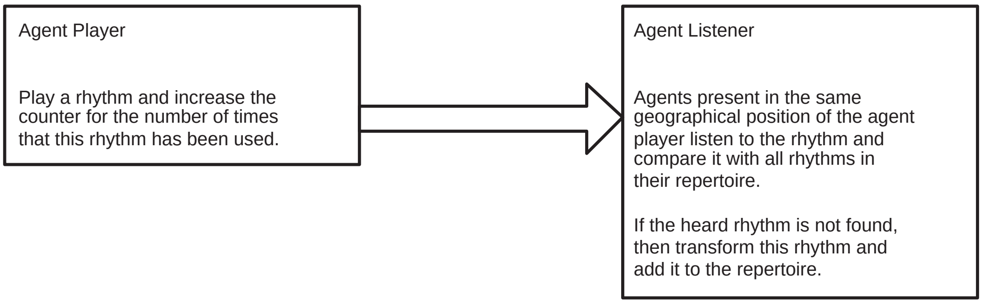


Figure 5.15: The transformation algorithm.

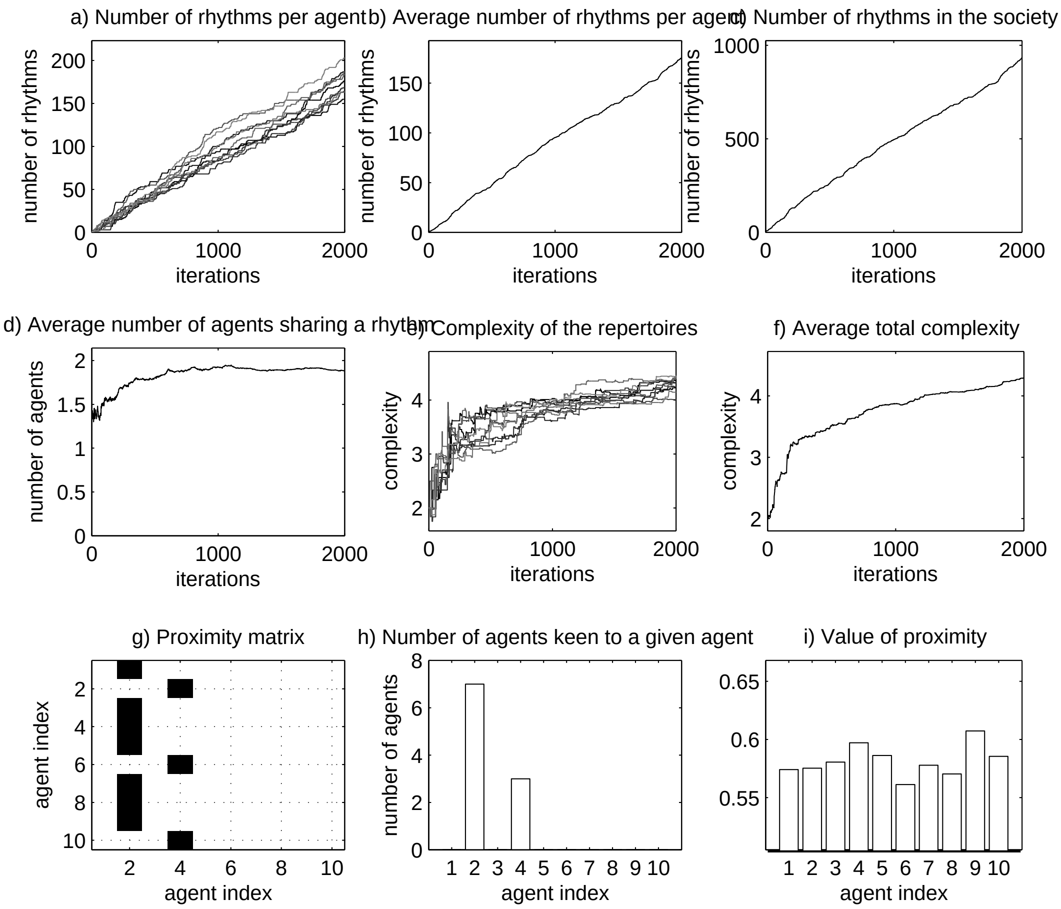


Figure 5.16: Results from a representative simulation using the transformation algorithm with 10 agents.

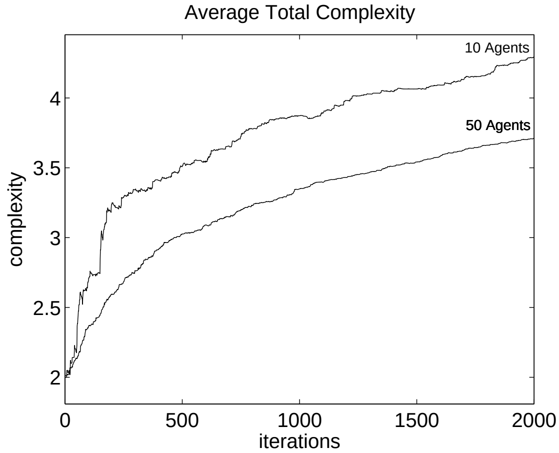


Figure 5.17: Average complexity evolution curves resulting from the transformation algorithm with 10 and 50 agents.

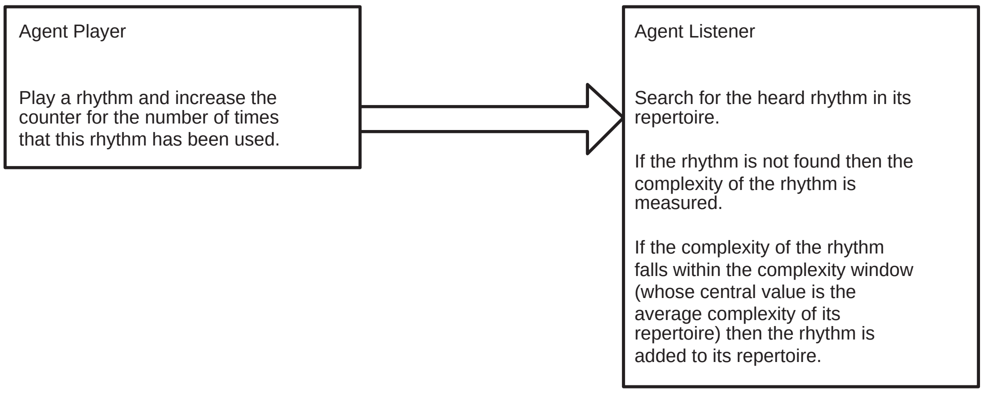


Figure 5.18: The complexity algorithm.

the average complexity development of the society we can observe two clearly differentiated growing rates before and after 200 iterations. When the algorithm is run with 50 agents we can also observe similar growing rates, although the initial rate is not as steep as it is with 10 agents, and the transition is smoothed (Fig. 5.17).

5.2.3 The complexity algorithm

The diagram of the complexity algorithm is shown in Fig. 5.18. The complexity algorithm studies the effect of preference for particular types of rhythm. In this case, the study aims at establishing whether the agents would show preference for rhythms with identical complexity, as defined in section 2.6.

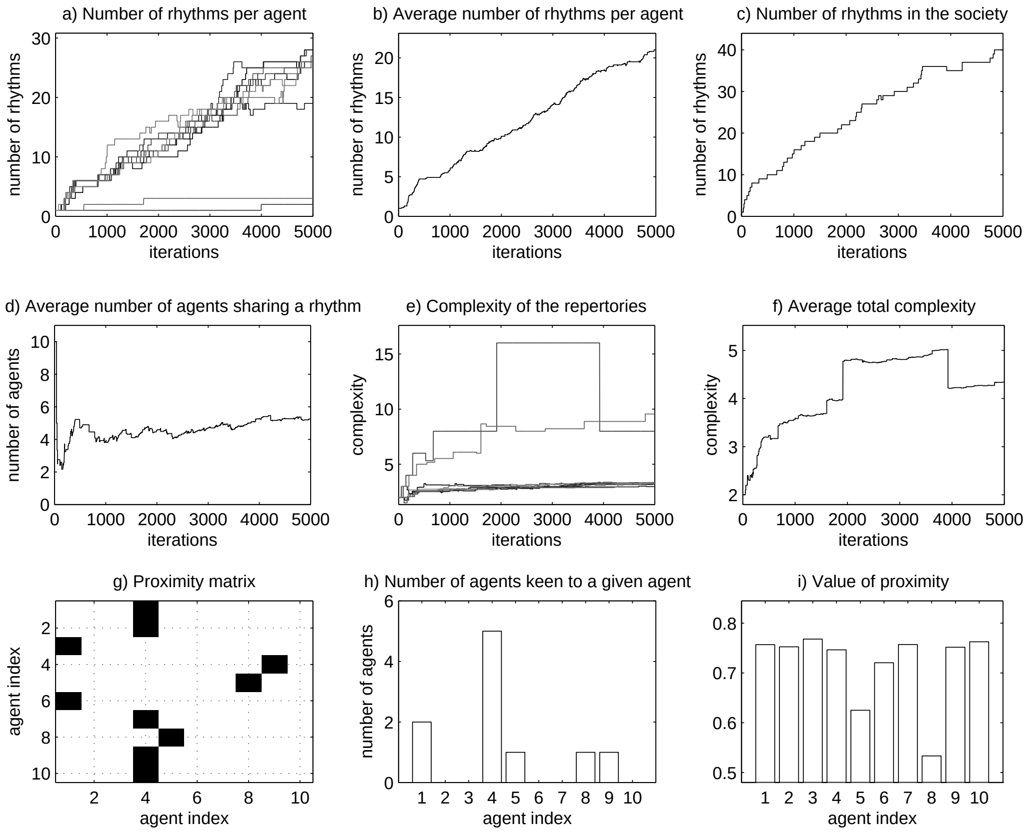


Figure 5.19: Results from a representative simulation using the complexity algorithm with 10 agents.

Here, the agents include in their repertoire only those listened rhythms that fall within a window of complexity centered in the average complexity of the rhythms of the listening agent. That is, all listened rhythms that are in the interval of $[AvComplexity - complexWindowRadius; AvComplexity + complexWindowRadius]$ will be included in the repertoire of the agent.

Fig. 5.19 displays the results from a run of the complexity algorithm with the same parameters as the run of the popularity algorithm shown in Fig. 5.10. The most interesting emergent behaviour that can be observed from the graphs in Fig. 5.19a, 5.19e and 5.19i, is the emergence of distinct repertoires developed by agents 5 and 8; they are distinct in terms of the complexity and number of developed rhythms. Although they are considered to have the smaller values of proximity in relation to the closer agent

(Fig. 5.19i), their development seems to be tightly connected.

It is seen here that initial small changes in complexity due to transformations can actually result in completely different developments between the agents.

The cluster tree for the results shown in Fig. 5.19 is given in Fig. 5.20. Two main clusters appear in the figure, separated by a value of $DistRep = 0.8$. Furthermore, the two agents that, at an early stage of the simulation, were able to perform transformations leading to sequences of higher complexity, remain more apart than the agents of the other cluster.

We also applied Principal Components Analysis (PCA) to study clustering. PCA is suitable here because the error between the multidimensional distances and its 2D reduction is relatively low (Fig. 5.21). Fig. 5.22 shows the development of agent complexity and repertoire size in a simulation where movement was linked to the success of the interaction process. One of the agents started to be more complex after 1500 iterations and then it joined the group with the larger amount of agents.

5.3 Summary

In this chapter, the first implementation of the multi-agent system A-Rhythm was presented and the results were described. This system comprises a virtual environment populated by artificial agents that exchange musical rhythmic units via three different algorithms of interaction. In the beginning of the chapter, the agents are characterised, including the representation for the rhythmic units and their corresponding parameters. Subsequently, the process of transformation of rhythmic sequences is described along with the specific measurements of similarity and complexity. The three algorithms, namely *popularity*, *transformation* and *complexity*, are characterised by the actions taken by the listening agents when facing a new rhythmic unit. The rhythmic repertoire of the three different algorithms develop from a single pulse to a set of diverse

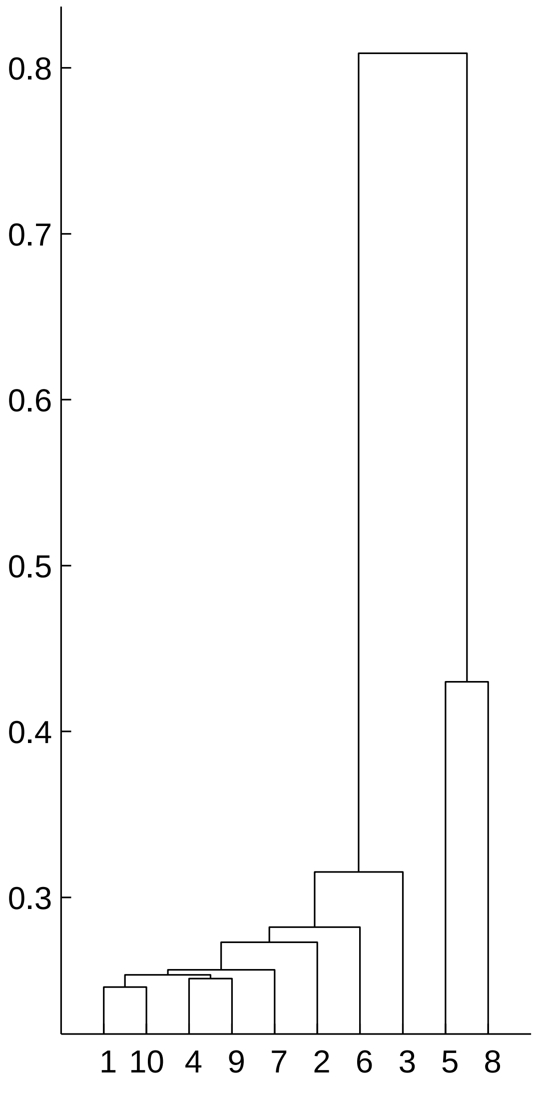


Figure 5.20: Dendrogram resulting from the hierarchical cluster analysis conducted in the context of the complexity algorithm with 10 agents.

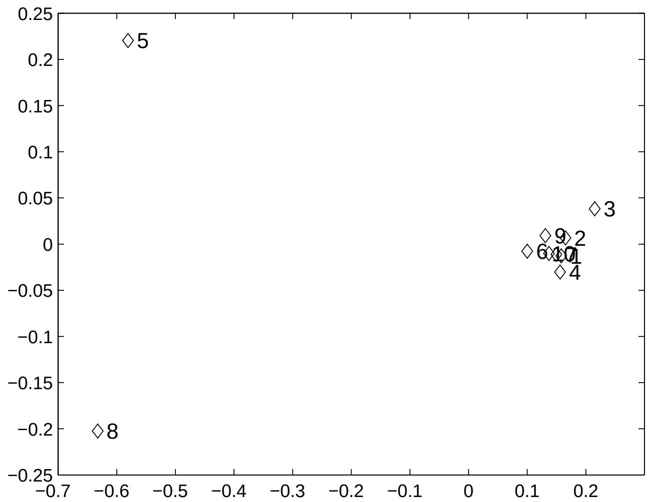


Figure 5.21: Principal Components Analysis (PCA) of the results from a representative simulation using the complexity algorithm with 10 agents after 5000 iterations.

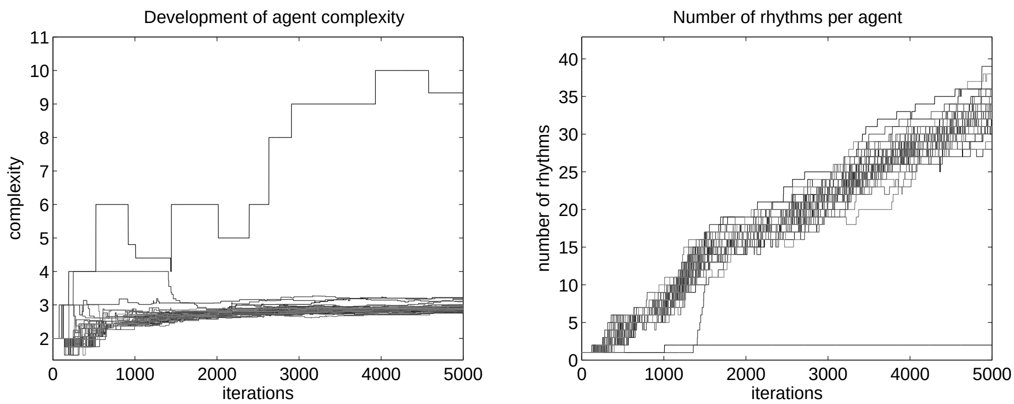


Figure 5.22: Development of complexity and of the number of rhythms per agent over the course of a run of the complexity algorithm with 50 agents, considering preference in terms of movement.

rhythms for each agent.

The results show that the popularity algorithm enables the evolution of a stable average number of rhythms at any given time during the simulation. Particular rhythms are in average shared by slightly more than half the population of the agents and the complexity of the repertoires remains within boundaries after a rise period of roughly 1000 iterations. The hierarchical cluster analysis based on the similarity of repertoires reveals several groups of agents, which have a correspondence to their geographical proximity. In a variation of the popularity algorithm, feedback is used to affect the next movement of the agents. As a result, the number of interactions increases, causing a corresponding increase in the popularity of the rhythms.

In the transformation algorithm, rhythmic transformations are performed over listened rhythms. The results show an increase without boundaries of the repertoire size and complexity. The agents become much more heterogenic in terms of their repertoire, showing large repertoires, but having only a few rhythms in common.

Finally, in the complexity algorithm, the agents choose to listen to rhythms based on complexity. In this algorithm, cluster analysis of the repertoires reveals that this mode of transmission enables differentiation of repertoires among groups of agents, as some agents exchange rhythms preferably between themselves. The binding of the groups is not permanent, and the emergence of new groups is observed throughout the simulation, as repertoires change due to new transformations.

The development of these algorithms paved the way for a new system which introduces the notion of a rhythmic grammar and will be explained in the next chapter.

Chapter 6

A-Rhythm: rhythmic structures

The natural progression from the version of A-Rhythm presented in Chapter 5 is to implement additional features to the agents, such as a grammar, in order to be able to produce rhythmic structures.

The aim of this exercise is to study the evolution of rhythmic grammars in an A-Life type of environment populated by interactive autonomous agents, endowed with rhythm perception/production capabilities. The agents also move in a virtual environment and interact with other agents by playing and perceiving rhythms. The interaction takes the form of music games involving groups of agents of variable size (depending on their location in the environment). Through these games, agents engage in a cultural exchange process which drives the evolution of a rhythmic grammar in a bottom-up process.

6.1 Hedonic values

In 1874, the german psychologist Wilhelm Wundt (Wundt 1874) presented a bell-shaped curve relating the arousal potential evoked by a stimulus and the resulting hedonic value. Berlyne (1971) proposed that this curve is produced by a combination of the primary reward system and the aversion system (Fig. 6.1).

The resulting hedonic value attributed to a stimulus would be obtained by the algebraic summation of the activity curves from the reward and aversion system.

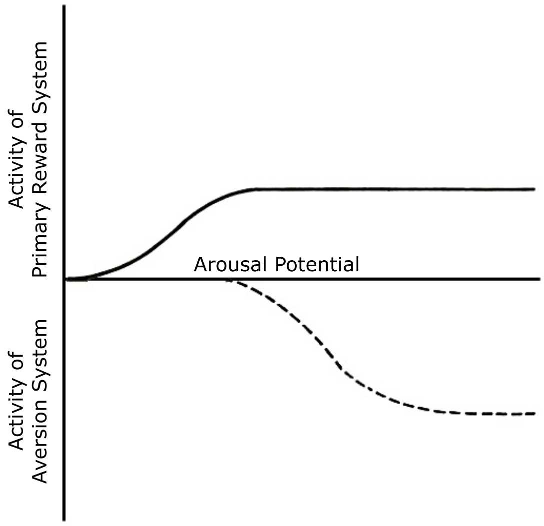


Figure 6.1: According to Berlyne (1971) the hedonic function results from the curves of the activity of the reward system combined with the activity of the aversion system.

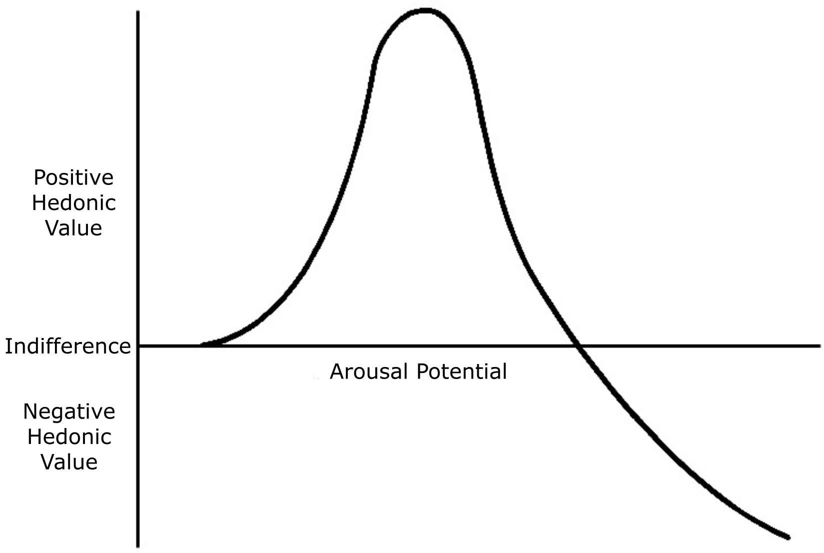


Figure 6.2: The resulting bell shaped curve derived from the algebraic sum of the curves depicted in Fig. 6.1.

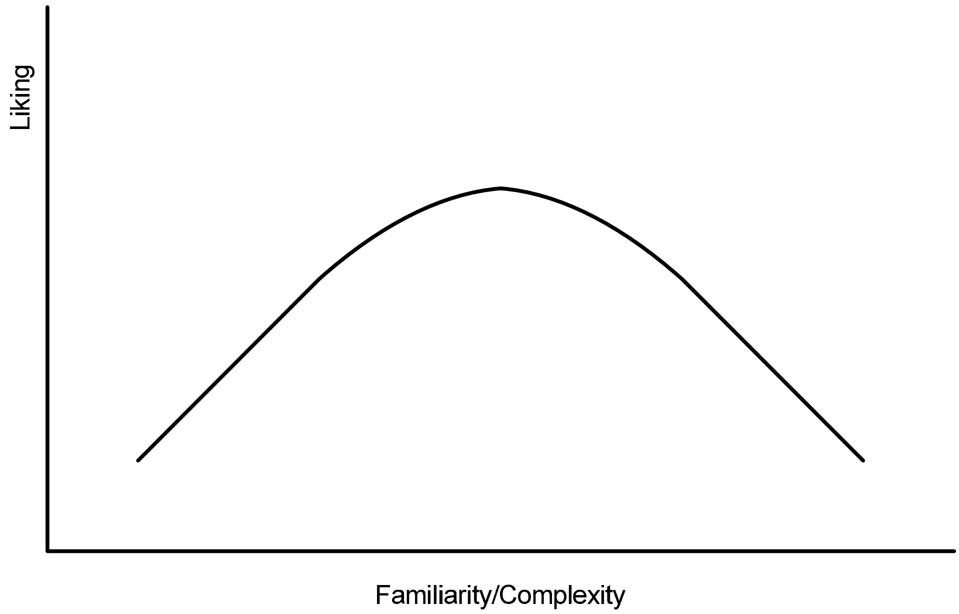


Figure 6.3: The bell shaped curve indicates preference for moderate values of complexity and familiarity.

Stimuli have several properties influencing arousal and these fall under three main categories: *psychophysical* properties, which depend on spatial and temporal distributions of energy (e.g. the brightness of lights or the loudness of sounds can have an influence on arousal and consequently have a corresponding hedonic value); *ecological* properties, involving associations with biologically noxious or beneficial conditions (e.g. modulation of arousal caused by physical pain or comforting physical contact); and lastly *collative* properties, named this way as it order to *collate* different sources of information to evaluate if the stimulus possesses this property (Berlyne 1971 p.69). This last group refers to stimulus properties such as novelty, surprisingness, complexity, ambiguity and puzzlingness, which are, according to Berlyne, the most relevant for aesthetical purposes. The claim that the hedonic value follows an inverted bell curve as a function of arousal caused by these properties (Fig. 6.3) is justified in his work with psychobiological arguments.

The theory of Berlyne had a strong impact in subsequent research. North and Hargreaves (2008) provide a review on the empirical work, both in laboratory and actual performance environments, that has been made since then.

In the analysis of music complexity provided by North and Hargreaves (2008 p.76) a link between complexity and a correspondent hedonic value is established. The theory is derived from the works of Berlyne. In the same work North and Hargreaves (2008 p.81) discuss how increased exposure can change the ratings of complexity and mention an unpublished study by Tuomas Eerola where ratings of albums of "The Beatles" seem to confirm this theory. The notion of subjective complexity was briefly described in section 2.6.

The sole existence of this psychological mechanism, when applied to musical parameters, might explain why some music pieces are more diffused than others. Therefore, it gains relevance in the context of this thesis. The theory of Berlyne, and simplified functions for the hedonic values, are included in the second implementation of A-Rhythm. The interaction algorithm considers two feedback values based on the evaluation of complexity and exposure of rhythmic units which will influence the process of transformation of rhythms in the repertoire.

In the second implementation of A-Rhythm, the transformation rate at which an individual agent will create a new sequence is contingent upon the hedonic evaluations of the previously generated rhythms.

6.2 Artificial environment

The artificial environment (AE) is a 2D discrete space with toroidal boundaries. A torus is a space which can be represented by a square, where each side of the square is connected to the exact opposite as shown in Fig. 6.4.

Agents are allowed to move freely in any direction, and each cell is a possible location. Multiple agents can be hosted in the same location simultaneously. Agents' positioning

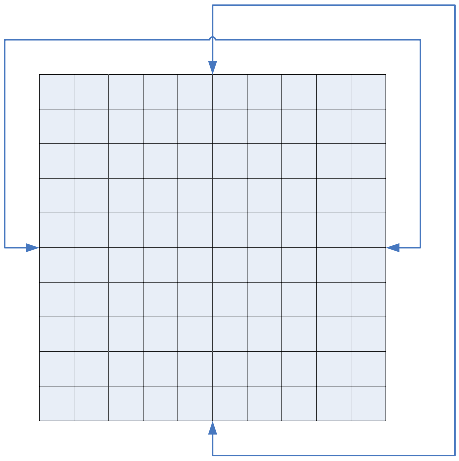


Figure 6.4: 2D world with toroidal boundaries.

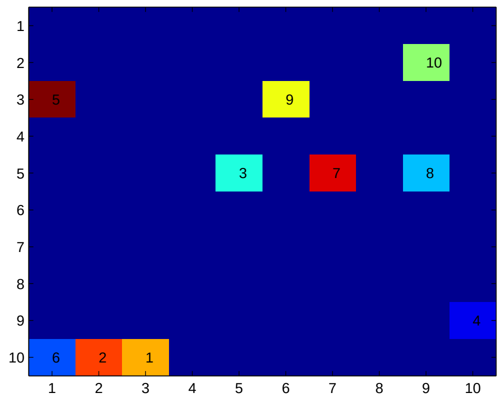


Figure 6.5: Artificial environment populated with 10 agents.

within the environment is a determinant factor, because the possibility of interacting with other agents, i.e., to perceive and play rhythms from/to other agents, depends on their actual location and neighbourhood radius (the interaction algorithm will be explained later in more detail).

6.3 Agents

The agents in this implementation of A-Rhythm present new features in relation to the previous model. In this section, the architecture and the life cycle of the agents are discussed.

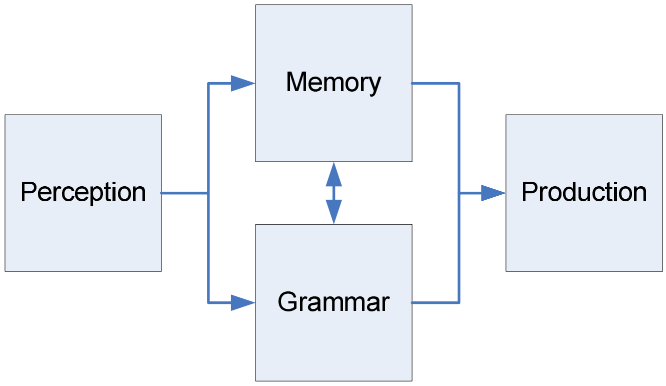


Figure 6.6: Agent's modules.

6.3.1 Agent's architecture

In this work each agent is composed of 4 different modules: perception, memory, grammar and production (Fig. 6.6). An expanded version of Fig. 6.6 is presented in Fig. 6.17.

Perception

The perception module enables the agent to sense and process acoustic stimuli from the environment. The acoustic stimuli consist of rhythmic sequences that agents are able to perceive and parse into rhythmic elements of shorter duration, which I will refer to as rhythmic units (RUs). An example of the parsing process appears in Fig. 6.7, where the sequence is divided into RUs of equal length and with a duration of two beats each. In the current system, in order to evaluate the complexity of the RUs according to the PS-Measure of complexity, each RU contains 4 beats. In addition, the perceptual model allows also the agent to extract the pattern of transitions between the RUs of each perceived sequence.

Memory

The perception module interacts with the memory when a rhythmic sequence is perceived from the environment. The memory module stores a corpora of RUs that mirrors

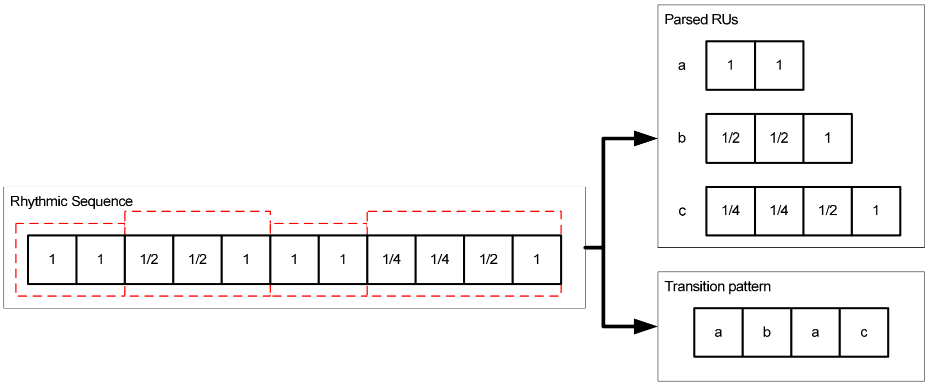


Figure 6.7: Parsing a sequence with the perception module.

the agent knowledge of rhythmic units and sequences. Furthermore, it also manages a set of parameters that typify the individual RUs, and their relation to the whole corpora (see Tab. 6.1. The structure of the memory module of the agent is depicted in Fig. 6.8).

The parameters of each RU are as follows:

1. *ioiVec*, contains the values of the relative Inter-Onset Intervals (IOIs) in the RU, thus representing its temporal structure. The term *relative* alludes to the fact that tempo is considered to be constant along the simulation, and the value of 1 for a IOI spans the duration of one element of the tactus. Due to the nature of the phenomena under study, expressive timing is not considered, and agents codify rhythmic elements like notes in a musical score. An example for a RU is depicted in Fig. 6.9;

2. *isDownBeat* is a boolean variable indicating with a 1 if the RU is to be played with an onset at the initial IOI (downbeat), or with a 0 if the first IOI is connected to the last IOI of the previous RU, when used in a sequence (syncopated).

3. *isCreated* is another boolean variable indicating whether the RU was created by this agent (1) or weather it was perceived from the environment (0), i.e. another agent playing it);

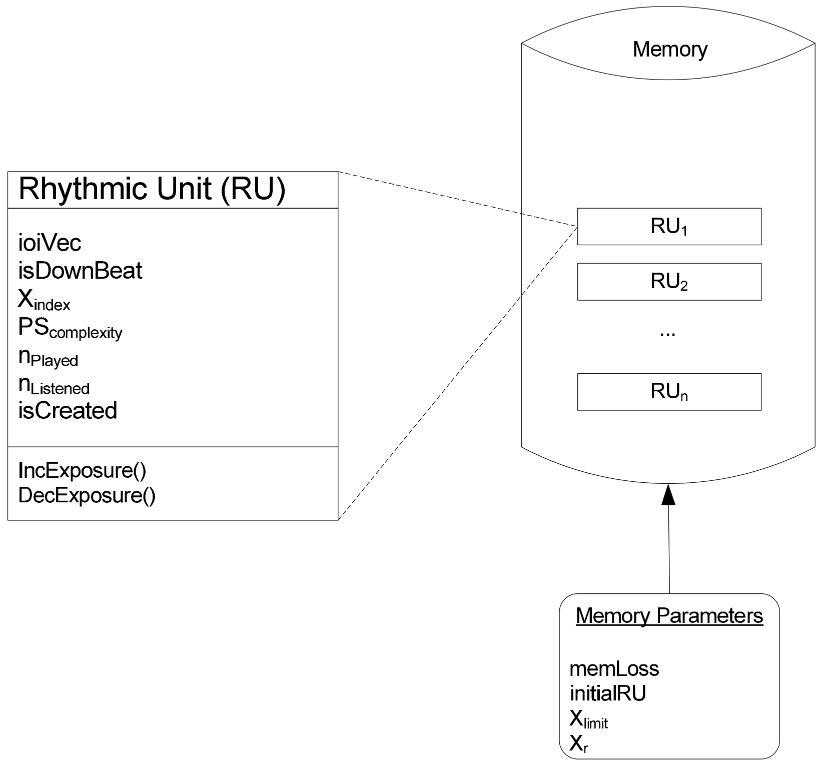


Figure 6.8: Memory structure for the agent.

4. $Iter_{in}$ stores the time step when the RU was integrated into the agent's memory;

5. *PS* is a complexity rating for the RU and is computed the first time it is created. The computation of complexity is dependent on the vector with the sequence of the IOIs (*ioiVec*), on the *isDownBeat* value and on the metric structure of a sequence. Complexity is measured for each unit with the PS-Measure (Shmulevich and Povel 2000) displayed in Eq. 2.3 and described in Sec. 2.3.3. This value may vary between 1 and 5 for sequences pertaining to the scope of this study. This measure differs from the implementation of A-Rhythm shown in chapter 5 as it is based in empirical studies;

6. n_{played} stores the number of times the RU was played by the agent;

7. $n_{listened}$ stores the number of times the RU was perceived by the agent;

8. X_{index} is exposure value of a given RU. It varies proportionally to $n_{listened}$ and is

Figure 6.9: Musical notation for a rhythmic unit

decreased as simulation time advances (the memory loss rate, which decreases exposure, is defined by the simulation parameter *memLoss*; see Sec. 6.5). If X_{index} drops below the threshold value of 0, then the RU is deleted from the agent's memory (although it may reappear later in the future when played by some other agent or recreated by the agent itself).

Every time an agent is created, it contains a single RU corresponding to an isochronous pulse in its memory.

The parameters computed for a particular RU contained in an agent's memory, and represented in Fig. 6.9 are summarised in Tab. 6.1.

RU Parameter	**Description**	**Example value**
ioiVec	Interonset intervals of RU	[1 0.5000 0.5000 1 1]
isDownBeat	Presence of even at the beginning of the RU	1
isCreated	If RU was created by this agent (1/0)	1
$Iter_{in}$	In which iteration the RU was integrated	2
PS	Value of Povel/Shmulevich complexity	13.490
n_{played}	Times played by this agent	15
$n_{listened}$	Times listened	22
X_{index}	Exposure value	0.9100

Table 6.1: Parameters of each Rhythmic Unit.

Grammar

The grammar module consists of the internal representations of the patterns of interconnectivity between RUs, that is, the way RUs are organized as rhythmic sequences. The grammar is implemented by a Markov model in which state transitions are defined by a matrix of probabilities between RUs. These probabilities are learned by the agent through exposure to rhythmic sequences while interacting with its environment or with

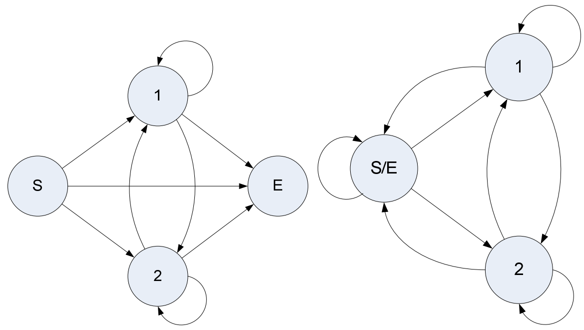


Figure 6.10: Markov transitions states. The numbered nodes represent Rhythmic Units and the directed arcs represent transition probabilities.

itself (this process will be further explained in Sec. 6.4).

The Markov model can be described by a state transition diagram of nodes representing the RUs and directed arcs representing transition probabilities of change between RUs, (Fig. 6.10a). There are two unnumbered nodes, which indicate the beginning of a sequence (S) and the point where it ends (E). This diagram can be shown in a more compact way by coinciding the *Start* and *End* nodes (Fig. 6.10b).

Sequences produced according to this grammar are created by selecting a path in the state transition diagram, where the next state is decided according to the transition probabilities, by using a roulette method in each transition. The process ends when the "End" state is reached.

The structure of the grammar module of the agent is depicted in Fig. 6.11, which shows that the grammar has two associated parameters. An internal parameter a_{mat}, and a received *exposureFeedback* parameter are responsible for the changes to the matrix.

Production

The production module endows the agent with the capacity to create new rhythmic sequences which can then be played to the environment or to itself. The production module recruits the memory module (to retrieve RUs) and the grammar (to sequence RUs) in the production of rhythmic sequences.

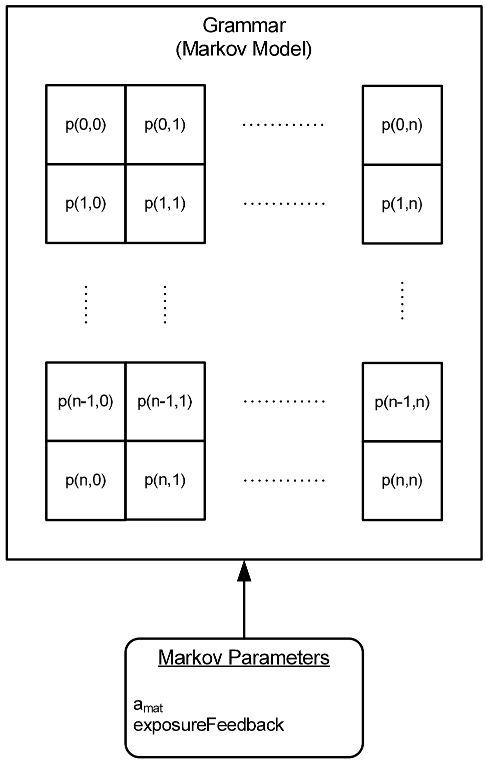


Figure 6.11: Grammar structure for the agent.

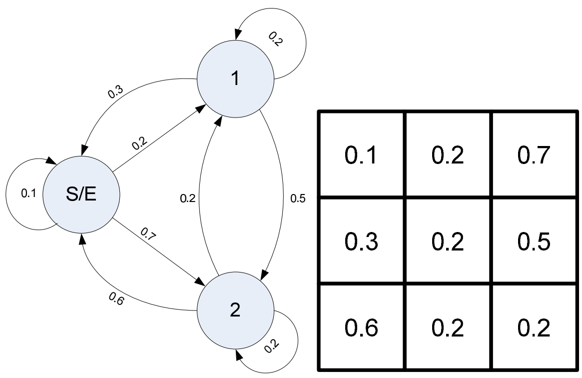


Figure 6.12: a) Markov transitions diagram with probability of transition discriminated in the arcs. b) Markov matrix, the transitions from the *start* state corresponding to the first line of the matrix and the transitions to the *end* state corresponding to the first column of the matrix.

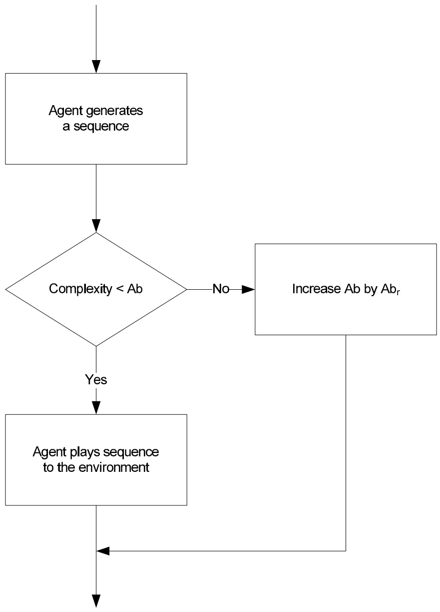


Figure 6.13: Agent process of rhythmic production concerning ability.

The production of rhythmic sequences varies from agent to agent, being a function of their ability to perform RUs of different complexity. Ability (*Ab*) is a parameter of each agent and determines the most complex sequence that one agent is able to play. Agents' ability increases (by a factor Ab_r) by repetitively attempting to play sequences which are more complex (as defined by the *PS* parameter) than the current *Ab* value (Fig. 6.13). New RUs with higher complexity than the player agent's ability are not played during that moment in simulation time, but each time an agent tries to play a RU with higher complexity than its ability, the ability parameter is increased to model the motor practice required to play a sequence. One consequence of this fact is that agents can have a memory of sequences more complex than they are able to play.

The production of rhythmic sequences may also involve the transformation of existing (stored in memory) RUs and sequences (as defined by the grammar). The transformations to which the RUs are submitted are a sub-set of the transformations used in the previous implementation of the system (chapter 5) and are explained in Sec. 5.1.1. In the current implementation, the length of the RUs is preserved, and therefore, only the

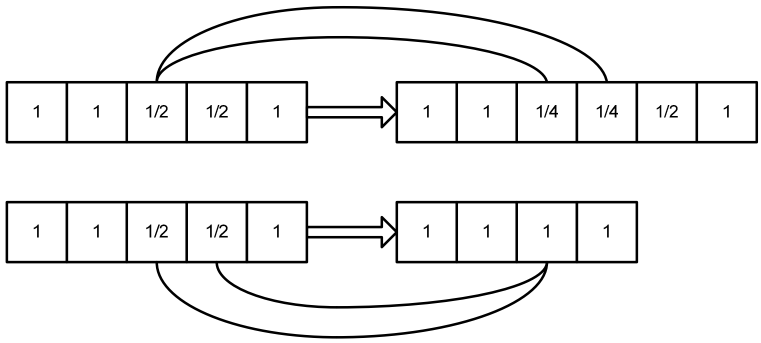


Figure 6.14: Examples of rhythmic transformations.

divide (Fig. 6.14a) and merge (Fig. 6.14b) transformations are applied.

Each player possesses a T_r parameter, which represents the probability of generating a new RU by transforming an existing RU. This parameter varies with the dynamics of the simulation and its variation is explained in Sec. 6.4. The set of parameters built into each agent is summarised in Tab. 6.2.

6.3.2 Agents' life-cycle and reproduction

In each iteration an agent has a probability of dying and being removed from the simulation. For populations with no renewal, the probability of dying is set to 0. Similarly, in each iteration an agent has probability of reproducing by generating new agents. The reproduction is asexual, and therefore we only consider mutation as a possible genetic variation. This is coherent with other studies and models of cultural evolution (Bown 2008; McElreath and Boyd 2007). An offspring of a particular agent does not directly inherits RUs or variables from the grammar, but inherits the preference values for exposure and complexity with a slight mutation in relation to its parent. The $MutExp$ parameter to mutate the exposure preference parameter is set to 0.05 and the $MutComp$ parameter to mutate the complexity preference parameter is set to 0.2 (Tab. 6.2). For populations with no renewal, the reproduction probability is set to 0.

Agent Parameter	**Description**	**Example value**
memLoss	Exposure decay constant with time	0.01
T_r	Probability of producing a transformation to a RU	0.003
a_{Tr}	Scale Factor for *changeRate*	0.1
a_{mat}	Adaptation weight for the Markov matrix	0.3
Ab_r	Ability increase rate	0.1
Ab_{init}	Initial agent's ability	1.10
Hc_{peak}	Peak value for complexity preference	3
C_{max}	Limit for complexity	5
X_r	Exposure increase rate when an RU is listened to	0.1
X_{max}	Limit for exposure	1
Hx_{peak}	Peak value for exposure preference	0.5
initialRU	Initial RU in agent's memory	[1 1 1 1]
$R_{interact}$	Radius of interaction	3
PMove	Probability of an agent moving	0.1
PRep	Probability of an agent reproducing	0.003
PDie	Probability of an agent dying	0.002
MutExp	Mutation value for exposure	0.05
MutComp	Mutation value for complexity	0.2

Table 6.2: Parameters for the agents.

6.4 Interaction

The interactions between the agents by which they develop new repertoires of rhythms takes the form of a *music game*. *Music games* are an analogy to the notion developed by Wittgenstein, in which two or more agents engage in *language games* as a way of attributing meaning to words (Wittgenstein 1958). Language games were further experimentally developed by Luc Steels to study the evolution of language with the help of robots (Steels 1997, 2003).

The interaction process involves the selection of a group of agents and the definition of their roles in the music game. The actual music game consists of the performance of rhythmic sequences by a player and their evaluation by listeners.

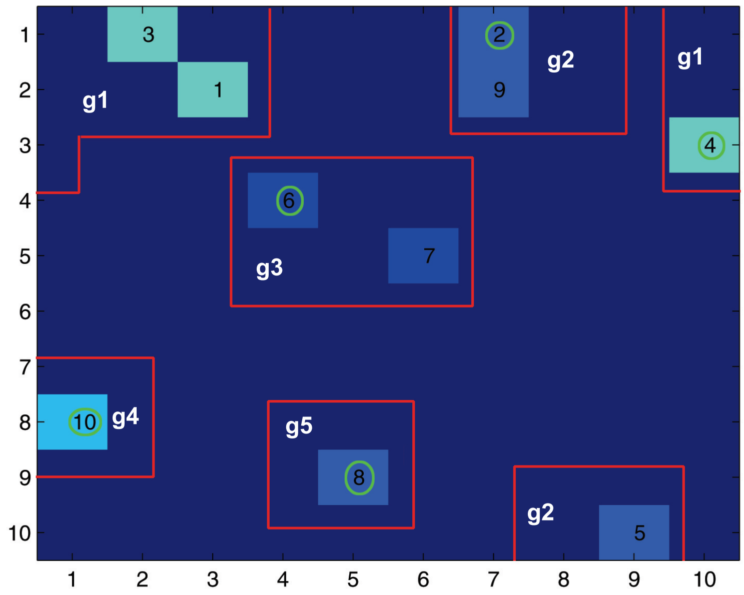


Figure 6.15: Definition of groups across society. Red lines mark the group boundaries and green circles mark the player agent.

6.4.1 Roles

Agents react only to neighbouring agents in what is sometimes called a limited perception model. This is similar to other biologically inspired models, such as swarms and other multi-agent systems where communication plays a significant role. Although agents can perceive rhythms from different sound sources due to their position, they will engage in only one music game per iteration.

This particular music game, is constituted by one agent player and several agents listening (Fig. 6.15).

The definition of an agent's role is randomly defined in the case of players and dependent on the distance to players in the case of listeners.

Choosing a random agent to be a player, automatically assigns the condition of listeners to neighbouring agents. This group will engage in the music game. The process is repeated with the remaining agents until all groups are formed (Fig. 6.15), and subsequently the games are initiated.

In the next section we explain in more detail the dynamics of the music game.

6.4.2 Music game

The music game is developed in four steps:

1. The player agent generates one rhythmic sequence and plays it to the listening agents;
2. All agents in a group listen to the rhythmic sequence and process it;
3. Listeners send feedback to the player;
4. All agents update their knowledge about the game.

These steps can be elaborated as follows:

In step 1) the rhythmic sequence is produced by generating a succession of RUs generated by the Markov matrix of probabilities. Then, the relevant RUs are retrieved from memory, organised in a sequence as defined by the Markov process, and played to the listening agents.

In step 2) incoming sequences are perceived by the agents (including the player) as described in the perception module (Sec. 6.3.1) and previously unheard RUs are included in the memory of the listeners. Subsequently the transition pattern between RUs is recorded and the grammar updated.

In step 3) the sequence is then evaluated by the listening agents in terms of exposure and complexity of the incoming RUs. These two evaluations follow a simplified Wundt curve explained in Sec. 6.4.5. All the evaluations of the listening agents are combined into two feedback values, $feedExposure$ and $feedComplexity$, which are returned to the player agent.

In step 4) the player receives the feedback values from the audience of listeners and adds its own feedback to the sequence it has just played. It then changes its internal representations in distinct ways. For low values of the complexity feedback, the agent increases the probability of making changes to the sequence. Conversely, for a high value of the complexity feedback, the inner probability of creating new RUs is diminished. The low exposure feedback causes the probabilities in the Markov transition matrix to be changed in order to be less likely to play that sequence in the future. In the case of the listeners, the incoming RUs increase the exposure value of the correspondent internal representations of the RUs (view Sec. 6.3.1). Also, the recorded transition pattern between RUs of the played sequence will affect the listeners' Markov transition matrices by reinforcing the probabilities of changes between the heard RUs, scaled by the parameter a_{mat} .

The next two sections explain in a more detailed way how the player agent deals with the feedback.

6.4.3 How feedback on complexity changes the RUs

Transformations of rhythmic units are explained in Sec. 6.3.1. Events inside RUs can either be divided or merged with other contiguous events.

The probability of changing an RU using a transformation is defined by the expression:

$$T_r = (1 - complexityFeedback) * 0.1 \tag{6.1}$$

Complexity feedback, denoted by the variable *complexityFeedback* and varying between 0 and 1, contains the feedback that the audience gives to a played sequence, along with the feedback that the player agent gives to its own played sequence.

6.4.4 How feedback on exposure affects the grammar

After having played a sequence, a player receives exposure feedback from neighbouring agents, and adds its own impression of the played sequence, with this value being computed by the same process as listeners use to compute their feedback. If this feedback value is high (closer to 1) then the Markov matrix is not changed, making it likely for the RUs to be played again in sequences with similar connection pattern in the future. Depending on the exposure feedback, the values in the columns of the Markov matrix associated to the played RUs are changed by a multiplicative value $multFact$, which is proportional to their frequency (*ruFrequency*) in the played sequence, to the feedback on exposure (*fE*) and to the parameter a_{mat} .

$$\begin{cases} multFact = 1 + a_{mat} * (1 - fE) * ruFrequency & expIndex < hedPeak \\ multFact = 1 - a_{mat} * (1 - fE) * ruFrequency & expIndex > hedPeak \end{cases} \quad (6.2)$$

6.4.5 Hedonic Values

Hedonic values are internal evaluations of the rhythmic sequences performed by the agents, these being functions of complexity and exposure (explained in detail in Sec. 6.1). There are two independent evaluations, one for complexity and one for exposure, that will affect future behaviour of the agents. The hedonic values are computed with a modified Wundt function, similar to the ones used in other systems, where agents are endowed with the capability of conducting aesthetic evaluations (Saunders 2002; Bown 2008). The function shown in Fig. 6.16 is characterised by a peak value of $X_{hedonicPeak}$ in the independent variable x, and has a normalised output between 0 and 1 corresponding to the hedonic value y. To each agent will be assigned two Wundt functions,

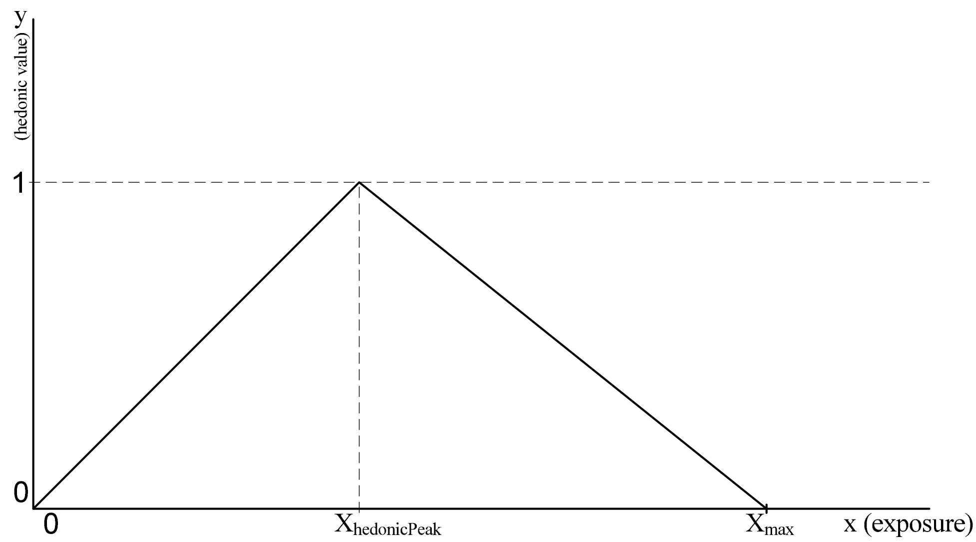


Figure 6.16: Simplified Wundt function to calculate the hedonic values for exposure. A similar curve is used for complexity.

characterised by the parameters Hc_{peak} for complexity and Hx_{peak} for exposure.

6.5 Simulation overview

The flow of information within an agent is displayed in Fig. 6.17.

One run of a simulation contains all the values for the parameters explained in previous sections of this chapter (Tabs. 6.1 and 6.2) along with the world size, number of agents and number of iterations. The summary of the parameters and their example values are displayed in Tab. 6.3.

6.6 Experiments

The study of the system was divided into two different stages. In the first stage, the agents during the simulation developed rhythmic sequences according to the mechanisms explained above. In this stage, the population in the society had no renewal, meaning that the agents interacted with no aging or reproductive constraints.

In the second stage, the simulation allowed for the agents to reproduce. The number of agents in the population was kept constant by letting one agent reproduce whenever

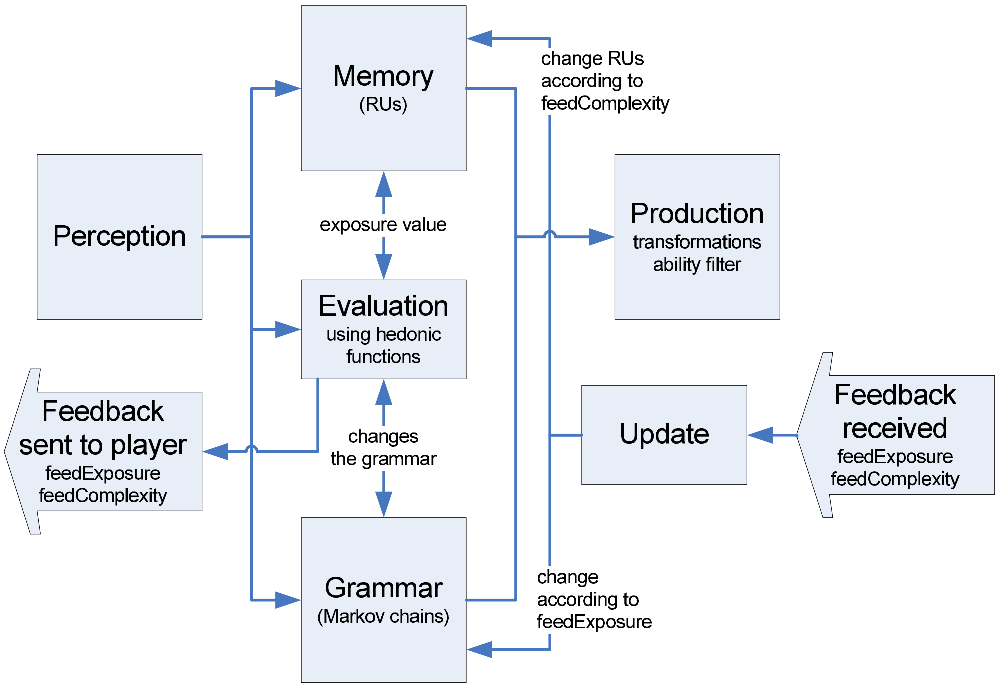


Figure 6.17: Agent's modules, including evaluation and feedback.

another agent died. This simulation conditions were aimed at studying the vertical transmission process of the rhythmic sequences. The first set of simulations studied the interaction between 2 agents, considering either constant interaction or interaction during a fraction of the time. For all the remaining simulations in study, a value of 7% of population density (number of agents/number of cells in the environment) was used. The case presented in this thesis considered 7 agents in a artificial environment (AE) of 10x10 cells. All the states in the simulation were recorded, namely the state of the agents and the musical games they participated in.

6.6.1 Interaction between 2 agents

The starting point from the simulations was the study from the interaction between two agents. In each iteration there was a player agent and a listener and the roles were selected randomly. Typically, as simulation time advances, new RUs are created by the agents using the process of transformation managed by the production module (Sec.

Parameter	Description	Example value
$ioiVec$	Interonset intervals of RU	[1 0.5000 0.5000 1 1]
$isDownBeat$	Presence of even at the beginning of the RU	1
$isCreated$	If RU was created by this agent (1/0)	1
$Iter_{in}$	In which iteration the RU was integrated	2
PS	Value of Povel/Shmulevxich complexity	13.490
n_{played}	Times played by this agent	15
$n_{listened}$	Times listened	22
X_{index}	Exposure value	0.9100
$memLoss$	Exposure decay constant with time	0.01
T_r	Probability of transforming a RU	0.003
a_{Tr}	Scale Factor for *changeRate*	0.1
a_{mat}	Adaptation weight for the Markov matrix	0.3
Ab_r	Ability increase rate	0.1
Ab_{init}	Initial agent's ability	1.10
Hc_{peak}	Peak value for complexity preference	3
C_{max}	Limit for complexity	5
X_r	Exposure increase when a RU is listened to	0.1
X_{max}	Limit for exposure	1
Hx_{peak}	Peak value for exposure preference	0.5
$initialRU$	Initial RU in agent's memory	[1 1 1 1]
$R_{interact}$	Radius of interaction	3
$PMove$	Probability of an agent moving	0.1
$PRep$	Probability of an agent reproducing	0.003
$PDie$	Probability of an agent dying	0.002
$MutExp$	Mutation for exposure	0.05
$MutComp$	Mutation for complexity	0.2
$initialAgents$	Initial number of agents	7
$maxIterations$	Number of iterations	5000
$worldSize$	Size of the virtual environment	[10 10]

Table 6.3: Simulation parameters and example values for a simulation with 7 agents in a 10x10 virtual world, with 5000 iterations.

6.3.1). The RUs are combined stochastically into sequences of rhythms that are played to each other and RUs are dropped if the exposure value of the RUs drops from a particular value. Only the RUs that are played or listened to will remain in memory.

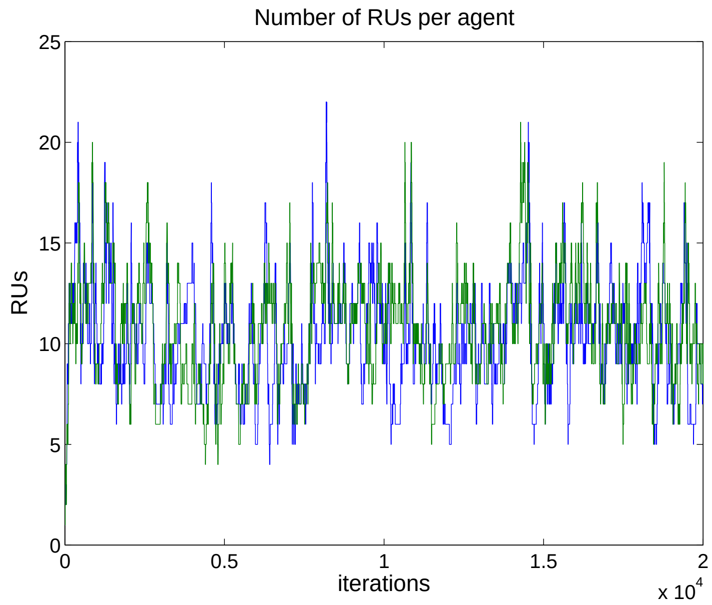


Figure 6.18: RUs in agents' memory.

Evolution of RUs and the grammar

Initially both agents have a regular pulse containing a cycle of four beats ([1 1 1 1]) which is the least complex rhythmic sequence. The number of RUs in the memory of the agents oscillates, depending on the exposure throughout the simulation (Fig. 6.18). The average number of rhythmic units for the whole length of the simulation is 7.4 for each of the agents. The grammar is a Markov matrix which determines the probability of transitions between RUs, when creating a new sequence. The first line of the matrix contains the probabilities of starting a rhythmic sequence with each of the RUs in memory and the first column contains the probability of finishing the sequence with the corresponding RU. The value in the cell (1,1) is the probability of not playing any sequence (or a sequence containing no RUs).

In Fig. 6.19 the grammar of both agents are compared at the beginning, in the middle and at the end of the simulation. The values above and left to the black lines (upper left quadrant) correspond to common RUs existing in the repertoire of both agents which are placed in order to be able to compare the values contained in the grammar.

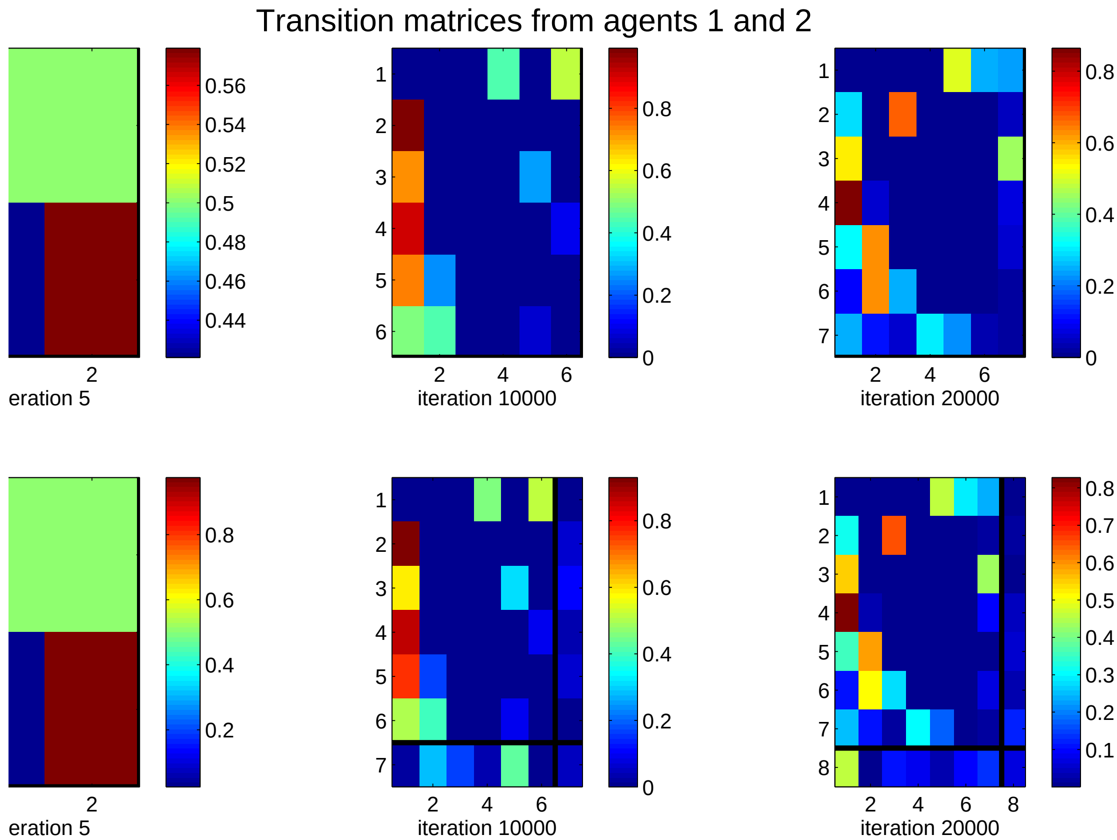


Figure 6.19: Values of the Markov matrices of two agents compared at iterations 5, 10000 and 20000. The values above and left to the black lines (upper left quadrant) correspond to common RUs existing in the repertoire of both agents which are ordered in order to be compared. The remaining columns correspond to RUs particular to each agent. From the analysis of the figure it can be seen that they are highly similar both in the common RUs as well as in the grammar, meaning that they interact frequently.

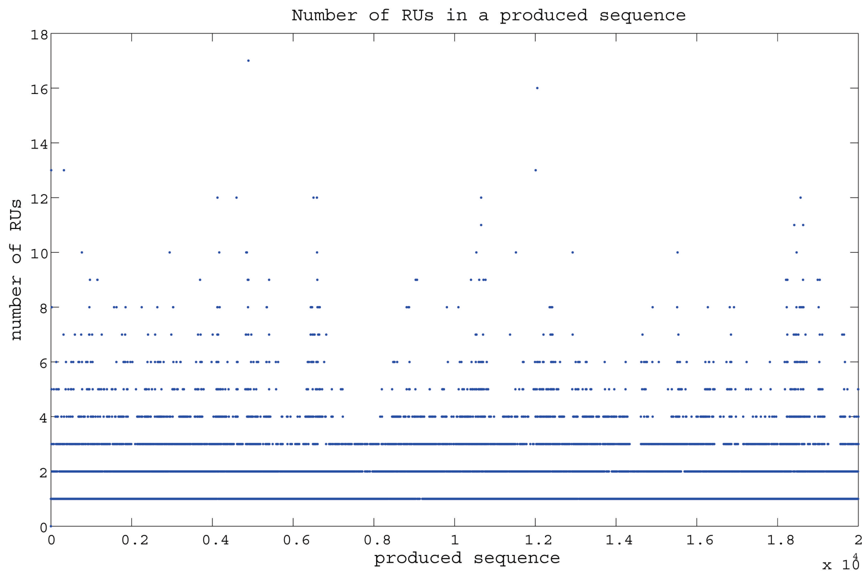


Figure 6.20: Number of RUs per sequence played in all iterations of the interaction between the two agents.

This stochastic process creates sequences of different lengths (Fig. 6.20). In order to avoid perpetual states, if the generated sequences contain more than 20 RUs, then the values in the first column are multiplied by 1.2, and then normalised, forcing the sequences to end sooner.

The complexity of the repertoire of the agents increases as new RUs are being created. Ability is a scalar parameter that influences the production of rhythmic sequences and its variation is described in Sec. 6.3.1. In Fig. 6.21 the evolution of complexity and ability are displayed for the repertoire of a single agent. The the stronger black line is the agent's ability to perform complex rhythmic sequences. The solid grey line corresponds to the average values of complexity in the agent's repertoire. The two dashed grey lines are respectively the maximum and the minimum values for complexity in the repertoire.

Often, it is the case that the agents have RUs in memory which they are not able to play, this being the reason why the line of maximum complexity overcomes the ability line.

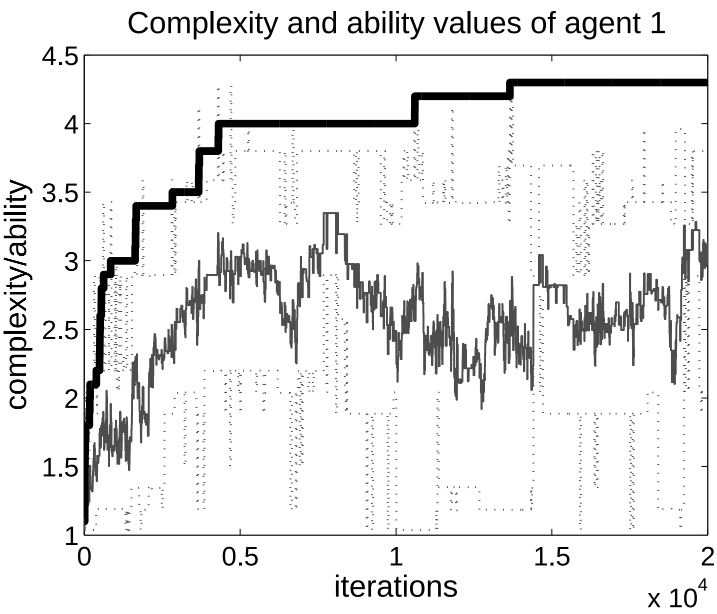


Figure 6.21: Complexity and ability values for agent 1; the stronger black line is the agent's ability to perform complex rhythmic sequences; the solid grey line corresponds to the average values of complexity in the agent's repertoire; the two dashed grey lines are respectively the maximum and the minimum values for complexity in the repertoire.

The agents develop complex repertoires as they transform existing rhythms in their memory or learn rhythms from other agents. The ability of the agents evolves with the repeated playing of more complex rhythms. Fig. 6.23 shows how ability increases as they try to play more complex sequences.

In each music game, each agent, including the player agent, computes two feedback values for *exposure* and *complexity*. The values for all the agents are summed and an average is taken, giving origin to the two feedback values of the group delivered to the player agent. In Fig. 6.24, it is possible to see the evolution of the feedback values received by player 1, along with a polynomial interpolation curve to better observe the trend. As both agents are constantly in the same interaction group, the graphs for agent 2 display the same trend.

In each iteration, the agents have a probability of creating new sequences by transform-

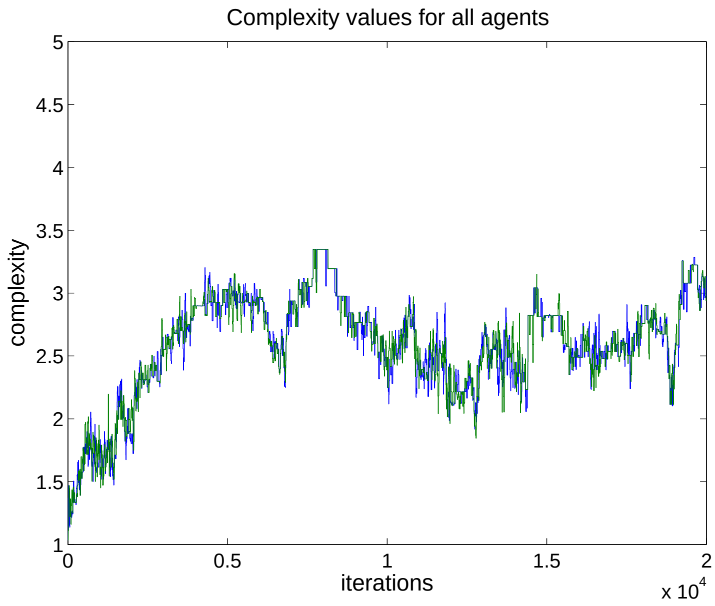


Figure 6.22: Complexity values for both agents.

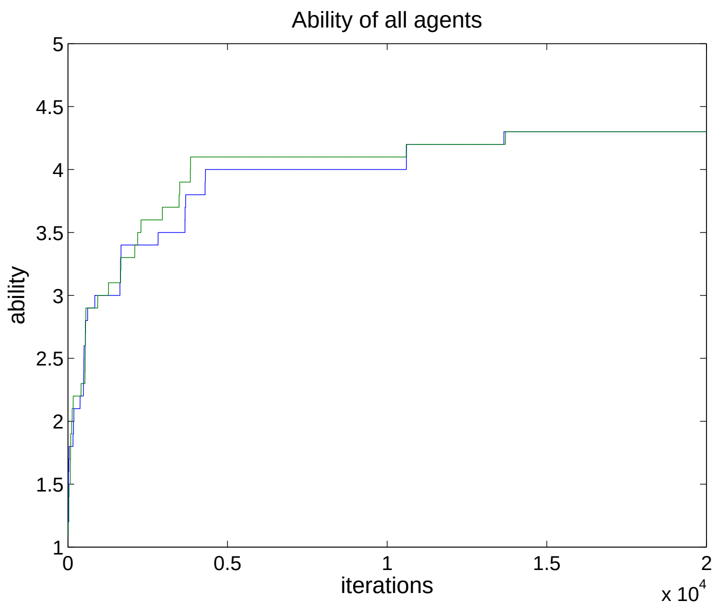


Figure 6.23: Ability values for both agents.

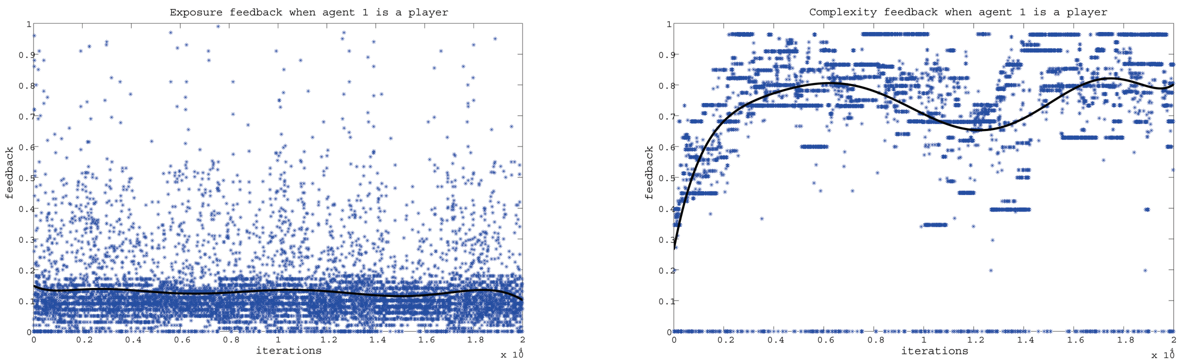


Figure 6.24: Exposure and complexity feedback values received by agent 1; The black lines are polynomial interpolations fitting the feedback values in blue.

ing existing ones. In Fig. 6.25 the variation of the change rate parameter, T_r, for all the agents is displayed. This parameter is affected by the complexity feedback of the listening agents. For this simulation, the parameter varies from 0, meaning no new RU is created, to a probability of 0.16 of creating a new RU in each iteration. It can also be seen that the agents have a higher probability of creating new RUs in the beginning of the simulation which then decreases as simulation time advances.

By comparing Fig. 6.25 with Fig. 6.24b) it is observed that there is a local minimum for the interpolated complexity feedback at around 12000 iterations, which in turn will influence the rate of the creation of new RUs by the agents. It was observed that a higher rate of transformation did not necessarily result in a growth in complexity (Fig. 6.22).

World size

Changes in the world size will reduce the amount of time the agents will be in contact with each other. In the case of 2 agents interacting 3% of the time (Fig. 6.26), changes in the amount of common RUs and on the values for the grammar will be observed. In Fig. 6.27 it is observed that the 2 agents will have some common RUs at iteration 10000, due to a period of interaction in previous iterations, but will have little in com-

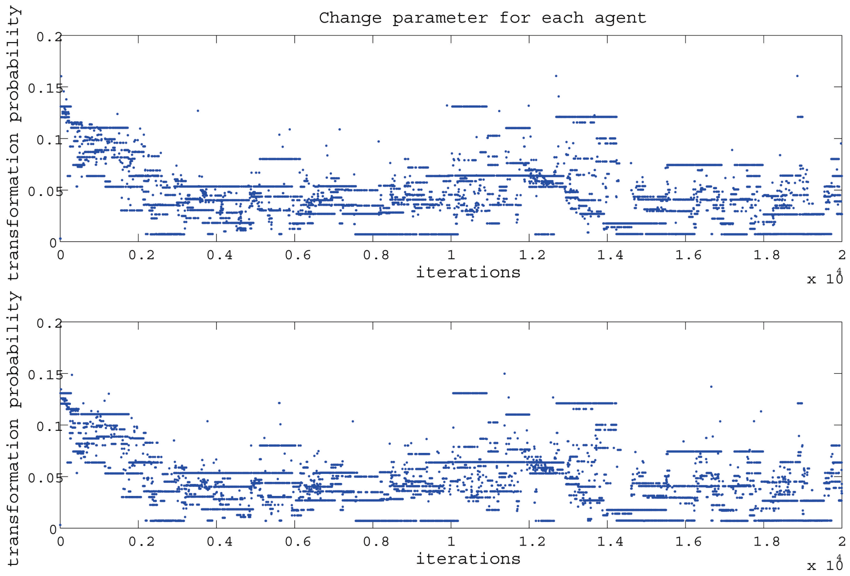


Figure 6.25: Probability of creating a new sequence in each iteration.

mon in the grammar. At the end of the simulation, the agents will no longer share RUs. It is observed that although the complexity of both agents will not follow exactly the same curve as in the constant interaction situation (Fig. 6.22), the short periods of interaction are enough to keep the complexity of the RUs in memory in similar trend (Fig. 6.28).

6.6.2 Interaction in a small community

For a simulation with 7 agents in an AE of 10x10 cells and 5000 iterations, the number of sequences emerged is presented in Fig. 6.29.

Diversity across space and time

The first target of investigation was how the process of cultural evolution shaped the repertoires of the agents in a society without population renewal.

In order to investigate the diversity of repertoires across space and time, three different moving conditions were used:

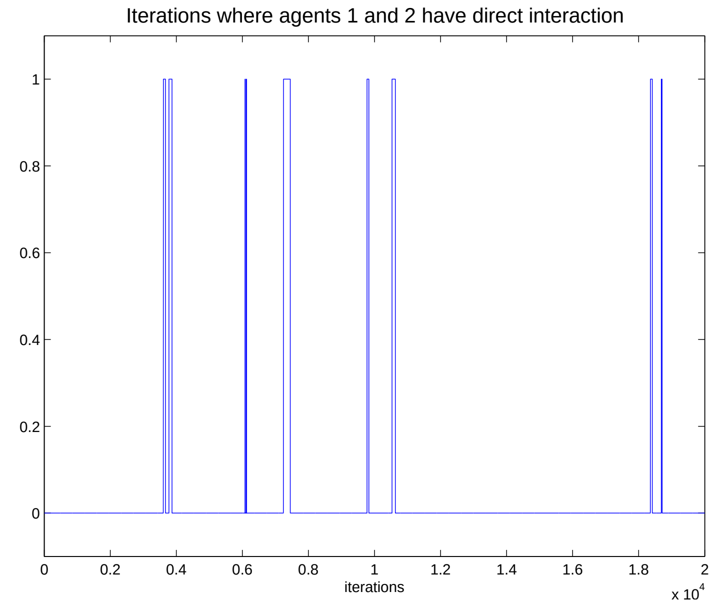


Figure 6.26: Simulation time steps where the the two agents interacted.

1. Agents do not move;
2. Agents have a probability of 1 of moving into another cell in each iteration;
3. Agents have a low probability of moving into another cell in each iteration.

The **synchronic analysis** studies the state of a system, or several systems, at a particular point in time. In this particular case, this analysis is aimed at verifying the similarity of the repertoires and the evolution of the grammars.

The **diachronic analysis** studies the evolution of a system by looking at different points in time. These two analysis are accomplished by doing a hierarchical cluster analysis (explained in Sec. 5.1.2) using the similarity between the agents, and also by looking at their position in space.

For the no-movement condition the results can be seen in Figs. 6.30 and 6.31. In Fig. 6.30a), two clusters, containing agents [1 2 3 4 5] and agents [6 7], constitute each a rhythm exchange network. The radius of interaction is 2 neighbouring cells, and although agent 3 and 2 are not in direct contact, they both are in contact with agent 5,

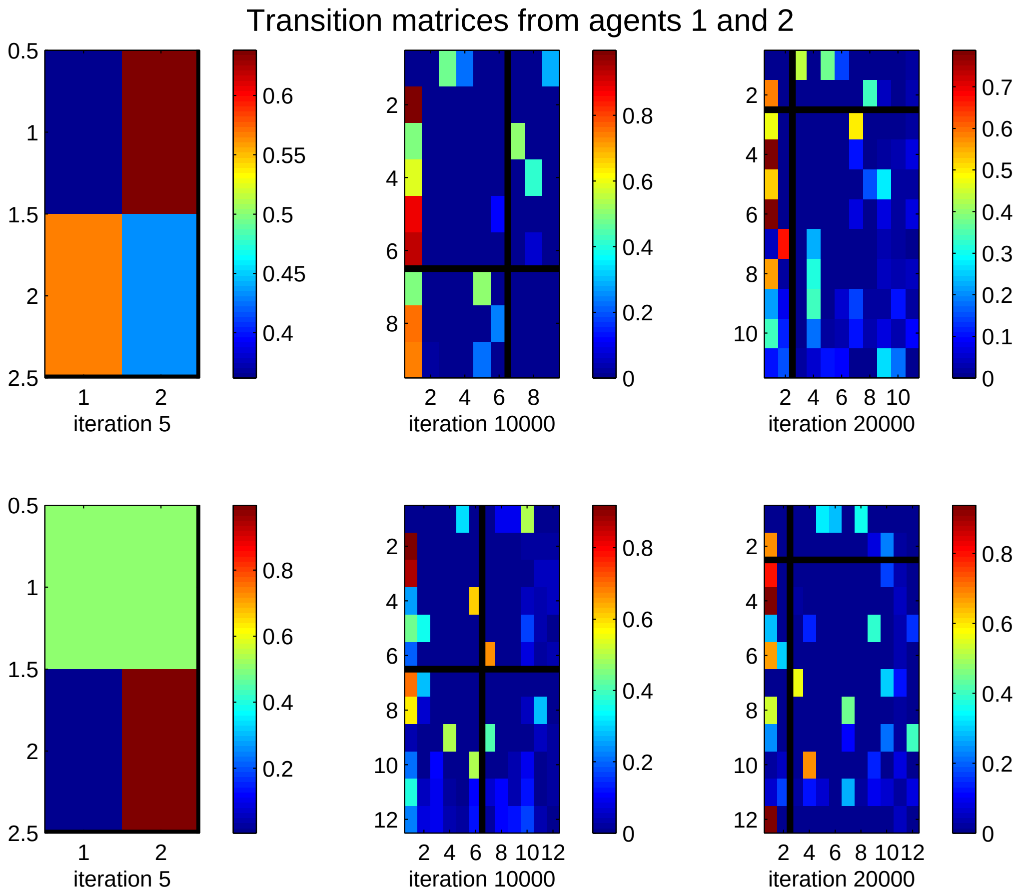


Figure 6.27: Values of the Markov matrices compared at iterations 5, 10000 and 20000. Simulation with two agents which interacted 3% of the time.

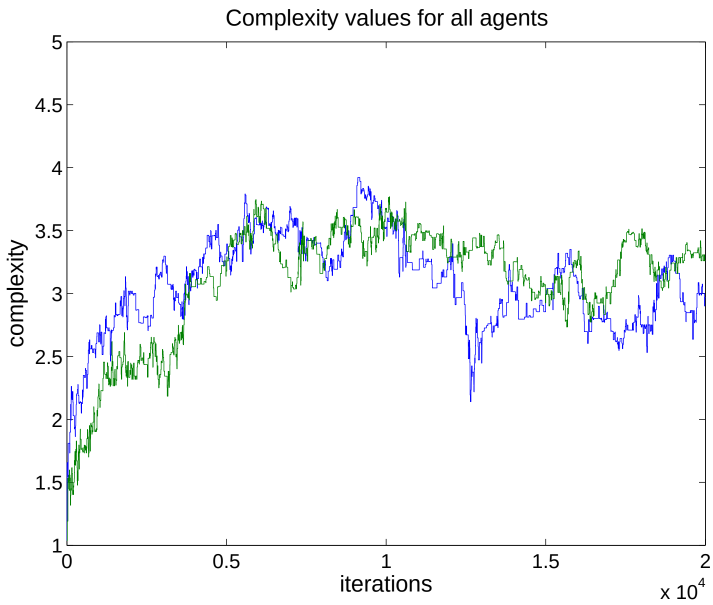


Figure 6.28: Complexity values for both agents.

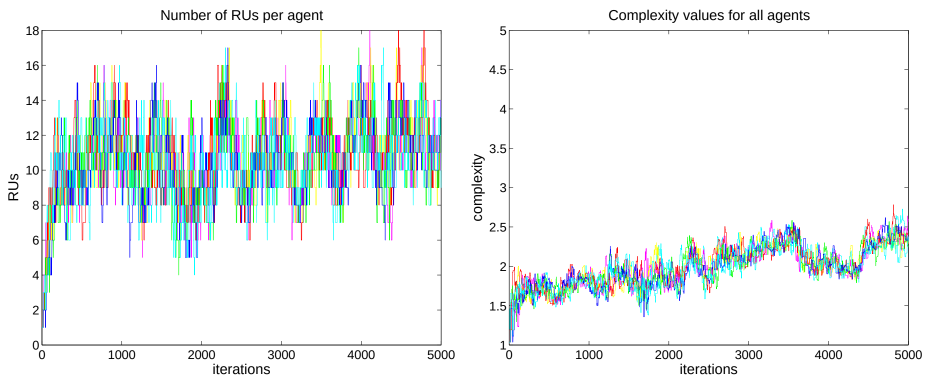


Figure 6.29: Number of rhythms and complexity values for 7 agents.

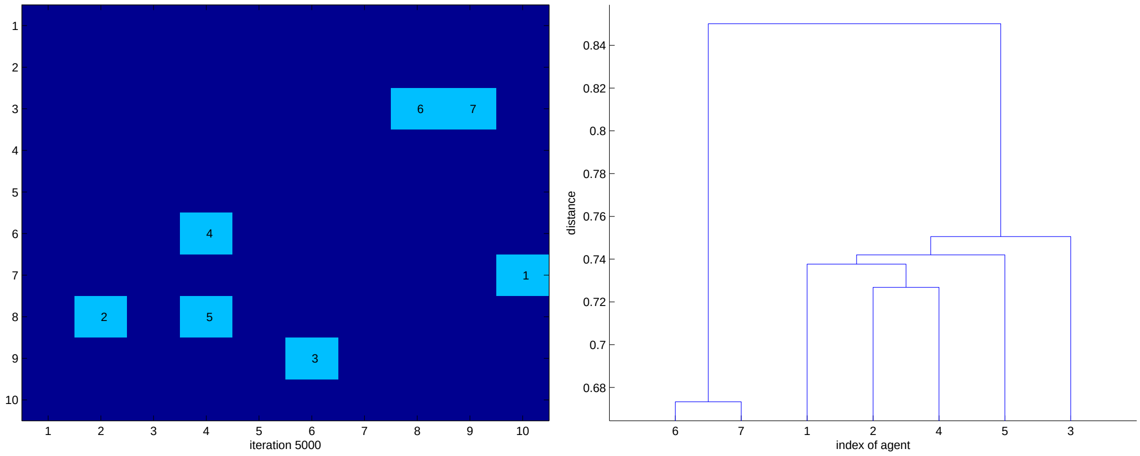


Figure 6.30: Synchronic analysis of 7 static agents.

which will guarantee that rhythms created by agent 3 may be learned by agent 2, and mediated by agent 5, when it becomes a player. As the artificial environment is a torus, agent number 1 will be in contact with agent 2. The analysis of the dendrogram shown in Fig. 6.30b), constructed by considering the distances between the repertoire of the agents taken at iteration 5000, shows that the repertoires of agents within the network are consistently related and the distance between the networks is considerably larger than the distance between individual repertoires within the network.

Using a diachronic analysis of the iterations 2500 and 5000 (Fig. 6.31), we can observe for each of the iterations considered, that the network structures are maintained within the compared repertoire of the agents, and that the temporal difference causes larger changes in repertoire than the spacial difference.

In the constant movement condition shown in Figs. 6.32 and 6.33, all the network structure disappears in the synchronic analysis, revealing no contribution of the spatial distribution of the agents. A dendrogram with a ladder shape means that no clusters were formed. It can also be observed that the agents have themselves changed more the repertoires across time than the individual differences between individuals at one time step.

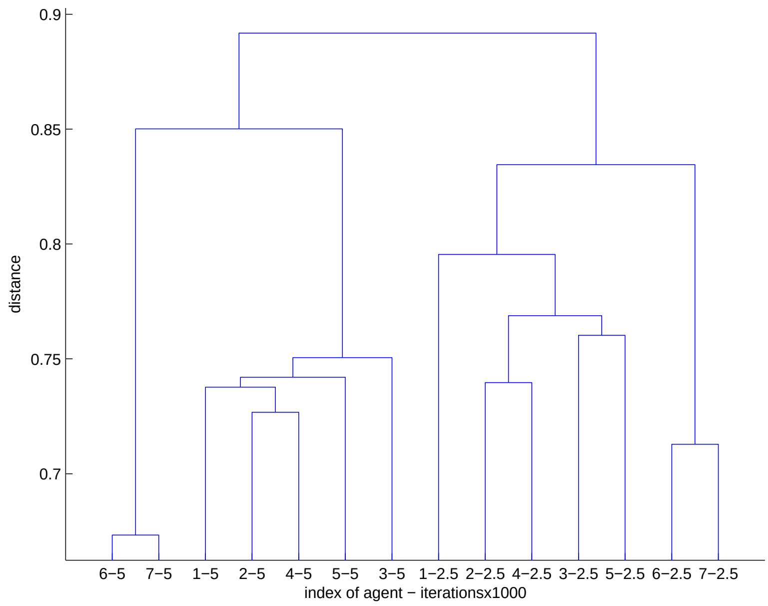


Figure 6.31: Diachronic analysis of 7 static agents.

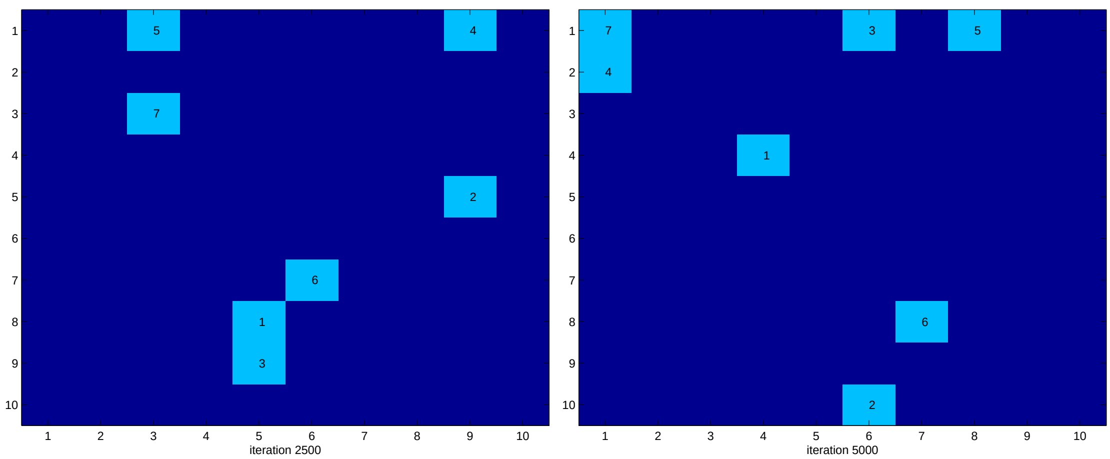


Figure 6.32: Snapshots of the world map with the 7 agents in the constant movement condition.

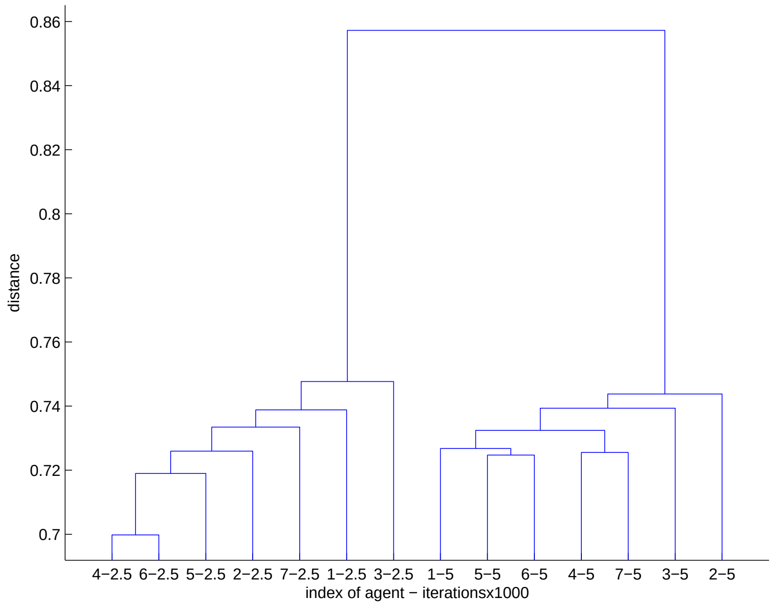


Figure 6.33: Diachronic analysis of the 7 agents in the constant movement condition.

When the agents have a probability of moving of 0.01 (Figs. 6.34 and 6.35), meaning that they move to an adjacent cell in every 100th iteration, we can observe slightly better clustering than in the constant movement case.

6.6.3 Population renewal

In this section, the results are presented for a typical simulation with population renewal. The population size is kept constant by letting one agent reproduce whenever another agent dies. Each agent is initialized with a probability of dying of 0, growing linearly afterwards by $PDieIncrease = 5e-7$ in each iteration. The main simulation had 10000 iterations, 7 agents and random preferences. The tree with the agent's life-cycle is displayed in Fig. 6.36.

The complexity evolution from all the agents is shown in Fig. 6.37. In the picture, it can be observed that newly born agents evolve complexity quicker than the initial agents, conditioned only by the ability parameter.

Fig. 6.38 shows the evolution of the complexity and exposure peaks. This led to the

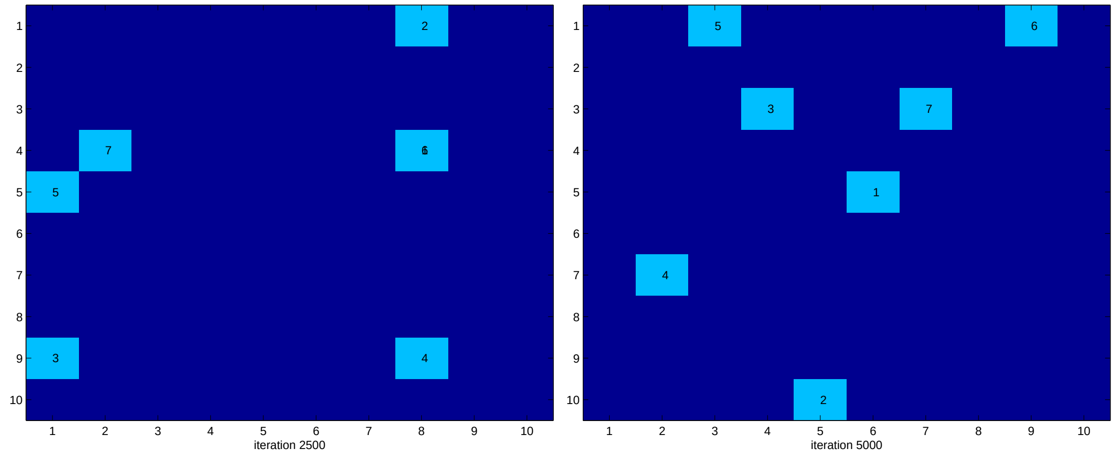


Figure 6.34: Snapshots of the world map with the 7 agents moving with a probability of 0.01 in each iteration.

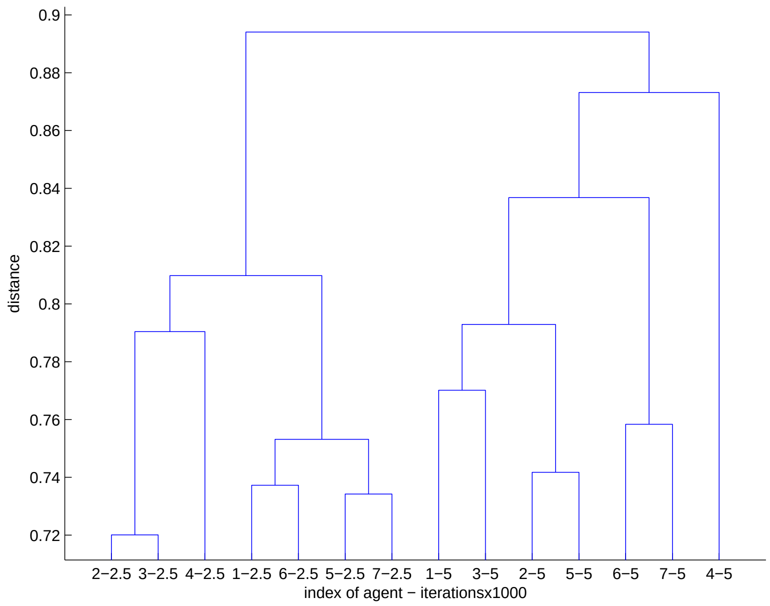


Figure 6.35: Diachronic analysis of the 7 agents moving with a probability of 0.01 in each iteration.

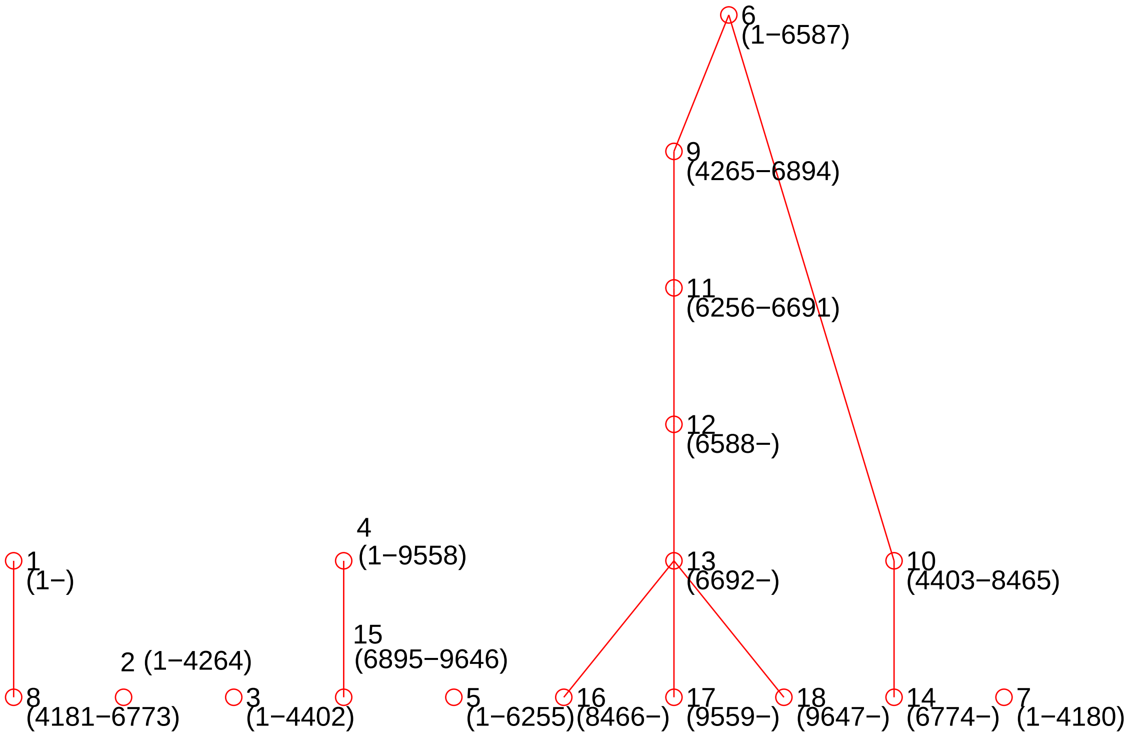


Figure 6.36: Tree indicating the life cycle of the agents. Agents are numbered and linked to their offspring. Iterations of birth and death are indicated between parentheses.

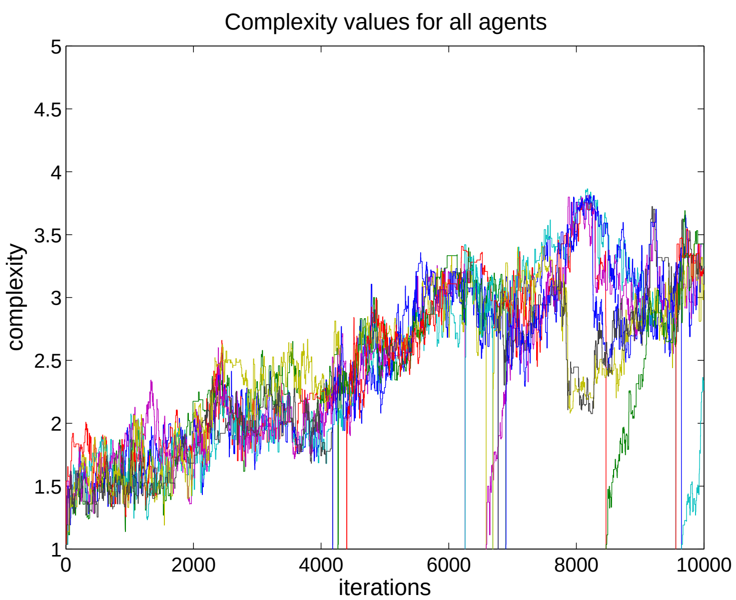


Figure 6.37: Complexity values for all the agents.

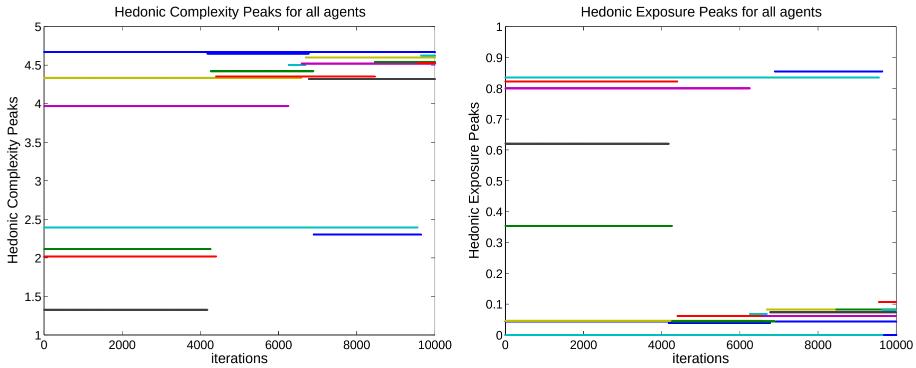


Figure 6.38: Hedonic complexity and exposure peaks for all the agents.

observation that hedonic peaks, although initially random, converge to clusters as new agents were born. The gradually more homogeneous population is caused by the inheritance by the offspring of the characteristics of the parents.

6.7 Summary

In this chapter, the second implementation of the multi-agent system A-Rhythm was presented and the results were described. This system is a new version of the previous, incorporating some of the elements of the previous A-Rhythm, such as the virtual environment, an essentially identical representation and some aspects of the interaction algorithms. New to this system is the incorporation of a grammar based in a Markov process, which increased the generative power of the system. Also new to the system was the Povel-Shmulevich measure of complexity (PS-Measure) used by the agents, which, due to constraints from the measure, limited the rhythmic units to the resolution of an eight-note and to a constant length of 4 beats on the tactus level (corresponding to 16 eight-notes). Still, these constraints enable the encoding of 65,536 rhythmic units, thus providing a high level of diversity.

From the field of bioaesthetics, a bell-shaped relationship between the intensity of a stimulus and the hedonic values seems to be determinant in artistic appreciation by individuals. The work done by Berlyne on the subject was followed in this thesis, by incorporating this notion in simulations, in order to test how these individual evaluations had an effect in the society.

The system was studied in terms of the synchronic and diachronic diversity of the rhythmic material produced in a a simulation. Different conditions were analysed, namely, the size of the world, number of agents, frequency of movement and population renewal.

The results show that shared rhythmic grammars evolve based on the interchange of rhythmic sequences between the agents. The learning of the grammar happens without direct access to the internal representation of other agents, but by listening to sequences created by their grammars. In the simulations with two agents, each agent was able to evolve a grammar and its similarity to the grammar of other agents depended on the number of interactions. The number of interactions varied with the size of the world, range of interaction, or probability of moving. Lower number of interactions revealed both a lower number of shared rhythmic units and substantial differences in the transition of the states. The synchronic and diachronic analysis were performed through a hierarchical clustering analysis in a society with 7 agents under different conditions of movement. These analysis revealed that although both space and time have influence in the variation of the repertoires, time seemed to have a stronger influence in the variation. It was also observed, that whenever agents moved constantly little or no clustering was observed at a particular time instant, whereas clusters of similar repertoires are found when agents are static or have a small probability of moving. Finally, in the case of population renewal, it can be observed that new agents evolve repertoires of complex rhythmic units rather quickly in comparison with the overall complexity trend of the society.

6.8 Discussion

A-Rhythm is a multi-agent system for the evolution of rhythmic structures, based in an evolutionary algorithm using music games between the agents. The rhythms are represented in sequences of inter-onset intervals in the memory of the agents, which can be evaluated for their complexity. The representation also enables similarity comparison between the repertoires of the agents. Transformation of the rhythms is made at the level of the individual agent, but it is the interaction algorithm and parameters of the simulation that condition the evolution of the repertoires. The exposure of the agents to the rhythms, as a direct result from the process of interaction, plays a significant role in the sustainment of those rhythms in the society. In both implementation it was shown that the rhythmic structures evolved in the individual agents, not in a random way, but in connection with neighbouring agents. A more detailed discussion is produced in chapter 8.

Chapter 7

A-Rhythm a computer music composition system

The previous chapters introduced the architecture and behaviour of A-Rhythm, and discussed the significance of the model. This chapter illustrates how A-Rhythms can be used by composers interested in computer-aided or generative music composition using evolutionary models.

7.1 Musical relevance

Evolutionary computation and Artificial-Life (A-Life) have proved to be a good source of material for composers and performers. The works of Miranda (2003), Beyls (1989), Blackwell (2007), Dahlstedt (2004), Kirke (2011), Gimenes (2009) and Biles (2007) have shown that the guided process of evolution makes sense in terms of applying into creative musical environments.

Total randomness and total predictability are two extremes that rapidly trigger a sensation of dullness into the listener. As it was briefly mentioned in chapter 3, the life like qualities of an evolutionary algorithm seem to achieve a diversity sweet-spot between these two extremes. It may also be the case that human beings have evolved to pay attention to sounds and patterns of nature, and evolutionary algorithms try to model natural processes. The key factor seems to be the process of self-organisation and emergence of behaviour.

In the case of the topic of this thesis, which is the study of cultural evolution, it also means trying to listen to imaginary cultures. Whereas History follows a continuous path though a past that never changes, A-Life provides alternative stories for some type of initial conditions, enabling us to look into how different the past might have looked, and even to see a little bit of the future.

7.2 Music and self-organisation

Self-organisation has played an important role in many musical environments. The music of West-Africa is noted for its interleaving patterns of melodic and rhythmic lines. There is no centralised control of the music and every performer has to attentively listen to the what the other musicians are doing (Arom 2004).

One composer for whom self-organisation has played an important role is Steve Reich. In his early piece *It's Gonna Rain* "two loops are lined up in unison and then gradually move completely out of phase with each other, and then back into unison" (Reich 2004 p.20). Commenting on the emergent composition, Reich says: "As you listen to the result, you seem to hear all kinds of words and sounds that you have heard before, and a lot of psychoacoustic fragments that your brain organises in different ways, and this will vary from person to person" (Reich 2004 p.21). This technique was defined by Reich as *phasing*, and it was used extensively in its early works (Tucker 2006). As in most self-organised systems, the process of composition is extremely simple, but the results are quite complex.

In the second half of the 20th century a new music current apeard: minimalism. This process involves the use of simple music patterns that are repeated, shifted and processed, leading to complex harmonies, rhythms and transitions. This approach was pursued in the United States by Steve Reich, Terry Riley, La Monte Young and Philip

Glass, and in Europe by Louis Andriessen, Michael Nyman, Henryk Górecki and Arvo Pärt.

In 2001, Brian Eno gave a lecture at the ICA in London about John Conway's Game of Life, a cellular automata where rules of interaction between cells lead to the emergence of unexpected moving shapes, and related this approach to the generative musical pieces (Toop 2006 p.241). Later, Eno stated that Steve Reich's pieces and Terry Reily's *In C*, or "anything where the composer doesn't specify a thing from the top down" were predecessors of generative computer models. He goes on saying: "Generative music is like trying to create a seed, as opposed to classical composition which is like trying to engineer a tree".

The earlier works on cellular automata and music can be traced back to 1986 when Xenakis composed his orchestral piece *Horos*. This process is documented in Hoffmann (2002). In the late 80s three composers were using cellular automata in their music pieces: Peter Beyls (Beyls 1989), David Millen (Millen 1990) and Eduardo Miranda (Miranda 2001).

Peter Beyls developed a composition system where pitches are assigned by the user to the cells in the space. Pitches that are active in consecutive iterations are linked, hence defining the rhythm of the notes. Later, he expanded the mapping possibilities of the system (Beyls 2004). *Drake Circus* consists of a virtual guitarist playing a cellular automaton generated piece. A computer program runs the automation and communicates to an algorithm specialised in harmonic articulation.

At the same time, Eduardo Miranda developed the CAMUS (Cellular Automata Music) system (Miranda and Kirke 2010). This software is based in a two dimensional cellular automata, where the cartesian space is mapped to the intervals between three note chords. *Entre o Absurdo e o Mistério*, for chamber orchestra, is a piece by Eduardo Miranda entirely composed with *Camus*.

Palle Dahstedt developed *Ossia*, a system that generates score fragments selected with a genetical algorithm. In the installation shown in 2002 at the Gaudeamus Music Week in Amsterdam, an entirely new piano composition was generated every three minutes and performed on the piano (Dahlstedt 2004).

Through the process of self-organisation, composers were able to produce pieces that have been performed in a variety of contexts and proved to be musically interesting. In this section, some pieces were described, which, along with the systems described in chapter 4, represent good examples of bottom-up processes used in algorithmic music composition.

7.3 Sonification and mapping

An artificial system can be used to create new rhythms which will be later used in an unrelated way to the progress of the simulations, with the sole goal of producing musical variation. But one of the challenges for the pieces produced with A-Rhythm is to reflect the simulation into the compositional process, in similar ways to the process of sonification. To accomplish this, the composer is faced with a problem of mapping.

Beyls (2004) defines mapping as "the establishment of a sensible connection between two areas of activity which creates meaning to a human or machine perceiver".

One concern in the mapping is the fact that the volume of data generated by a simulation is too large for all the material to be used. The most immediate step is to choose some form of selection.

The goal is to produce pieces that are able to navigate through the generated material and trace the evolution of the rhythms within the artificial society. This can be accomplished by using parameters such as complexity, intra- and inter-community similarity, and other possible features such as popular rhythms, or agents that produce rich musical output.

7.3.1 Mapping dimensions

The evolutionary simulations have several well defined dimensions and some other more implicit dimensions. For the purpose of this thesis, the pieces should reflect some of the dimensions in the simulation.

The *direct* dimensions are the ones which concern space, simulation time and the agents life cycle: simulation time; agent variables (identification, role, age, parental lineage); space (position, direction).

The *indirect* dimensions are the ones that concern the cultural substrate of the agents and their musical behaviour: complexity; exposure; repertoire size; similarity.

Music has also several dimensions. Some of them are subjective qualities with contributions from many physical variables represented in more or less quantifiable scales, and others are rather difficult to quantify. Examples of these dimensions are: duration; pitch; timbre; loudness; envelope; articulation; accent; dynamics; melodic contour; harmony.

The first four musical variables have a strong correlation with a physical variable, but there are contributions from other variables to create the perceptual result (Rossing et al. 2001). Duration is measured in seconds but it is dependent from the envelope of the sound. Pitch is dependent on the frequency of the first partial but it also dependent on the timbre, or spectral content. Loudness is dependent on the energy of the sound but also on the envelope. The last four dimensions from this list are not attributes of single notes but result from the context of a group of notes.

For the purpose of the simulations in this thesis we consider only individual notes. Some musical dimensions are mapped into variables of the simulations, such as duration, timbre, pitch and dynamics, and some others become emergent properties or are arbitrarily defined by the composer.

Rhythm is the object of research and this dimension is reflected by taking the samples of the played rhythms during the interaction between the agents. The remaining mapping process is explained in the sections concerning each of the pieces.

7.4 Flickering Pleiades

In chapter 5 we have seen how a group of rhythmic agents can evolve a shared repertoires of rhythms of increasing complexity. The simulations and, within those, several modes of transmission were studied.

“Flickering Pleiades” is a piece for a solo percussionist playing seven different sources of unpitched percussion. It follows the evolutionary process used in the simulations for the *complexity* algorithm and the underlying music structure is constructed by taking samples of the rhythms played during the simulation.

Each percussion instrument is assigned to one of the agents. When two or more agents meet in the same cell at a given time step, one agent is assigned with the role of the player, and the others with the role of listeners. The rhythms used by the players during the simulation form the structure of the music piece in its simplified version presented in this chapter. These rhythms are marked in with an *f* for *forte*, a standard music dynamics notation for playing a loud passage, whereas the remaining rhythms selected from the neighbouring agents' repertoires are marked with a *p* for *piano*, meaning softly (Fig. 7.1). The rhythms from the repertoire of the listening agents are included in order to have an idea of the number of agents in the proximity of the playing agent, and how similar are their repertoires.

The simulation used in this piece contained 7 agents with the parameters summarised in Tab. 7.1.

The virtual environment has 9 cells and multiple agents within one cell form a group. As explained in chapter 5, in each group one agent plays and the others listen to it. If the

Parameter	Initial values	Description
initialAgents	7	Initial number of agents
maxIterations	5000	Number of iterations
worldSize	[3 3]	Size of the virtual environment
complexWindowRadius	0.8	Radius for the complexity window
transThresh	5	Threshold of transformation of sequences

Table 7.1: Parameters used in the *complexity* algorithm to create the piece "Flickering Pleiades".

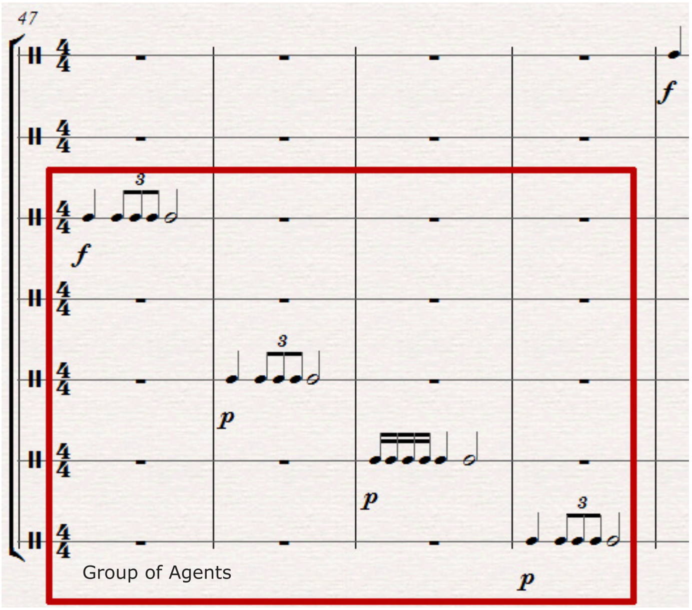


Figure 7.1: Bars 47-50 of "Flickering Pleiades" with player agent marked as *forte* and neighbouring agents with *piano*.

complexity of the played rhythm falls within the window of complexity of the listening agent ($[AvComplexity - complexWindowRadius; AvComplexity + complexWindowRadius]$ and $complexWindowRadius$ is defined in Tab. 7.1), then the rhythm is incorporated. The rhythms are transformed by the player agent after having gone passed the transformation threshold ($transThresh$ defined in Tab. 7.1). This value is a quotient between the number of times the rhythm was played and its complexity.

The full length of the piece is printed in the appendix and contains also rhythms taken from the memory of the agents that were interacting with player agents at that moment.

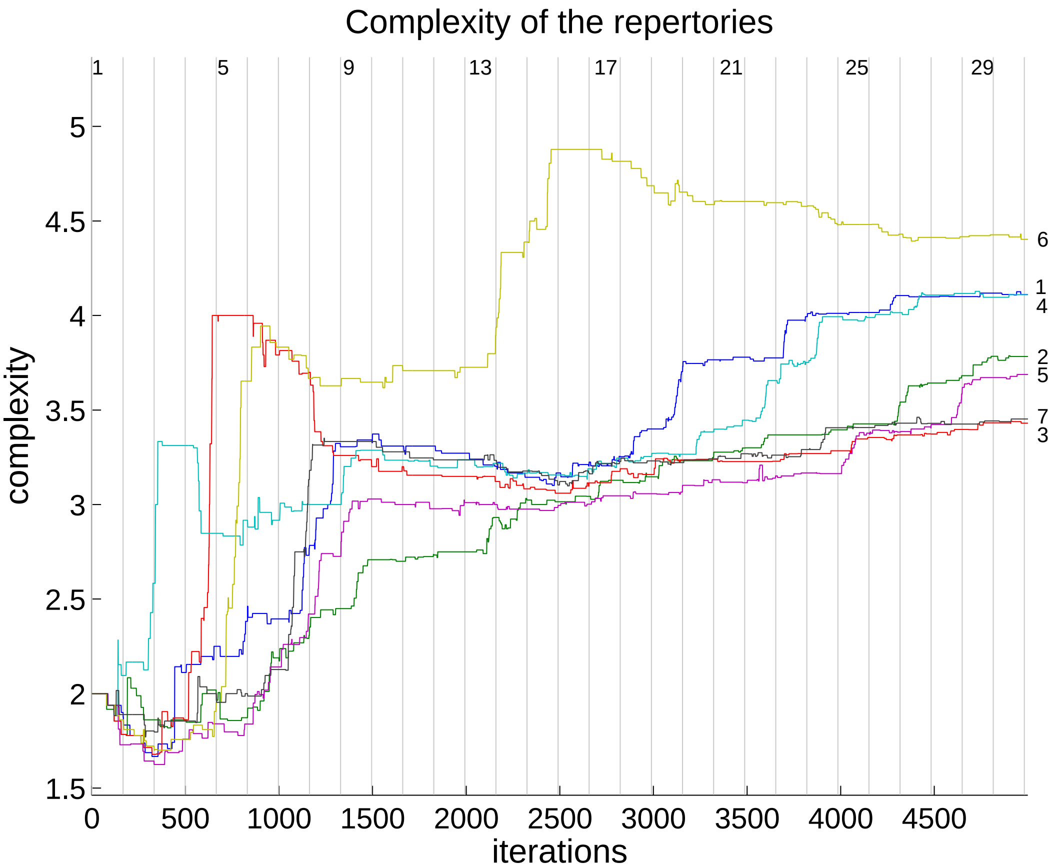


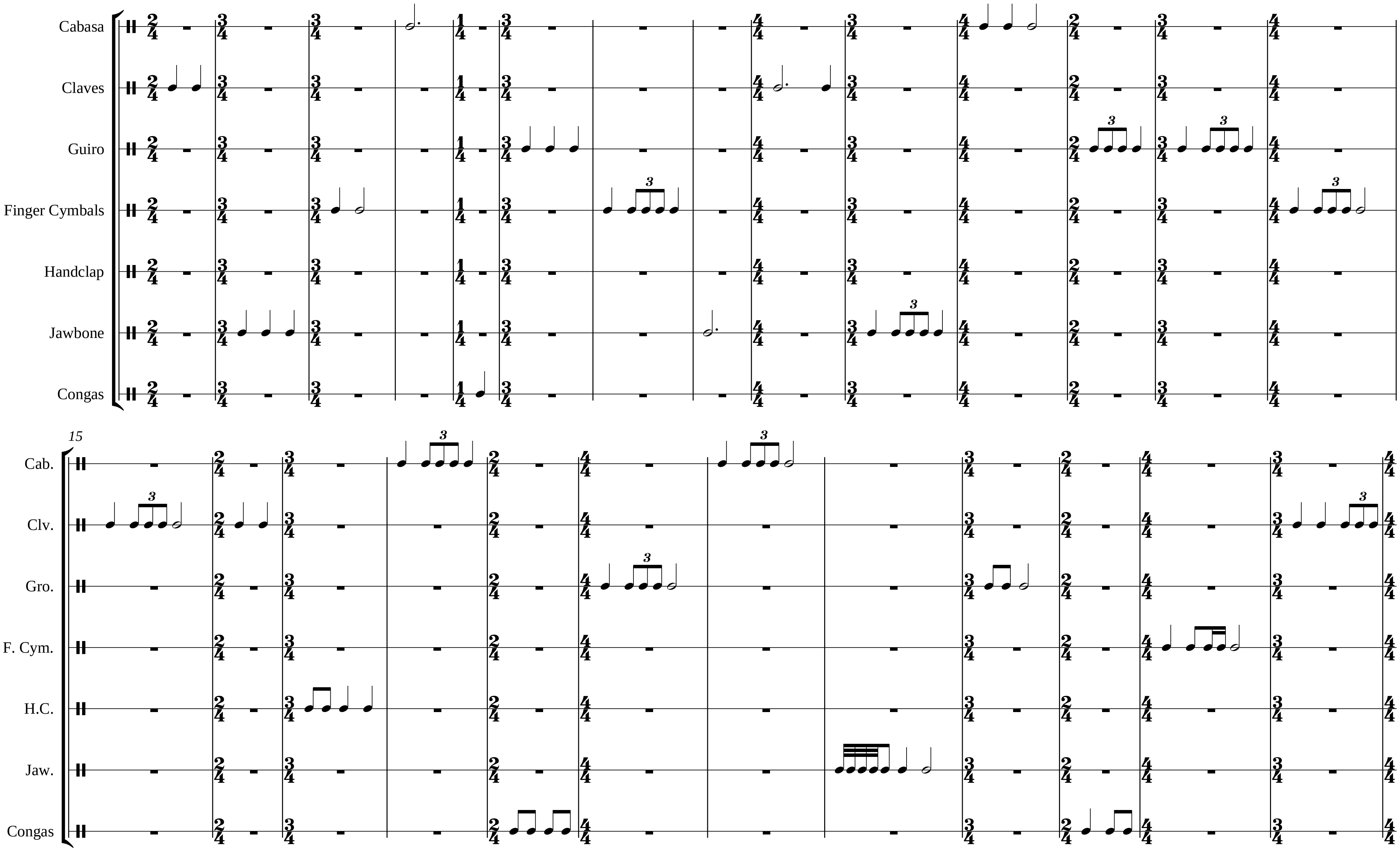
Figure 7.2: Evolution of complexity for the simulation originating Pleiades. The vertical lines mark the points where samples of the simulation were taken. The numbers on the right represent the agents' indices.

Fig. 7.2 displays the evolution of complexity in the society of agents that generated the rhythms. The vertical lines represent the points sampled in the simulation and the used rhythms are the ones that were played by the agents, having a direct correspondence to the bar number in the simplified version.

A-Rhythm is open for use by other composers to generate material for their own compositions. The system has been used by Eduardo Miranda in the 2nd movement (Evolve) of the piece "Mind Pieces" premiered in 2011 at the Contemporary Music Festival in Plymouth. The original rhythms used in "Mind Pieces" were composed using the *popularity* algorithm (chapter 5). In the program of the piece Miranda wrote: *I started with a set of computer-generated rhythms, which were generated by means of a simulation of evolution and transmission of rhythmic memes; memes are the cultural equivalent of genes. These rhythms, which are played on the snare drum, form the backbone of the whole movement*.

Flickering Pleades

Joao Martins (2011)




27
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas


38
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas

7.5 Music Games

In chapter 5, the first version of A-Rhythm was shown, and in chapter 6 new features were added to the agents; namely, a probabilistic grammar and perceptual module that rates rhythms. Also in chapter 6, a study of the horizontal vs. vertical modes of transmission was conducted.

The second composition presented in this thesis shows the evolution of structured sequences of rhythmic units and how the new repertoires of rhythms evolve in terms of complexity. The notion of complexity presented in the second piece is related to syncopation and it is explained in chapter 2.

The piece "Music Games" is composed for a set of 5 instruments. The note range of all instruments is idiomatic for the guitar but other instrumentation can be used.

Each instrument is assigned to a region of the virtual world (10x10) and the instruments play the rhythms from the agents which are located in the corresponding region of the world (5 adjacent columns of size 10x2). In this way, the music sources will give a rough idea of the agents location.

An average of 7 agents are present at the same time and each pitch represents one agent. The single pitch representation for each agent was chosen to permit the perceptual identification of the agents (Fig. 7.3).

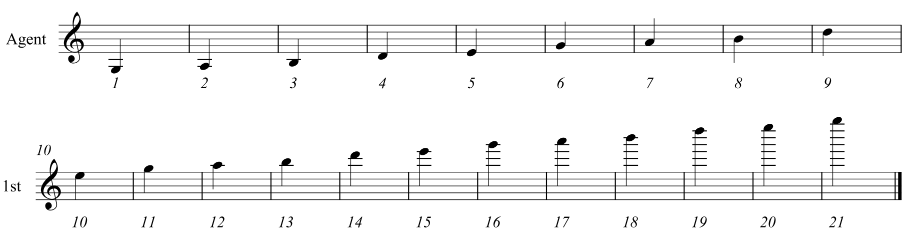


Figure 7.3: Pitch as an identifier of the agents

The generation and life-cycle of the sequences are contingent on the social interaction, and depend on values of complexity preference and received exposure of the RUs.

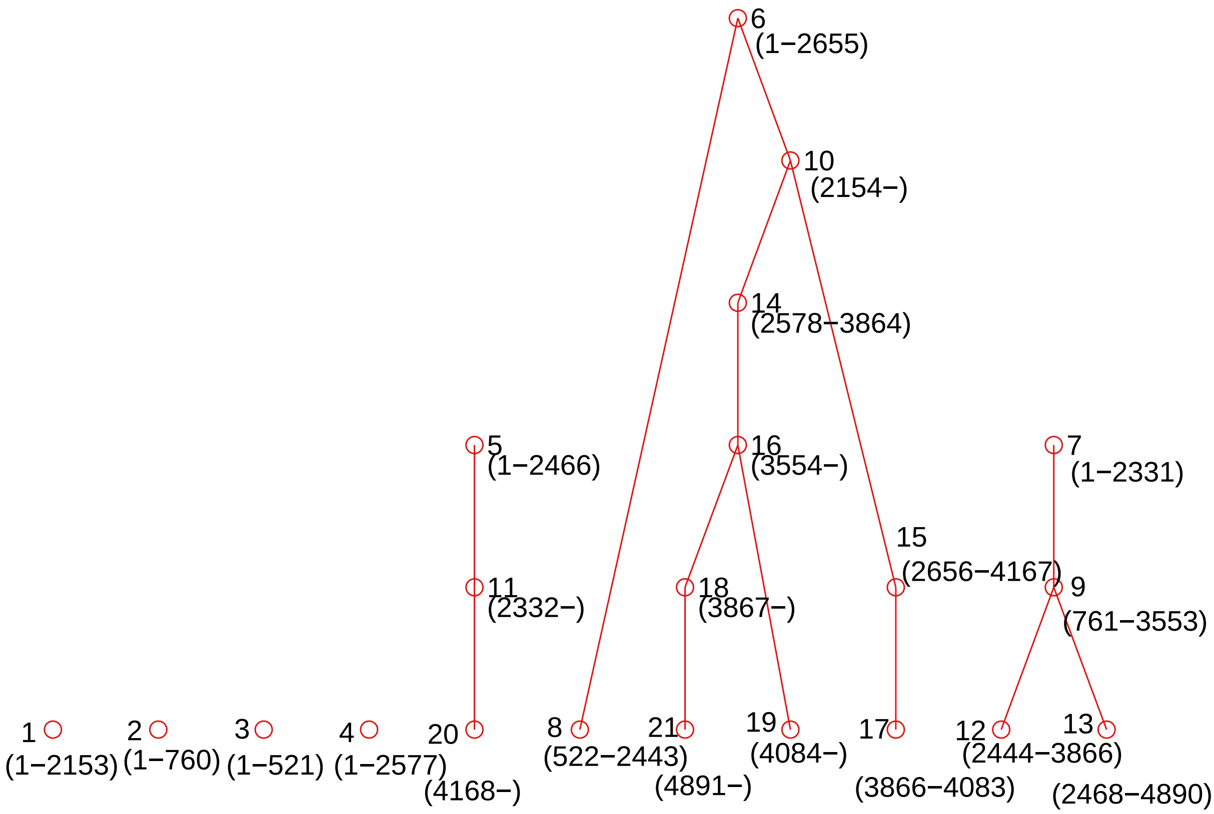


Figure 7.4: Tree indicating the life cycle of the agents used in the piece "Music Games".

The grammar determines the order of the sequences and decision process of the agents regarding what to play in an group. Each played sequence is generated in the basis of the grammar and finishes with a whole note to delimitate the sequence.

The rhythmic material for this piece was taken from a simulation of the system described in chapter 6 with the parameter values presented in Tab. 7.2.

The list of parameters used for the simulation used to generate material for the piece "Music Games" is presented in Tab 7.2. For a description of the parameters please see Sec. 6.3.

After running the simulation for 5000 iterations with population renewal, a tree with the life-cycle of the agents was produced (Fig. 7.4). From the analysis of this figure we can

Parameter	Description	Example value
$memLoss$	Exposure decay constant with time	0.01
a_{Tr}	Scale Factor for $changeRate$	0.2
a_{mat}	Adaptation weight for the Markov matrix	0.3
Ab_r	Ability increase rate	0.1
Ab_{init}	Initial agent's ability	1.10
Hc_{peak}	Peak value for complexity preference	3
C_{max}	Limit for complexity	5
X_r	Exposure increase when a RU is listened to	0.1
X_{max}	Limit for exposure	1
Hx_{peak}	Peak value for exposure preference	0.5
$initialRU$	Initial RU in agent's memory	[1 1 1 1]
$R_{interact}$	Radius of interaction	2
$PMove$	Probability of an agent moving	0.01
$PDie$	Initial Probability of dying	0
$PDieIncrease$	Increase in probability of dying	5.0000e-007
$MutExp$	Mutation for exposure	0.05
$MutComp$	Mutation for complexity	0.2
$initialAgents$	Initial number of agents	7
$maxIterations$	Number of iterations	5000
$worldSize$	Size of the virtual environment	[10 10]
$nBars$	Size of the composition in number of bars	100
$iterJump$	Iteration step where rhythms are probed	50
$nTracks$	Divisions of the virtual space played by instruments	5

Table 7.2: Parameters for a simulation with 7 agents in a 10x10 virtual world, with 5000 iterations giving origin to "Music Games".

see that there are 21 agents in total, spanning through a maximum of six generations. The number of bars in the composition is given by variable *nBars* and this defines the points in simulation time where the rhythms will be probed. In the case of the current piece rhythmic sequences would be probed in steps of 50 iterations ($iterJump = 50$).

Fig. 7.5 presents a detail from the piece beginning at the bar 50. In this screenshot it is possible to observe how the agent corresponding to the pitch E4 moves from the first region of the space to the second region of the space. It can also be seen that the agent with the pitch A4, in the 5th region of the space, has produced a rhythmic structure with

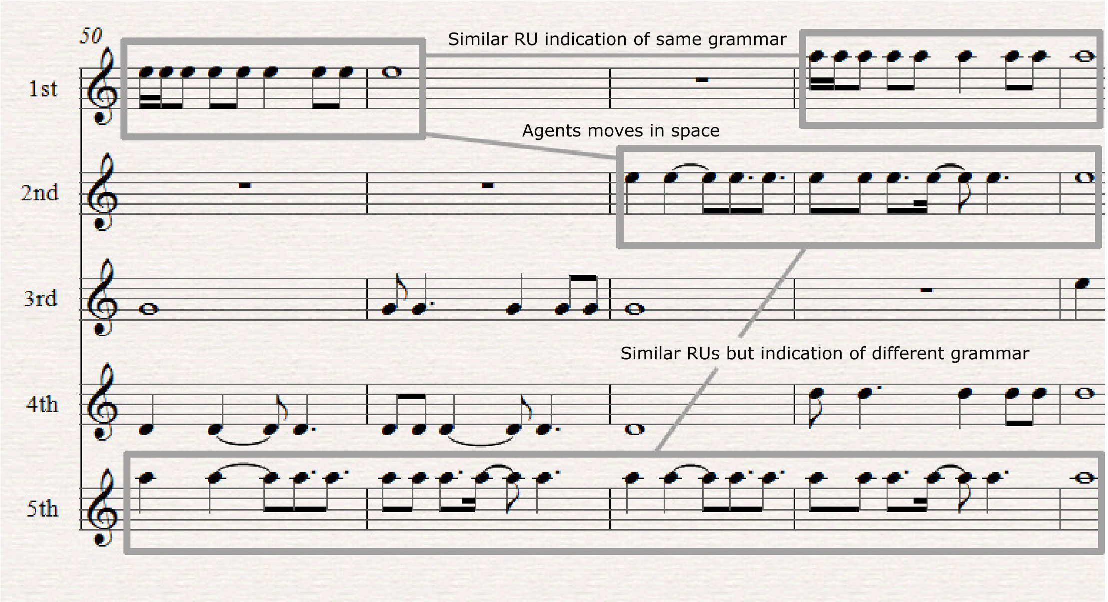


Figure 7.5: Bars 50-53 of "Music Games" with highlights of rhythmic structures

four rhythmic units on bar 50 and agent E4 uses the same rhythmic units but indicating to have a slightly different grammar. The fact that both use the same RUs and that the grammar has some similarities, indicates that they have interacted in the past.

In bar 53, the agent with the pitch A4 moves into the first region of the space, playing an exact reproduction of the structure by E4 in bar 50, therefore confirming that both agents have interacted recently.

Music Games

João Martins (2011)

7.6 Summary

As a composer I don't feel compelled to obey the rules of any particular music style. I try to create musical pieces that are new and interesting to me and hopefully someone will also find them interesting. Nevertheless, it is impossible to escape the context where one lives, and I try to be as aware as possible of the processes of cultural transmission and their effects on individuals. Much more than to be able to conform to a musical current, it is important for me to understand the processes that are hidden behind the emergence of music style, to understand how expectations are built and how these factors contribute to surprise. In A-Rhythm, the compositional goal is to find musical games and new interaction forms for musicians and non musicians to express themselves creatively.

The principle of self-organisation, or the theoretical constructions and discoveries associated to it, present a new way of looking into all domains of music. These domains can comprise: composition, in the way people design new music; performance, with new forms of interactions on stage and via new means of communication; learning, using self-organisation to show complex concepts and motivate music students with new exercises.

Most of the composers, even those who do not use computers, have used algorithms in their composition process and the usage of generative processes does not mean that the resulting music is devoid from the "hand of the composer". A carefully designed evolutionary algorithm can take longer to produce than to write a piece note by note. Also, an algorithm often has free parameters which can be changed in order to extract new behaviours from it. Finally, an evolutionary produces a material for a piece. The composer is always free to make *a posteriori* adjustments and select the parts which meet his aesthetical goals.

In this chapter, it was shown how A-Rhythm can be used in composition. Two pieces of music were composed, one for each implementation of the system, namely “Flickering Pleiades” and “Music Games”. In the beginning of the chapter, the musical relevance of this approach to composition was discussed, framing the system within the paradigm of evolutionary computation. Several examples of self-organisation in composition processes were presented.

In addition, the issue of mapping between the results provided by the simulations and the musical score was considered. The results can take the form of direct and indirect variables. The direct variables concern space, simulation time and the agents life cycle, whereas the indirect variables concern the cultural substrate of the agents and their musical behaviour. On the side of the musical score, the variables try to cover the most common compositional effects present in western classical music.

“Flickering Pleiades” is a piece for a set of 7 percussion instruments, containing rhythmic units from particular points of interaction during a typical simulation taken from the first implementation of A-Rhythm, and run with the complexity algorithm. This piece shows how this algorithm can evolve complex rhythms and what is the role of the players in clusters of agents. There is a strong geographical component, which is brought into evidence by the imitative nature of the game. The used rhythms are product of self-organisation in the sense that they are created the evolutionary algorithm.

“Music Games” is a piece for 5 guitars, containing rhythmic structures from particular points of the interaction during a typical simulation taken from the second implementation of A-Rhythm. The virtual space in the simulations is mapped into the 5 instruments and to each agent corresponds a pitch in the score. Each instrument only plays one rhythmic structure at a time which is delimited by a whole note.

The mapping options considered in “Flickering Pleiades” piece differ substantially from the “Music Games”. Whereas in the former piece each instrument corresponded to a particular agent, in the latter each instrument will play the rhythmic structures from agents placed in a region of the space. This brought consequences to the used instrumentation, hence “Flickering Pleiades”contains only percussion instruments, which are natural rhythmical instruments, and “Music Games” has a set of five identical melodic and harmonic instruments.

The cultural transmission process becomes evident by the appearance of common rhythmic units in different playing agents. In the case of “Music Games” we can also observe similar interconnection patterns between the rhythmic units. This is an example of the self-organised process. The development of the scores, with regard to the player agents, follow the time steps of the simulations and the performance of the pieces can be considered a sonification of the evolutionary process.

Finally, I would like to add that this system opens new possibilities for the performers and listeners to interact with the compositional process. Using the simulations as departing point, musicians can choose to play the structures produced by particular agents, or load the rhythmic structures that are more in tune with their aesthetical tastes. The internet will enable distant people to experiment with this framework and obtain results that are dependent on their culture.

Chapter 8

Discussion and Conclusions

The central focus of this thesis is the study of rhythm as an emergent property in the context of evolutionary computation.

In the introduction, the goals were established as to study rhythm, not only as a product of cognition, but also as a product of cultural evolution by means of computer modelling and simulations. In particular, how rhythm is transmitted between individuals and what motivates the emergence of particular rhythmic structures. It also looks into how rhythmic structures remain active in society for long periods and how music changes in the course of space and time.

To accomplish these objectives, a multi-agent systems was developed, with a first implementation focusing on algorithms of interaction that motivate rhythmic transformation and a second implementation focusing on the evolution of rhythmic complexity in syntactically constructed rhythmic structure. These structures were created by a grammar that was part of the agents cognitive apparatus. The repertoires of the rhythms and the grammar are shaped by self-organisation processes due to the interaction between agents in the virtual environment.

The fundamental reasons why the multi-agent system framework is chosen are four: there is no centralised control over the process of cultural evolution; the system focus on the rules of interaction between the individuals; individuals can show different

Criteria	A-Rhythm
Composition	x
Musicology	x
Rhythm	x
Complexity	x
Exposure	x
Spatial	x
Grammar	x
Horizontal Transmission	x
Vertical Transmission	x
Genetic	x

Table 8.1: Criteria satisfied by A-Rhythm.

characteristics; it resembles an actual society.

In order to design a system that studies the evolution of rhythm, one has to have a set of necessary criteria. These criteria are useful in the sense that they may help in the design process of the system and enable the comparison of related systems in the literature. The meaning of the criteria is explained in chapter 4 and the criteria are concatenated on a table in Sec. 4.4 which is reproduced here (Tab. 8.1).

Tab. 8.1) shows that all criteria were taken into account in the design of the system, meaning that A-Rhythm satisfies the necessary conditions for a multi-agent systems that models the evolution of rhythms, thus providing an answer to question number 1 (Q1) posed in the introduction.

8.1 Analysis

In the next two sections, a summary of the analysis of the two versions of A-Rhythm is presented:

8.1.1 Interaction games

The first implementation of the system focused on how different algorithms had an influence in the evolution of a rhythmic culture, namely the emergence of repertoires of rhythmic units that were exchanged between agents. Three different types of agent behaviours are explored according to the manipulation of the exchanged material during an interaction: popularity evaluation, transformation algorithm, and complexity based choice. The models were developed independently to observe the contribution of different behaviours to the process of repertoire growth and stability.

Complexity and repertoire tend to constantly grow unless the life-time of the agents is limited, originating population renewal. This is an alternative to have a memory loss procedure. The results of the simulations show the emergence of a coherent repertoire across society and clusters of agents can be observed. Fundamental differences were found in the evolution of complexity between the different algorithms and on the quantity of rhythms that emerged in each agents' repertoire. Using different conditions of the same algorithm, it was observed that a small subset of agents concentrates the preference of most of the population. In the third algorithm, it was observed that there was some tendency to generate a big cluster of agents in terms of complexity and average number of rhythms, but smaller groups of agents were found in which the complexity evolves with an inverse tendency of the repertoire growth. One of the obstacle that was found during the process of developing this system was the difficulty in defining measurements for similarity and complexity of rhythmic sequences. The analysis of the system was conducted using a measurement of similarity that can compare sequences of different lengths and also account for similarity between similar sub-sequences. The rhythmic complexity measure used was based on the number of events and their distribution inside a rhythm. These formal measures were developed based on the ideas taken from the literature and hence are inspired, but do not necessarily correspond to very accurate models of human perception.

8.1.2 Perception and Grammar

The second implementation of A-Rhythms focused on the use of perceptually relevant measures of complexity and on the evolution of a grammar. The aim of this exercise was to study the evolution of rhythmic grammars in an A-Life type of environment populated by interactive autonomous agents, endowed with rhythmic perception/production capabilities.

The agents move in a virtual environment and interact in groups defined by their position in the environment. The game of interaction consists of one agent playing a rhythmic sequence to the rest of the group. The generated rhythmic sequence consists of rhythmic units stochastically combined via a grammar which is defined by a Markov process.

After processing the sequence, the listening agents, as well as the player agent, generate two feedback values which are functions of the social evaluation of the sequence for complexity and exposure. The resulting feedback values influence the internal representation of the grammar pattern and the creation of new rhythmic units. An individual ability parameter introduces constraints to the production mechanism of the agents, enabling some sequences to be in memory, although not necessarily being playable by that agent. This parameter varies with the practice of complex rhythmic units.

The first observation was that some sequences became very long due to the existence of absorbent states in the Markov matrix (values near to 1 in the diagonal), or loops that perpetually repeated. This problem was overcome by slightly increasing the values in the first column of the matrix every time a sequence was generated with more than 20 rhythmic units. This increased the probability of ending the sequence earlier.

Different conditions were analysed, namely, the size of the world, number of agents, frequency of movement and population renewal.

The number of RUs was kept within bounds in every observed condition and complexity did not always present a steady growth. It was observed that a higher rate of transformation did not necessarily result in a growth in complexity. The size of the world had a direct influence in the number of interactions between the agents. A larger world meant that the agents would be interacting less frequently, given a particular pattern of movement. The analysis of the common RUs and grammar values between 2 agents shows a correlation between similarity of repertoires and grammar on the interaction periods and a divergence in the periods where the agents did not interact.

Subsequently an analysis of the synchronic and diachronic diversity was performed in larger groups of agents to evaluate whether the rhythmic units of the agents and corresponding grammars would converge. In the constant movement situation, all the network structure disappeared in the synchronic analysis, revealing no contribution of the spatial distribution of the agents to the repertoire formation. In the case of the "no movement" condition the agents developed repertoires and transition patterns which were highly dependent on their spatial distribution. With a small probability of moving, a higher degree of stability of the repertoires were observed. Diachronic analysis shows that, for this system, the changes across time are more significant than the differences between every agent at a particular time point.

Finally, population renewal was considered, with the hedonic peak values being genetically transmitted. This led to the observation that there is a convergence of the hedonic peaks, as new agents are born from existing ones, leading to a more homogeneous population.

8.2 Pieces

Chapter 7 illustrated how A-Rhythm can be used to compose music with. For this purpose, two pieces were composed using the material provided by the two implementations of the system. In this chapter, the problems of mapping the variables of the simulations onto the musical domain were addressed. The first piece, "Flickering Pleiades", used one of the considered algorithms, namely the *complexity* algorithm. The time steps of the simulation were divided into the number of bars to create the structure of the piece, using the rhythms played by the player agents at that particular time step. For each bar, additional bars were created containing rhythmic units with the same metric structure, from the repertoires of each of the listening agents situated in the neighbouring region.

The second piece, "Music Games", involved a more extensive mapping of the simulation dimensions where the virtual space from the simulations is mapped into 5 musical instruments. To each agent corresponds a pitch in the score and the rhythmic structures are all taken from player agents.

8.3 Contributions of the thesis

This thesis develops a set of tools for the study of the evolution of rhythms in a multi-agent system, set in an artificial environment. This framework permits the study of synchronic and diachronic diversity of the rhythms. A link between the theories of rhythmic complexity in the literature and the studies on hedonic values by Berlyne was established, to test the notion of cultural evolution in a controlled artificial environment. During the course of the research, a new similarity measure was developed that enables the comparison of rhythms of different lengths and with structural relations on the sub-sequence level.

Music itself cannot be carbon-dated and compared to the age of notation, as the earliest examples of music writing are too recent. There is another way of reconstructing ancient music by studying the history of music instruments, but if these can provide information in relation to pitch and timbre, rhythm is totally out of grasp.

If we want to figure out how music could have been, we need to look into the human mind, which is thought to be structurally unchanged for many millennia, and see how we process and create rhythms and how new patterns emerge. Then, we also need to look into generative models and study how the rhythms are selected and passed from generation to generation. The system developed in this thesis is a small contribution in that direction.

Surely some information is irrevocably lost in history, but there is a lot to be gained in understanding the puzzles in both ancient and recent history by looking at the missing links.

8.3.1 Research answers

In this section I will address the questions proposed in the outset of this thesis individually:

- Q1. Is it possible to use multi-agent systems to model the evolution of rhythms? What would be the criteria for the design of such a system?

A-Rhythm was build with a set of criteria in mind. It was shown in chapters 5 and 6 that repertoires of rhythmic units, along with grammars, can evolve in an artificial environment guided by an evolutionary process. Moreover, it was shown that the individual repertoires do not evolve randomly, but will develop in close connection to the repertoires of the surrounding agents. In the first implementation of A-Rhythm, different behaviours were studied with regard to three different interaction games and the results were presented.

- Q2. How can rhythms be represented and the product of this representation be recombined in a creative manner?

The process by which rhythms are represented in A-Rhythm is by using rhythmic units which are subject to transformations. These units are represented using short sequences of inter-onset intervals (IOIs) of durations indexed to a pulse. This representation was chosen based on the notion, derived from the literature, that sequences of IOIs convey the highest amount of information regarding rhythm. Although other potentially relevant information such as amplitude of sound, envelope, or timbre, might be important in determining accents and rhythmic complexity, the characterisation of a rhythm does not dispense the information about IOIs, and this information will be sufficient to study the most important characteristics of a rhythm. This representation, together with the notion of metre, enabled the categorisation of the rhythms in terms of complexity and similarity.

In the second implementation of A-Rhythm, the previously studied rhythmic units were combined in structures using a grammar. This grammar takes the form of a Markov process in which the rhythmic units can be recombined in potentially infinite ways.

- Q3. What motivates change in music?

This question was dealt by carrying a literature review in the mechanisms of change, by experimenting with different behaviours and became a corner stone in defining the conditions for the simulations contained in A-Rhythm. Traditionally, culture change is seen by anthropologists as being a product of either some form of generation and selection, or acculturation, by importing elements from other cultures. In this thesis both processes are inherent in the system. On one side, individual agents generate new sequences by cumulative transformations to rhythmic units, which are then reinforced or neglected due to the internal processes of the agents. On the other hand, agents are subject to external influence, and themselves influence the surrounding environment. As a consequence, by considering movement, one can observe significant processes of acculturation. The agents in this system are very limited regarding the cognitive capabilities, and therefore we can only deal with a very limited amount of behaviours, leaving behind others, that definitely have an impact in musical change. Further study will enable a better understanding on how music change happens in society and some of these questions are addressed in section 8.4 of this chapter, in the recommendations for further work.

- Q4. How can such volatile things as unwritten pieces of music hold for several generations, or spread to large parts of the population, and what are the features in oral transmission that enable these processes to take place?

The answer to this question relies on memory reinforcement provided by exposure, and in the process of interaction. As it can be seen from most algorithms, the number of rhythmic units in an agent's memory will rise, from the single pulse, up to a constant value dictated by the rate of transformation, the amount of interactions and by memory loss. The rhythms are then transmitted to other agents which will subsequently become part of their repertoire. In both implementations of A-Rhythms it was observed that some rhythms were transmitted to offspring of the agents in a vertical transmission process. This iterated process of transmission meant that some of the rhythms persisted long after their creators disappeared.

- Q5. Is it possible to use the output of such a system in composition?

In chapter 7 it was shown that A-Rhythm can also be used as composition tool. Two compositions were created, "Pleiades" and "Music Games", that reflect the evolutionary process for both implementations of A-Rhythm. The system does not present finished compositions, but instead generates rhythmic material and data from the simulation, which need to be mapped into the musical space. One advantage of this system is the fact that the agents possess a great deal of diversity in its history, and yet have many commonalities with other agents. This means that each agent can be treated as a single musical entity generating rhythms, but can also create interesting effects in dialogs with other agents.

8.4 Recommendations for future work

Many fields of research are left open when dealing with experiments of forms of cultural transmission and definition of musical style.

In the course of the research, I realised that there could be other pathways and aspects that could have been investigated, but they do not address directly the posed questions and another sub-set of tasks needed prior attention.

With the current system it is possible to study the kinship relations between the agents and the influence these relations have in the evolution of the repertoires and general trends in the society. Studying the relations of parent to offspring, as well as relations between siblings enable a distinction between vertical, horizontal and oblique social transmission paths.

In this study, similarity is only used in the analysis process, whereas the actual algorithm of transmission of repertoire discards information on similarity, considering only perfect matches when computing exposure values. In other studies with imitation games, both in music (Miranda 2002*a*) and language (Boer 1999), similarity plays an important role in discarding new elements which are considered to be in the same category. Similarity can also be considered as defining a musical style.

Another possible line of research is to try the grammar model with different metrical structures, using other types of transformations. To accomplish this task, it will be necessary to previously quantify metrics for complexity, given the new metrical patterns, and perform verification with listening tests.

Another challenge is to bring these systems into a real-time framework and letting actual musicians interacting with it. The present study is directed to composers who are making decisions without time constraints, but in principle it could also be applied to an improvisation setup.

One of the most pressing goals for the future, although probably one of the most difficult solution, is the parametrization of the system to study the current changes in media technology. This was briefly mentioned in the introduction, but due to its complexity, remains a topic for further research. In order to study how the internet and how the current changes in music industry will affect our music listening, one can use such the system with higher radius of interaction, or by adding links between distant agents, and compare it to the results produced when the radius of interaction is smaller. It should be possible to observe if the structures evolved would become more complex or if more structures would be present in the individual repertoires.

I hope this research will be useful to musicologists and composers who struggle with understanding the complex mechanism by which music changes and becomes part of our culture.

Appendix A

Record of Activities

A.1 Published papers

- Martins, J. M. and Miranda, E. R. (2008). "Engineering the Role of Social Pressure: A New Artificial Life Approach to Software for Generative Music". *Journal on Software Engineering*, Vol. 2, No. 3, pp.31-42.

- Martins, J. M. and Miranda, E. R. (2008). "Breeding Rhythms with Artificial Life", *Proceedings of the 5th Sound and Music Computing Conference*, Berlin (Germany).

- Martins, J. M. and Miranda, E. R. (2007). "Emergent Rhythmic Phrases in an A-Life", *Proceedings of the Workshop on Music and Artificial Life - European Conferecence of Artificial Life (ECAL)*, Lisbon (Portugal).

- Martins, J. M. and Miranda, E. R. (2006). "Evolution of rhythm in a society of virtual agents", *Rhythm, time and temporal organisation conference*, Edinburgh (UK).

- Martins, J. M. and Miranda, E. R. (2006)."A Connectionist Architecture for the Evolution of Rhythms", *Proceedings of EvoWorkshops 2006*, Lecture Notes in Computer Science, Berlin: Springer-Verlag, Budapest (Hungary).

- Martins, J. M. and Miranda, E. R. (2005). "Societies of rhythmic agents: A neural representation of rhythm", *Proceedings of the 10th Rhythm Perception and Production Workshop (RPPW2005)*, Bilzen (Belgium).

- Martins, J. M., Gimenes, M. , Manzolli, J. and Maia Jr, A. (2005). "Similarity Measures for Rhythmic Sequences", *Proceedings of the 10th Brazilian Symposium of Musical Computation (SBCM)*, Belo Horizonte (Brazil).

- Coutinho, E., Gimenes, M., Martins, J. and Miranda, E. R. (2005). "Computational Musicology: An Artificial Life Approach", *Proceedings of the 2nd Portuguese Workshop on Artificial Life and Evolutionary Algorithms Workshop*, Covilhã (Portugal).

A.2 Research lectures without proceedings

- "New developments in the study of the evolution of rhythm", Talk given at the Language Evolution and Computation Research Unit, University of Edinburgh, April 2007.

- "Evolution of Rhythms in Artificial Worlds", Talk given at the Live Algorithms for Music research network (LAM), Goldsmiths College, London, December 2005

- "Evolutionary Rhythms", Talk given at the University of Plymouth, November 2005

- "Societies of rhythmic agents: The synchronization problem and neural representation of rhythm" Talk given at the Max Planck Institute for Human Development, Berlin, June 2005

- "Towards a computational Model of the Evolution of Musical Grammars", Talk given at the University of Plymouth, May 2005.

A.3 Other

- “DECOI 2006 - Summer School on the Design Of Collective Intelligence”, Vrije Universiteit Amsterdam, August 7-11 2006

Appendix B

Pieces

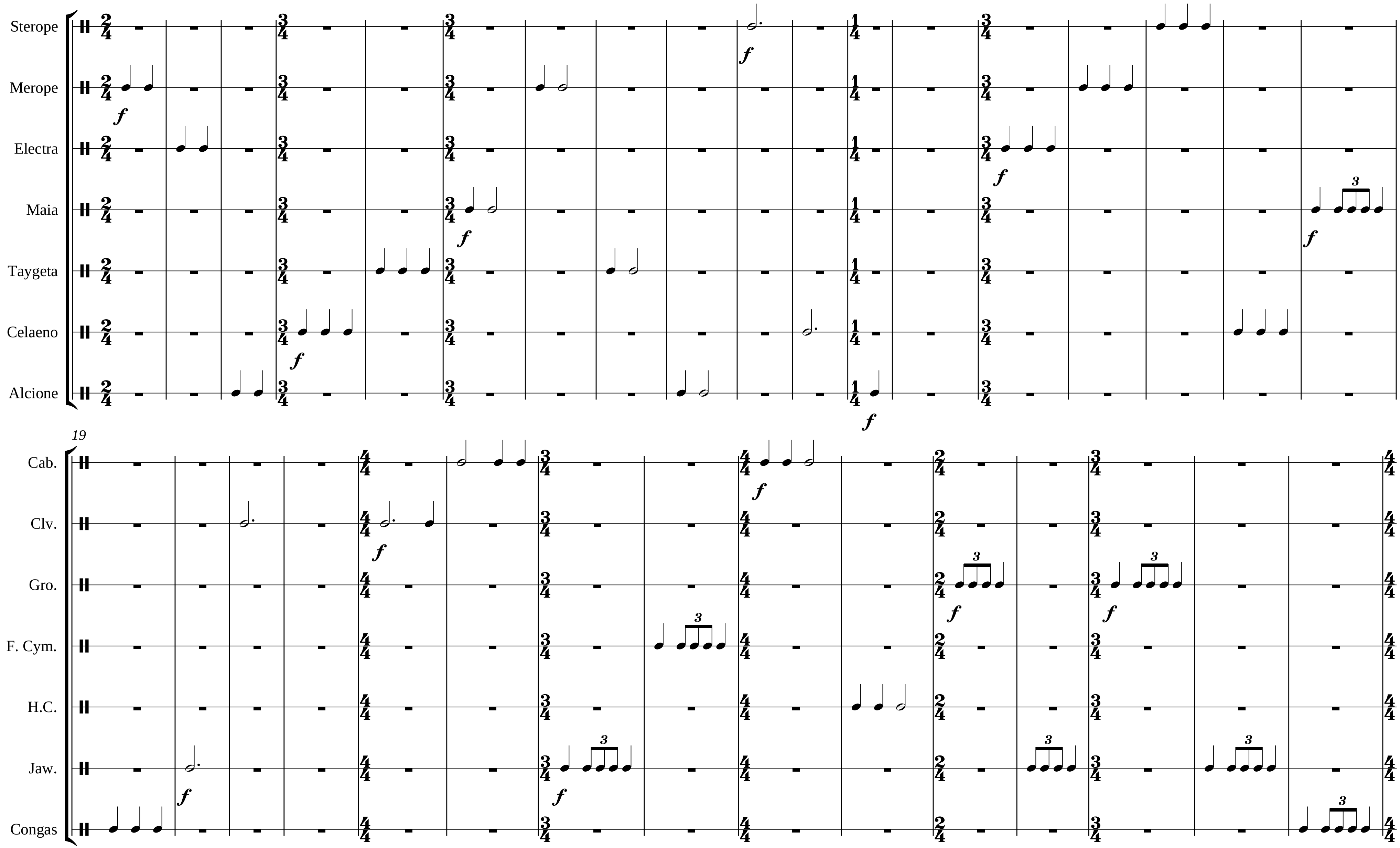
In this section of the appendix the completes pieces “Flickering Pleiades” and “Grammar Games”. The first piece is for a set of 7 percussion instruments and its composition process is described in chapter 7 with rhythmic material extracted from the simulations described in chapter 5. This piece focus on the development of complexity in the repertoires.

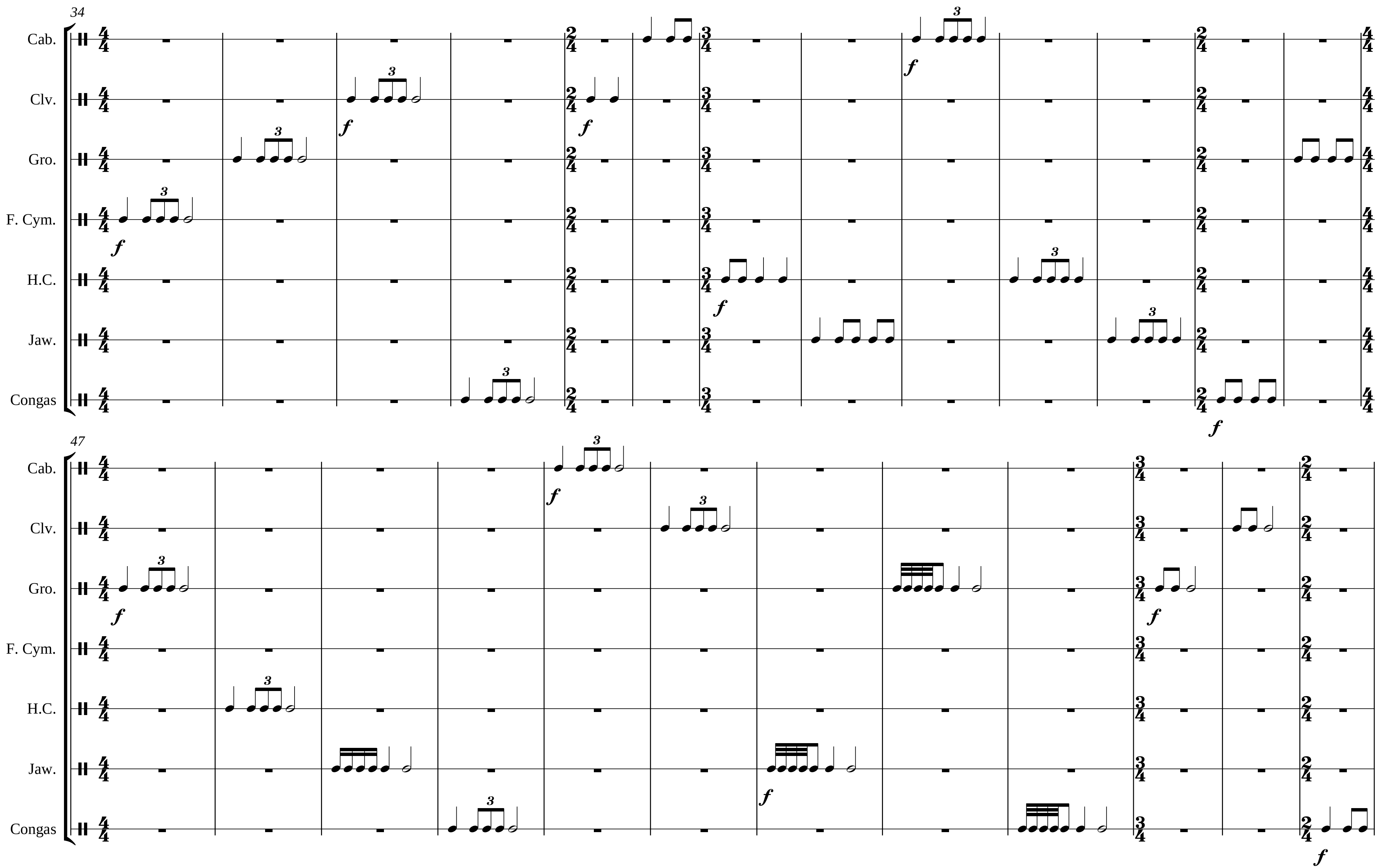
The second piece shows how agents can incorporate a grammar based on the principles of imitation and transformation investigated in the system described in chapter 6. Each instrument is assigned to a region of the space and each pitch corresponds to one agent.

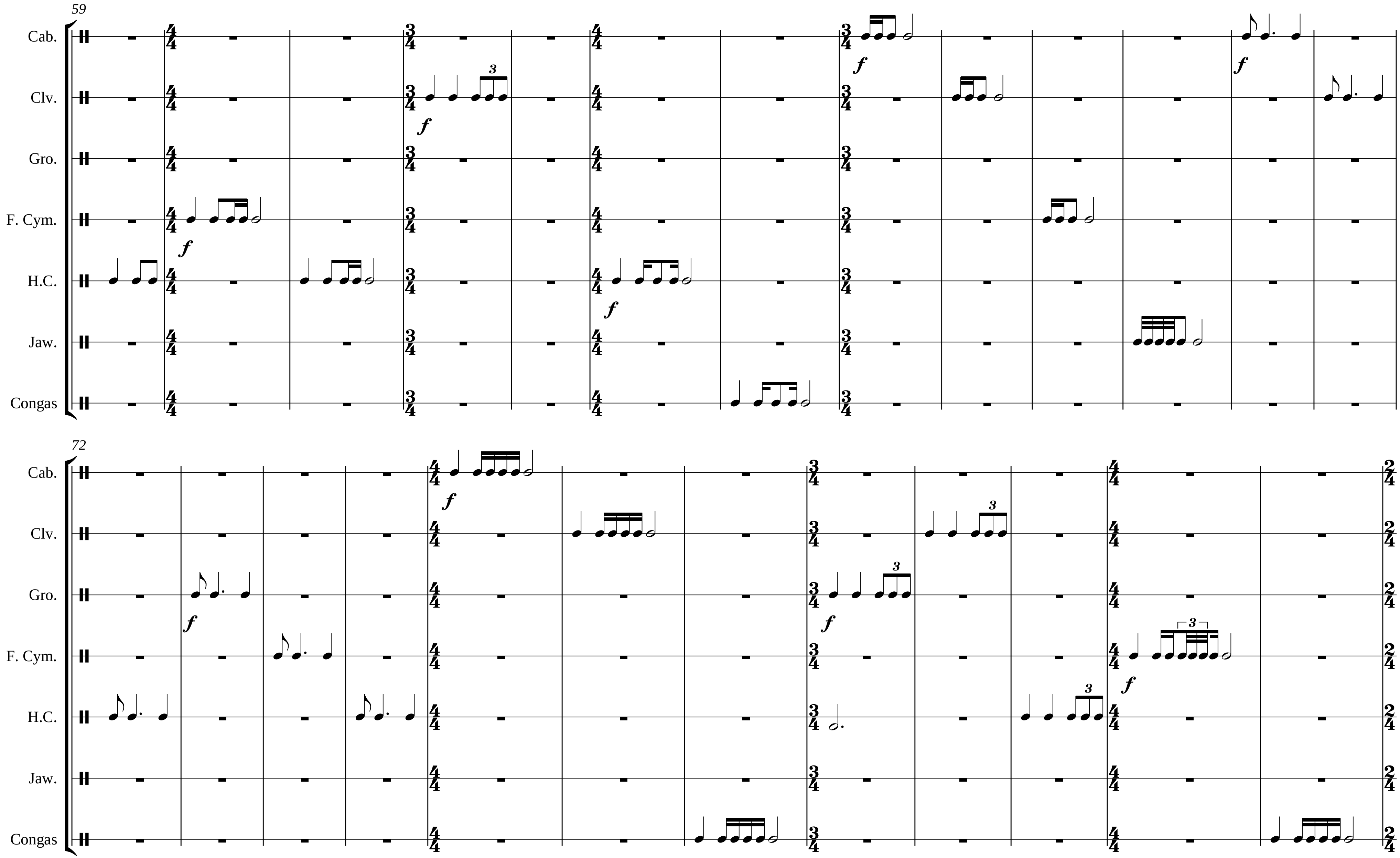
The two pieces are published in a CD that is enclosed with this thesis.

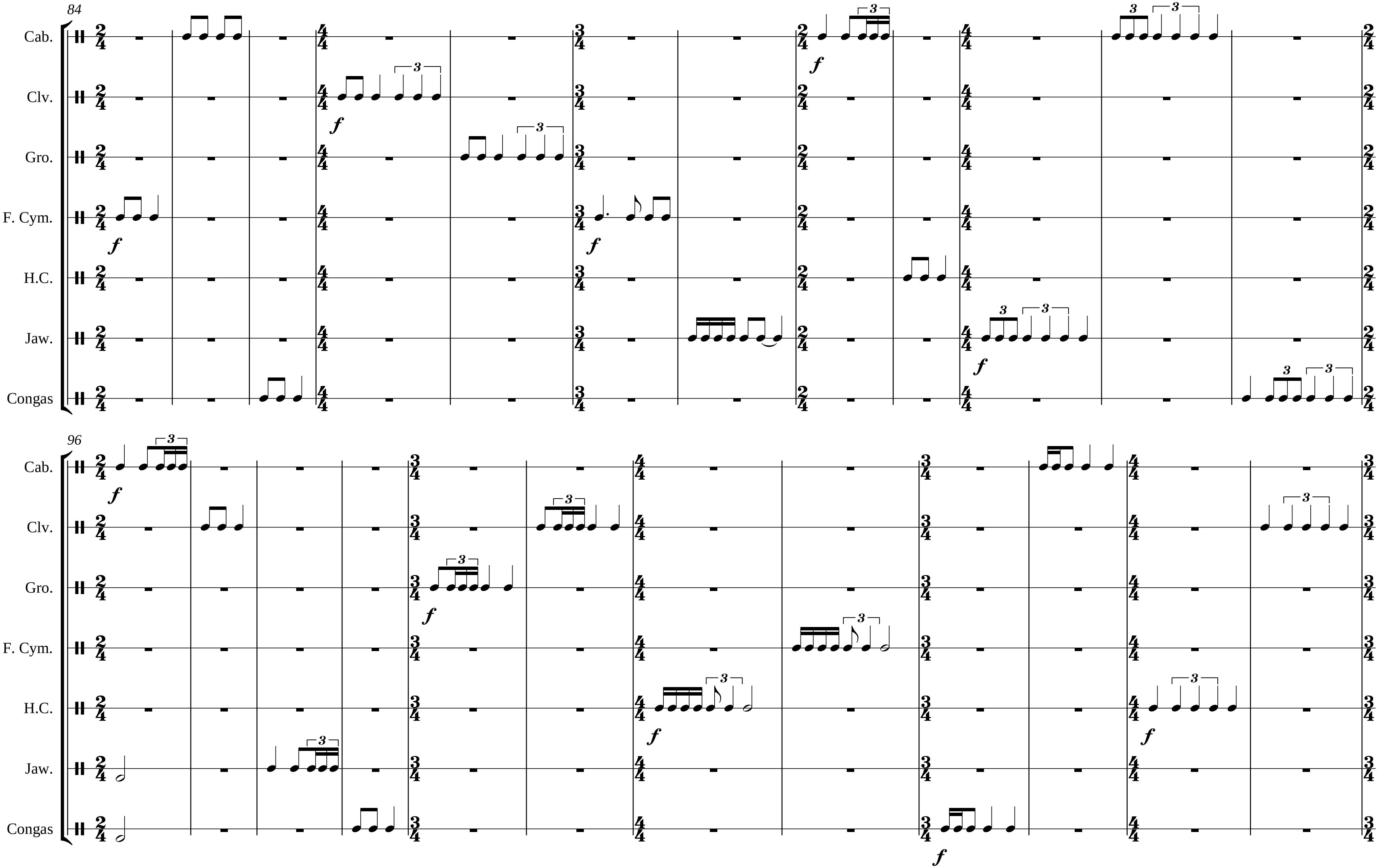
Flickering Pleiades

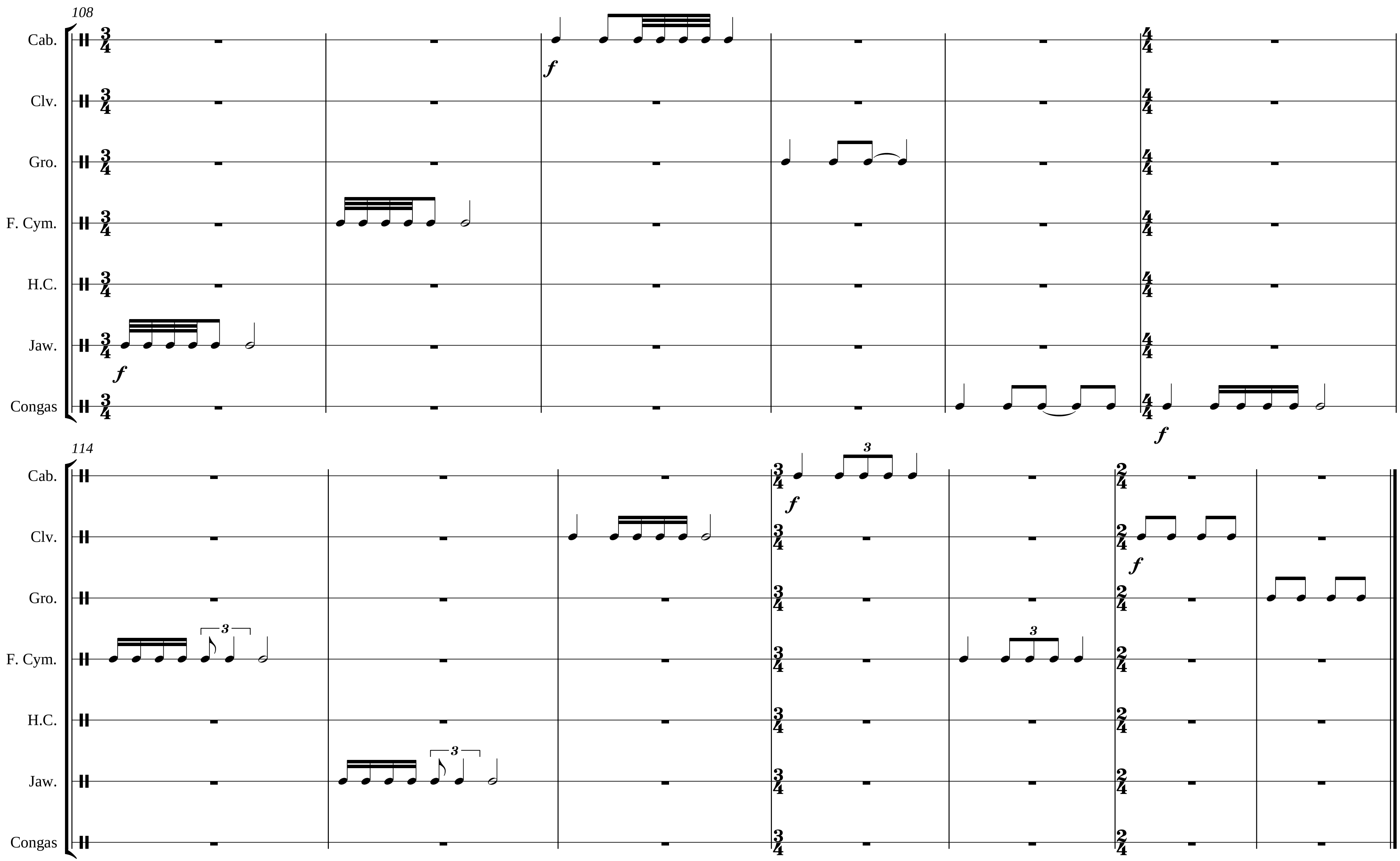
Joao Martins (2011)




34
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas
47
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas


59
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas
72
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas


84
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas
96
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas



108
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas
114
Cab.
Clv.
Gro.
F. Cym.
H.C.
Jaw.
Congas

Music Games

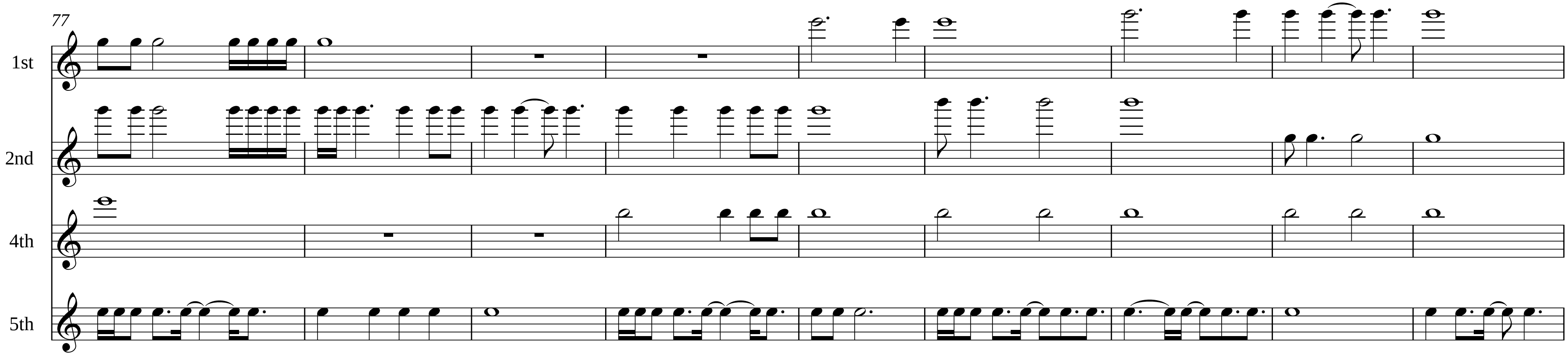
João Martins (2011)


21
1st
2nd
3rd
4th
5th

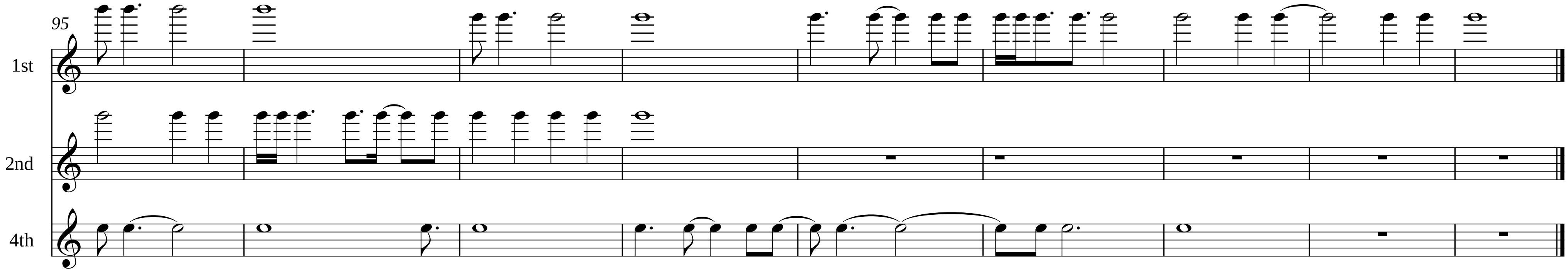

31
1st
2nd
3rd
4th
5th


41
1st
2nd
3rd
4th
5th
50
1st
2nd
3rd
4th
5th


59
1st
2nd
3rd
4th
5th
68
1st
2nd
3rd
4th
5th


77
1st
2nd
4th
5th


86
1st
2nd
3rd
4th
5th


95
1st
2nd
4th

Glossary

Anacrusis - Often referred to as "Pick up notes" are starting notes before the main beat of a metric hierarchy (bar).

A-Life - Artificial Life is a field of computer science developing "life like" algorithms, inspired by biology, psychology and the social sciences, creating models that draw some insight into complex processes or help solving computer science problems.

Accent - "An event that is marked for consciousness" (Cooper and Meyer 1963 p.8).

Bar - Element of music notation that defines the boundaries of a metric hierarchy.

Beat - Structural element of a metric hierarchy.

Downbeat - Points in time where the metrical accents have higher weights. Typically the first beat in a measure.

Emergence - Complex patterns or behaviours displayed by systems of interacting parts subject to self-organisation.

Exposure - Loosely defined as the amount of times a particular musical material is used in a given context.

Finite acceptor - A kind of finite automaton with a univocal mapping of the set of states to the set {"acceptable","unacceptable"} (Bel and Kippen 1992 p.398).

Legato - Playing style where the musician intends to blend. the melodic line as much as possible.

Mapping - The establishment of a sensible connection between two areas of activity which creates meaning to a human or machine perceiver citeBeyls2004.

Meme - A basic unit of cultural transmission in the same way that a genes, in biology, are units of genetic information. This term was invented by Richard Dawkins in the Selfish Gene (Dawkins 1989).

Metre - “A perceptually emergent property of a musical sound, [...] a form of entrainment [...] , a synchronisation of some aspect of our biological activity with regularly recurring events in the environment” (London 2004 p.4).

Metric grid - Grid of discrete time points that cyclically represent the idealised points of metrical accent in the musical flow or music notation.

Metrical preference rules - According to the GTTM (Lerdahl and Jackendoff 1983), these are metrical rules used to solve ambiguities created by parsing of the groups in the music surface.

Metrical Well-formness rules - According to the GTTM (Lerdahl and Jackendoff 1983), these are a set of rules that explain the temporal regularities captured by notation, as well as those implicit in musical notation and musical practice.

Music surface - Is the actual musical line described by notation regardless of the metric grid perceptually superimposed.

Rhythm - “... a movement marked by the regulated succession of strong or weak elements” (London 2001).

Self-organisation - Property of systems of interacting parts, which show complex patterns or behaviours as a result of relatively simple modes of interaction.

Sonification - The use of non-speech audio to convey information or perceptualize data (Kramer 1993).

Staccato - Playing style where the note is damped immediately after it is played, opposed to legato.

Stigmergy - A biological process that involves animals communicating with others over long time scales, by modifying the environment.

Syncopation - Rhythms with beats deviating from the metrical structure of a piece. If a metrical position without onset has a higher metrical level than the next sounding note, then that pair of notes is syncopated(Longuet-Higgins and Lee 1984).

Rest - Notation figure or period of time corresponding to the absence of sound.

Tactus - Isochronous pulse used as a reference for the relative durations contained in a rhythm.

Tempo - Value in beats per minute (bpm) that defines the rate of events, or speed, of the tactus.

Upbeat - Points in time where the metrical accents have lower weights.

List of references.

Allers, R. and Minkoff, R. (1994), '(Dirs.) The Lion King'.

Ames, C. and Domino, M. (1992), Cybernetic composer: an overview, *in* M.Balaban, K. Ebcioglu and O.Laske, eds, 'Understanding Music with AI:Perspectives on music Cognition', The Mit Press, pp. 186–205.

Anders, T. (2009*a*), A computational model that generalises schoenberg's guidelines for favourable chord progressions, *in* '6th Sound and Music Computing Conference'.

Anders, T. (2009*b*), A model of musical motifs, *in* T. Klouche and T. Noll, eds, 'Mathematics and Computation in Music', Springer-Verlag.

Anders, T. and Miranda, E. R. (2011), 'A survey of constraint programming systems for modelling music theories and composition', *ACM Computing Surveys* **43**(4).

Arom, S. (2004), *African Polyphony and Polyrhythm: Musical Structure and Methodology*, Cambridge University Press.

Barry, N. H. and Hallam, S. (2002), Practice, *in* R. Parncutt and G. E. McPherson, eds, 'The Science and Psychology of Music Performance', OUP.

Bartok, B. (1938), 'The so-called bulgarian rhythm', *Enekszo* **6**, 537–541.

Basharin, G. P., Langville, A. N. and Naumov, V. A. (2003), 'The life and work of a. a. markov', online. http://www.maths.tcd.ie/ nhb/talks/Markov.pdf (last access in Apr 2011).

Beek, P. J., Peper, C. E. and Daffertshofer, A. . (2000), Timekeepers versus nonlinear oscillators: how the approaches differ, *in* P. Desain and L. Windsor, eds, 'Rhythm Perception and Production', Swets and Zeitlinger, pp. 9–34.

Bel, B. (2008), The bol processor project: musicological and technical issues, *in* 'Virtual Gamelan Graz: Rules - Grammars - Modeling', Shaker Verlag GmbH Germany.

Bel, B. and Kippen, J. (1992), Bol processor grammars, *in* M.Balaban, K. Ebcioglu and O.Laske, eds, 'Understanding Music with AI:Perspectives on music Cognition', The Mit Press, pp. 64–109.

Berlyne, D. E. (1971), *Aesthetics and Psychobiology*, Appleton Century Crofts.

Beyls, P. (1989), The musical universe of cellular automata, *in* 'Proceedings of the 1989 International Computer Music Conference', pp. 34–41.

Beyls, P. (2004), Cellular automata mapping procedures, *in* 'Proceedings of the International Computer Music Conference (ICMC2004)'.

Biles, J. A. (1994), Genjam: A genetic algorithm for generating jazz solos, *in* P. Anderson and K. Warwick, eds, 'Proceedings of the International Computer Music Conference', International Computer Music Association, Aarhus(Denmark), pp. 131–137.

Biles, J. A. (2007), Improvizing with genetic algorithms: Genjam, *in* E. Miranda and A. Biles, eds, 'Evolutionary Computer Music', Springer Verlag.

Bilotta, E., Miranda, E., Pantano, P. and Todd, P. M. (2002), 'Artificial life models for musical applications: Workshop report', *Artificial Life* **8**(1), 83–86.

Blackwell, T. (2008), Swarm granulation, *in* J. Romero and P. Machado, eds, 'The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music', Springer Verlag, chapter 5, pp. 103–122.

Blackwell, T. and Jefferies, J. (2005), Swarm techtiles, *in* R. R. et al., ed., 'EvoWorkshops 2005', LNCS 3449, Springer-Verlag, Berlin, p. 468Ű477.

Blackwell, T. M. (2007), Swarming and music, *in* E. Miranda and A. Biles, eds, 'Evolutionary Computer Music', Springer Verlag.

Blackwell, T. M. and Young, M. W. (2004), Swarm granulator, *in* G. R. e. a. Raidl, ed., 'EvoWorkshops 2004', LNCS 3005, Springer-Verlag, Berlin, p. 399Ű408.

Blackwell, T. and Young, M. (2005), 'Live algorithms', *Society for the Study of Artificial Intelligence and Simulation of Behaviour Quarterly* **122**, 7.

Boer, B. d. (1999), Self-Organisation in Vowel Systems, PhD thesis, Vrije Universiteit Brussel AI-lab.

Bolton, T. L. (1894), 'Rhythm', *American Journal of Psychology* **6**, 145–238.

Bown, O. (2008), Theoretical and Computational Models of Cohesion, Competition and Maladaptation in the Evolution of Human Musical Behaviour, PhD thesis, University of London (Goldsmiths College), UK.

Boyd, R. and Richerson, P. J. (1988), *Culture and the Evolutionary Process*, University Of Chicago Press.

Brown, A. R. (2005), Exploring rhythmic automata, *in* F. e. a. Rothlauf, ed., 'Proceedings of the 3rd European Workshop on Evolutionary Music and Art', Springer Verlag, Lausanne(Swizerland), pp. 551–556.

Burton, A. R. and Vladimirova, T. R. (1999), 'Generation of musical sequences with genetic techniques', *Comput. Music J.* **23**(4), 59–73.

Burton, T. (1992), '(Dir.) Batman Returns'.

Camazine, S., Franks, N. R., Sneyd, J., Bonabeau, E., Deneubourg, J.-L. and Theraula, G. (2001), *Self-Organization in Biological Systems*, Princeton University Press, Princeton, NJ, USA.

Cangelosi, A. and Parisi, D., eds (2002), *Simulating the Evolution of Language*, Springer Verlag, London.

Cangelosi, A., Smith, A. and Smith, K., eds (2006), *The Evolution of Language*, Singapore: World Scientific.

Cayne, B. S., ed. (1990), *The new Lexicon dictionary of the English language*, Lexicon, New York.

Chernoff, J. M. (1979), *African Rhythm and African Sensibility*, The University of Chicago Press.

Chomsky, N. (1956), 'Three models for the description of language', *IRE Transactions on Information Theory* **2**, 113Ű124.

Chomsky, N. (1957), *Syntactic Structures*, 2nd revised edition (2002) edn, Mouton de Gruyter.

Chomsky, N. (1959), 'On certain formal properties of grammars', *Information and Control* **2**(2), 137–167.

Christiansen, M. and Kirby, S., eds (2003), *Language Evolution.*, Oxford University Press.

Conklin, D. and Witten, I. H. (1995), 'Multiple viewpoint systems for music prediction', *Journal of New Music Research* **24**, 51–73.

Cooper, G. and Meyer, L. (1963), *The Rhythmic Structure of Music*, University Of Chicago Press.

Cope, D. (1991), *Computers and musical style*, A-R Editions, Inc., Madison, WI, USA.

Cope, D. (1992*a*), A computer model of music composition, *in* S. Schwanauer and D. Levitt, eds, 'Machine Models of Music', MIT Press, Cambridge, MA, USA, pp. 263–286.

Cope, D. (1992*b*), 'Computer modelling of musical intelligence in EMI', *Computer Music Journal* **16**(2), 69–83.

Cope, D. (1992*c*), On algorithmic representation of musical style, *in* M.Balaban, K. Ebcioglu and O.Laske, eds, 'Understanding Music with AI:Perspectives on music Cognition', The Mit Press, pp. 355 – 363.

Cope, D. (1996), *Experiments in Musical Intelligence*, Vol. 12 of *Computer Music and Digital Audio Series*, A-R Editions; Bk&CD-Rom edition.

Cope, D. (1997), 'The Composer's Underscoring Environment: CUE', *Computer Music Journal* **21**(3), 20–37.

Cross, I. (2003), 'Music and evolution: Consequences and causes', *Contemporary Music Review* **22**(3), 79–89.

Cross, I. (2010), The evolutionary basis of meaning in music: some neurological and neuroscientific implications, *in* F. C. Rose, ed., 'Proceedings of the Medical Society of London's Mansell Bequest Symposium - The neurology of music', Imperial College Press, London.

Dahlstedt, P. (2004), Sounds Unheard of: evolutionary algorithms as creative tools for the contemporary composer, PhD thesis, Chalmers University of Technology, Götteborg, Sweden.

Dali, S. (1970), *Dali by Dali*, Abrams.

Darwin, C. (1871), *The descent of man and selection in relation to sex*, John Murray.

Dawkins, R. (1989), *The The Selfish Gene*, Oxford University Press, London.

Dawkins, R. (1996), *The Blind Watchmaker*, W. W. Norton & Company.

de Campo, A. and Dayé, C. (2006), Navegar é preciso, viver nao é preciso, *in* 'Proceedings of the 12th International Conference on Auditory Display', London (UK).

Demaine, E. D., Gomez-Martin, F., Meijer, H., Rappaport, D., Taslakian, P., Toussaint, G. T., Winograd, T. and Wood, D. R. (2009), 'The distance geometry of music', *Comput. Geom. Theory Appl.* **42**(5), 429–454.

Díaz-Bañez, M., Farigu, G., Gómez, F., Rappaport, D. and Toussaint, G. T. (2004), El compás flamenco: a phylogenetic analysis, *in* 'In Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science', Southwestern College, Kansas,.

Dissanayake, E. (1992), *Homo Aestheticus: Where Art Comes From and Why*, University of Washington Press.

Drake, C. (1998), 'Psychological processes involved in the temporal organization of complex auditory sequences: Universal and acquired processes.', *Music Perception* **16**, 11–26.

Drake, C. and Bertrand, D. (2001), 'The Quest for Universals in Temporal Processing in Music', *Ann NY Acad Sci* **930**(1), 17–27.

Drake, C. and Gérard, C. (1989), 'A psychological pulse train: how young children use this cognitive framework to structure simple rhythms.', *Psychol Res.* **51**(1), 16–22.

Ebcioglu, K. (1992), An expert system for harmonizing four-part chorales, *in* S. Schwanauer and D. Levitt, eds, 'Machine Models of Music', MIT Press, Cambridge, MA, USA, pp. 385–401.

Eigenfeldt, A. (2009), The evolution of evolutionary software: Intelligent rhythm generation in kinetic engine, *in* 'Proceedings of EvoMusArt 09, the European Conference on Evolutionary Computing'.

Essens, P. (1995), 'Structuring temporal sequences: Compason of models and factors of complexity.', *Perception and Psychophysics* **57**(4), 519–532.

Fitch, W. T. (2005), 'The evolution of music in comparative perspective', *Annals of the New York Academy of Sciences* **1060**, 1749–6632.

Fitch, W. T. (2006), 'On the biology and evolution of music', *Music Perception* **24**, 85–88.

Fitch, W. T. and Rosenfeld, A. J. (2007), 'Perception and production of syncopated rhythms', *Music Perception* **25**, 43–58.

Fraisse, P. (1956), *Les structures rythmiques*, Studia psychologica (Louvain, Belgium).

Gabrielsson, A. (1973), 'Similarity ratings and dimension analyses of auditory rhythm patterns i', *Scand. J. Psychol.* **14**, 138–160.

Gartland-Jones, A. and Copley, P. (2003), 'The suitability of genetic algorithms for musical composition', *Contemporary Music Review* **22**(3), 43–55.

Gimenes, M. (2009), An Approach to Machine Development of Musical Ontogeny, PhD thesis, School of Computing, Communications and Electronics, University of Plymouth.

Gimenes, M., Miranda, E. and Johnson, C. (2007), Emergent musical environments: An artificial life approach, *in* 'Proceedings of the Workshop on Music and Artificial Life (ECAL2007)', Lisbon (Portugal).

Gimenes, M., Miranda, E. R. and Johnson, C. (2005), A memetic approach to the evolution of rhythms in a society of software agents, *in* 'Proceedings of the 10th Brazilian Symposium of Musical Computation (SBCM)', Belo Horizonte (Brazil).

Gjerdingen, R. O. (1989), 'Meter as a mode of attending: A network simulation of attentional rhythmicity in music', *Intégral* **3**, 67–91.

Gong, T., Zhang, Q. and Wu, H. (2005), Music evolution in a complex system of interacting agents, *in* 'Proceedings of the 2005 IEEE Congress on Evolutionary Computation', Vol. 2, Edinburgh (UK), pp. 1111 – 1118.

Grube, M. and Griffiths, T. D. (2009), 'Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences', *Cortex* **45**, 72–79.

Haggerty, M. (1991), 'Evolution by esthetics, an interview with william latham and stephen todd', *IEEE Computer Graphics and Applications* **11**, 5–9.

Handel, S. (1989), *Listening: An Introduction to the Perception of Auditory Events*, The MIT Press.

Handel, S. and Lawson, G. R. (1983), 'The contextual nature of rhythmic interpretation.', *Percept Psychophys.* **34**(2), 103–120.

Handel, S. and Oshinsky, J. S. (1981), 'The meter of syncopated auditory polyrhythms', *Perception and Psychophysics* **30**, 1–9.

Hewlett, W. B. and Selfridge-Field, E., eds (2005), *Music Query: Methods, Models, and User Studies*, MIT Press.

Heyduk, R. G. (1975), 'Rated preference for musical compositions as it relates to complexity and exposure frequency', *Perception and Psychophysics* **17**, 84–91.

Hiller, L. and Isaacson, L. (1959), *Experimental Music: Composition with an Electronic Computer.*, Greenwood Press.

Hiller, L., Isaacson, L. and Baker, R. (1967), 'Computer music from the university of illinois', Heliodor - Vynil (LP). Hiller's notes on reverse of the album.

Hoffmann, P. (2002), 'Towards an Şautomated artŤ: Algorithmic processes in xenakisŠ compositions', *Contemporary Music Review* **21**(2/3), 121–131.

Honing, H. (2011), *Musical Cognition: A science of listening*, Transaction publishers.

Honing, H., Ladinig, O., Háden, G. P. and Winkler, I. (2009), 'Is beat induction innate or learned?', *Annals of the New York Academy of Sciences* **1169**(1), 93–96.

Horner, A. and Goldberg, D. (1991), Genetic algorithms and computer assisted music composition, *in* R. Belew and L. Booker, eds, 'Proceedings of the 4th International Conference on GAs', p. 437Ű441.

Johanson, B. and Poli, R. (1998), Gp-music: An interactive genetic programming system for music generation with automated fitness raters, Technical report, School of Computer Science, The University of Birmingham.

Jones, M. R. (1976), 'Time, our lost dimension: Toward a new theory of perception, attention, and memory', *Psychological Review* **83**(5), 323–355.

Juslin, P. and Sloboda, J. (2010), *Handbook of music and emotion: theory, research, applications*, Oxford University Press.

Kelso, J. A. (1984), 'Phase transitions and critical behavior in human bimanual coordination', *Am J Physiol Regul Integr Comp Physiol* **246**(6), R1000–1004.

Kim, K.-J. and Cho, S.-B. (2006), 'A comprehensive overview of the applications of artificial life', *Artif. Life* **12**(1), 153–182.

Kippen, J. (1988), *The Tabla of Lucknow: a Cultural Analysis of a Musical Tradition*, Cambridge University Press.

Kippen, J. and Bell, B. (1992), *Computer Representations and Models in Music*, Academic Press, chapter Modelling Music with Grammars - Formal Language Representation in the BOL Processor, pp. 207–238.

Kirby, S. (2002), 'Natural language from artificial life', *Artificial Life* **8**(2), 185–215.

Kirke, A. (2011), Application of Intermediate Multi-Agent Systems to Integrated Algorithmic Composition and Expressive Performance of Music, PhD thesis, University of Plymouth.

Kirke, A., Freeman, S., Miranda, E. and Ingram, S. (2011), Application of multi-agent whale modelling to an interactive saxophone and whale duet, *in* 'Proceedings of International Computer Music Conference (ICMC 2011)'.

Komar, A. J. (1971), *Theory of Suspensions*, Princepton University Press.

Kramer, G. (1993), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Perseus Publishing.

Ladinig, O. (2009), Temporal expectations and their violations., PhD thesis, Universiteit van Amsterdam.

Large, E. W. and Jones, M. R. (1999), 'The dynamics of attending: How people track time-varying events', *Psychological Review* **106**(1), 119–159.

Lempel, A. and Ziv, J. (1976), 'On the complexity of finite sequences', *IEEE Transactions on Information Theory* **22**, 75–81.

Lerdahl, F. and Jackendoff, R. (1983), *A Generative Theory of Tonal Music*, MIT Press, Cambridge, Massachusetts.

Lindblom, B. and Sundberg, J. (1972), *Musical acoustics and music theory*, STL-QPSR 4.

Lippo, C. P., O'Brien, M. J., Collard, M. and Shennan, S. J., eds (2006), *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory*, Evolutionary Foundations of Human Behavior Series, Aldine Transaction.

London, J. (2001), *The New Grove Dictionary of Music and Musicians*, Macmillan, chapter Rhythm, pp. 277–309.

London, J. (2004), *Hearing in Time - Psychological Aspects of Musica Meter*, Oxford University Press.

Longuet-Higgins, H. and Lee, C. (1984), 'The rhythmic interpretation of monophonic music', *Music Perception* **1**(4), 424–441.

Macedo, L., Pereira, F., Grilo, C. and Cardoso, A. (1996), Experimental study of a similarity metric for retrieving pieces from structured plan cases: its role in the originality of plan case solutions, *in* 'Second International Conference on Case-Based Reasoning', pp. 575–586.

Machado, F. P. (2006), Inteligência Artificial e Arte, PhD thesis, Universidade de Coimbra.

Machado, P. and Cardoso, A. (1998), Computing aesthetics, *in* 'Proceedings of the XIVth Brazilian Symposium on Artificial Intelligence; Advances in Artificial Intelligence', Vol. 1515 of *LNCS*, Springer, Porto Alegre, Brazil.

Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W. and Davis, R. B. (2005), 'Zipf's law, music classification, and aesthetics', *Comput. Music J.* **29**(1), 55–69.

Martin, J. G. (1972), 'Rhythmic (hierarchical) versus serial structure in speech and other behavior.', *Psychological Review* **79**(6), 487–509.

Martins, J. M., Gimenes, M., Manzolli, J. and Maia Jr., A. (2005), Similarity measures for rhythmic sequences, *in* 'Proceedings of the 10th Brazilian Symposium on Computer Music (SBCM)', Belo Horizonte (Brazil).

McElreath, R. and Boyd, R. (2007), *Mathematical Models of Social Evolution: A Guide for the Perplexed*, University Of Chicago Press.

Merker, B. (1999), Synchronous chorusing and human origins., *in* N. L. Wallin, B. Merker and S. Brown, eds, 'The Origins of Music', MIT Press, Cambrige (MA), USA, pp. 315–328.

Merriam, A. P. (1964), *The Anthropology of Music*, Northwestern University Press.

Meyer, L. B. (1956), *Emotion and meaning in music*, University of Chicago Press.

Millen, D. (1990), Cellular automata music, *in* 'Proceedings of the 1990 International Computer Music Conference'.

Miller, G. (1999), Evolution of music through sexual selection., *in* N. L. Wallin, B. Merker and S. Brown, eds, 'The Origins of Music', MIT Press, Cambridge (MA), USA, pp. 329–360.

Minsky, M. (1981), 'Music, mind, and meaning', *Computer Music Journal* **5**(3), 28–44. available at http://web.media.mit.edu/ minsky/papers/MusicMindMeaning.html.

Minsky, M. (1988), *Society of Mind*, Simon & Schuster.

Miranda, E. (2001), *Composing Music with Computers*, Focal Press.

Miranda, E. R. (2002*a*), 'Emergent sound repertoires in virtual societies', *Computer Music Journal (MIT Press)* **26**(2), 77–90.

Miranda, E. R. (2002*b*), Mimetic development of intonation, *in* 'Proceedings of the Second International Conference on Music and Artificial Intelligence', ICMAI '02, Springer-Verlag, London, UK, pp. 107–118.

Miranda, E. R. (2003), 'On the music of emergent behaviour - what can evolutionary computation bring to the musician?', *Leonardo* **36**, 55–58.

Miranda, E. R. (2004), 'At the crossroads of evolutionary computation and music: Self-programming synthesizers, swarm orchestras and the origins of melody', *Evolutionary Computation* **12**(2), 137–158.

Miranda, E. R. (2008), Autonomous development of singing-like intonations by interacting babbling robots, *in* 'Proceedings of International Computer Music Conference - (ICMC2008)', Belfast (UK).

Miranda, E. R. and Biles, J. A., eds (2007), *Evolutionary Computer Music*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Miranda, E. R., Kirby, S. and Todd, P. M. (2003), 'On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars', *Contemporary Music Review* **22**(3), 91–111.

Miranda, E. R. and Kirke, A. (2010), Game of life music, *in* A. Adamatzky, ed., 'Game of Life Cellular Automata', Springer Verlag, pp. 489–501.

Miranda, E. R., Kirke, A. and Zhang, Q. (2010), 'Artificial evolution of expressive performance of music: An imitative multi-agent systems approach', *Computer Music Journal* **34**(1), 80–96.

Mongeau, M. and Sankoff, D. (1990), 'Comparison of musical sequences', *Computers and the Humanities* **24**(3), 161–175.

Mont-Reynaud, B. and Goldstein, M. (1985), On finding rhythmic patterns in musical lines., *in* 'Proceedings of the International Computer Music Conference', San Francisco, CA (US).

Mozer, M. C. (1994), 'Neural network music composition by prediction: Exploring the benefits of psychophysical constraints and multiscale processing.', *Connection Science* **6**, 247–280.

Murray-Rust, D. (2003), Virtualatin - agent based percussive accompaniment, Master's thesis, University of Edinburgh.

Murray-Rust, D. (2008), Musical Acts and Musical Agents: theory, implementation and practice, PhD thesis, University of Edinburgh.

Murray-Rust, D., Smaill, A. and Edwards, M. (2006), Mama: An architecture for interactive musical agents, *in* 'ECAI', pp. 36–40.

Murray-Rust, D., Smaill, A. and Maya, M. C. (2005), Virtualatin - towards a musical multi-agent system, *in* 'Proceedings of the Sixth International Conference on Computational Intelligence and Multimedia Applications', IEEE Computer Society, Washington, DC, USA, pp. 17–22.

Nauert, P. (2007), 'Division- and addition-based models of rhythm a computer-assisted composition system', *Computer Music Journal* **31**(4), 59–70.

Nierhaus, G. (2009), *Algorithmic Composition: Paradigms of Automated Music Generation*, Springer-Verlag.

North, A. and Hargreaves, D. (2008), *The Social and Applied Psychology of Music*, Oxford University Press.

Orpen, K. S. and Huron, D. (1992), 'Measurement of similarity in music: A quantitative approach for non-parametric representations', *Computers in Music Research* **4**, 1–44.

Oudeyer, P. (2006), *Self-organization in the evolution of speech*, Studies in the evolution of language, Oxford University Press.

Pachet, F. (2000), Rhythms as emerging structures, *in* 'Proceedings of the International Computer Music Conference', Berlin, Germany.

Pachet, F. and Roy, P. (1998), Formulating constraint satisfaction problems on part-whole relations: The case of automatic musical harmonization., *in* 'Constraint Techniques for Artistic Applications Workshop. ECAI'98', Brighton, UK.

Pachet, F. and Roy, P. (2001), 'Musical harmonization with constraints - a survey', *Constraints* **6**(1), 7–19.

Page, L. (2001), 'Method for node ranking in a linked database'. United States patent US 6,285,999 B1.

Parncutt, R. (1994), 'A perceptual model of pulse salience and metrical accent in musical rhythms', *Music Perception* **11**(4), 409–464.

Patel, A. D. (2007), *Music, Language, and the Brain*, Oxford University Press, USA.

Pereira, F., Grilo, C., Macedo, L. and Cardoso, A. (1997), A structured framework for representing time in a generative composition system, *in* 'TIME '97: Proceedings of the 4th International Workshop on Temporal Representation and Reasoning (TIME '97)', IEEE Computer Society, Washington, DC, USA, p. 168.

Pinker, P. (1997), *How the mind works*, Norton, New York.

Povel, D.-J. (1981), 'Internal representation of simple temporal patterns.', *Journal of Experimental Psychology: Human Perception and Performance.* **7**(1), 3–18.

Povel, D. J. and Essens, P. (1985), 'Perception of temporal patterns', *Music Perception* **2**(4), 411–441.

Povel, D. J. and Okkerman, H. (1981), 'Accents in equitone sequences.', *Percept Psychophys* **30**(6), 565–572.

Pressing, J. (1998), 'Cognitive complexity and the structure of musical patterns.', *Noetica* **3**, 1–8.

Reich, S. (2004), *Writings on Music, 1965-2000*, Oxford University Press.

Repp, B. (2009), 'A flled duration illusion in music: Efects of metrical subdivision on the perception and production of beat tempo', *Advances in Cognitive Psychology* **5**, 114–134.

Reynolds, C. W. (1987), 'Flocks, herds and schools: A distributed behavioral model', *SIGGRAPH Comput. Graph.* **21**, 25–34.

Roads, C. (1996), *The Computer Music Tutorial*, The MIT press, Cambridge, Massachusetts 02142.

Roads, C. (2001), *Microsound*, The MIT press, Cambridge, Massachusetts 02142.

Romero, J. and Machado, P., eds (2007), *The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music*, Natural Computing Series, Springer Berlin Heidelberg.

Rossing, T. D., Moore, R. F. and Wheeler, P. A. (2001), *The Science of Sound*, 3rd edn, Addison-Wesley.

Sachs, C. (1953), *Rhythm and Tempo*, W. W. Norton and Company, Inc.

Sampaio, P. A., Ramalho, G. and Tedesco, P. A.-c. (2008), 'CinBalada: a multiagent rhythm factory', *Journal of the Brazilian Computer Society* **14**, 31–49.

Santorini, B. and Kroch, A. (2007), 'The syntax of natural language: An online introduction using the trees program'. http://www.ling.upenn.edu/ beatrice/syntax-textbook (last access in 16 Apr 2011).

Sasahara, K. and Ikegami, T. (2003), Coevolution of birdsong grammar without imitation, *in* 'ECAL03', pp. 482–490.

Sasahara, K. and Ikegami, T. (2004), Song grammars as complex sexual displays, *in* 'Artificial Life IX'.

Saunders, R. (2002), Curious Design Agents and Artificial Creativity = A Synthetic Approach to the Study of Creative Behaviour, PhD thesis, University of Sydney.

Saunders, R. and Gero, J. (2002), How to study artificial creativity, *in* 'Proceedings of C&C'02', Loughborough, United Kingdom.

Schottstaedt, B. (1989), Automatic species counterpoint, *in* M. V. Mathews and J. Pierce, eds, 'Current Directions in Computer Music Research', MIT Press, pp. 225–262.

Schwanauer, S. M. and Levitt, D. A., eds (1993), *Machine models of music*, MIT Press, Cambridge, MA, USA.

Shannon, C. (1950), 'Prediction and entropy of printed english', *The Bell System Technical Journal* **30**, 50–64.

Shmulevich, I. and Povel, D. J. L. (2000), 'Measures of temporal pattern complexity', *Journal of New Music Research* **29**(1), 61–70.

Sims, K. (1993), 'Interactive evolution of equations for procedural models', *The Visual Computer* **8**, 466–476.

Smith, L. M. and Honing, H. (2006), Evaluating and extending computational models of rhythmic syncopation in music, *in* 'Proceedings of the 2006 International Computer Music Conference', pp. 688–691.

Snyder, B. (2001), *Music and Memory: An Introduction*, MIT Press.

Spector, L. and Klein, J. (2002), Complex adaptive music systems in the breve simulation environment, *in* B. et al., ed., 'Proceedings of the 8th IntŠl Conf. on the Simulation and Synthesis of Living Systems', University of New SouthWales, Sydney, pp. 163–170.

Steedman, M. J. (1977), 'The perception of musical rhythm and metre', *Perception* **6**(5), 555–569.

Steels, L. (1997), 'The synthetic modeling of language origins', *Evolution of Communication* **1**(1), 1–34.

Steels, L. (2003), 'Evolving grounded communication for robots', *Trends in Cognitive Science* **7**(7), 308–312.

Sternberg, S. and Knoll, R. L. (1982), Timing by skilled musicians, *in* D. Deutsch, ed., 'Psychology of music', Academic Press, New York, pp. 181–239.

Stevens, L. T. (1886), 'On the timse sense', *Mind* **11**(43), 393–404.

Sundberg, J. and Lindblom, B. (1992), Generative theories in language and music descriptions, *in* S. Schwanauer and D. Levitt, eds, 'Machine Models of Music', MIT Press, Cambridge, MA, USA, pp. 263–286.

Takeda, H., Nishimoto, T. and Sagayama, S. (2003), Automatic rhythm transcription from multiphonic midi signals, *in* 'ISMIR 2003'.

Tanguiane, A. S. (1993), *Artifical Perception and Music Recognition*, Springer-Verlag.

Temperley, D. (2004), *The Cognition of Basic Musical Structures*, The MIT Press.

Thomas, E. A. C. and Brown, I., J. (1974), 'Time perception and the filled-duration illusion.', *Perception & Psychophysics* **16**, 449–458.

Thul, E. and Toussaint, G. T. (2008), Rhythm complexity measures: A comparison of mathematical models of human perception and performance., *in* J. P. Bello, E. Chew and D. Turnbull, eds, 'ISMIR', pp. 663–668.

Todd, P. (1999), Simulating the evolution of musical behavior, *in* N. L. Wallin, B. Merker and S. Brown, eds, 'The Origins of Music', MIT Press, Cambridge (MA), USA, chapter 20, pp. 361–388.

Todd, P. M. and Werner, G. (1999), Frankensteinian methods for evolutionary music composition, *in* N. Griffith and P. M. Todd, eds, 'Musical networks: Parallel distributed perception and performance', MIT Press/Bradford Books, Cambridge (MA), USA, pp. 313–339.

Toop, D. (2006), The generation game: Experimental music and digial culture, *in* C. Cox and D. Warner, eds, 'Audio Culture: readings in modern music', Continuum.

Toussaint, G. T. (2002), A mathematical analysis of african, brazilian, and cuban clave rhythms, *in* 'In Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science', Southwestern College, Kansas,.

Tsang, C. P. and Aitken, M. (1991), Harmonizing music as a discipline of constraint logic programming (, *in* 'Proceedings of ICMC '91', Montréal.

Tucker, M. (2006), 'Steve reich - the south bank show', ITV documentary.

Unemi, T. and Bisig, D. (2005), Playing by interaction among two flocking species and a human, *in* 'Proceedings of the Third IntŠl Conf. on Generative Systems in Electronic Arts', Melbourne, Australia, p. 171Ű179.

Vallejo, E. E. and Taylor, C. E. (2005), Interactions between learning and evolution in simulated avian communication, *in* 'ECAL05', Springer, pp. 634–643.

van den Broek, E. and Todd, P. (2009), 'Evolution of rhythm as indicator of mate quality', *Musicae Scientiae* **Special Issue 2009-2010**, 369–386.

van den Broek, E. and Todd, P. M. (2003), Piep piep piep - ich hab' dich lieb: Rhythm as an indicator of mate quality., *in* 'ECAL03', pp. 425–433.

von Vierordt, K. (1868), *Der Zeitsinn nach Versuchen*, H. Laupp.

Walker, E. L. (1970), 'Complexity and preference in animals and men', *Annals of the New York Academy of Sciences* **169**, 619–652.

Wallin, N., Merker, B. and Brown, S., eds (2001), *The origins of music*, Bradford Books, MIT Press.

Wearden, J. H., Norton, R., Martin, S. and Montford-Bebb, O. (2007), 'Internal clock processes and the filled-duration illusion.', *J Exp Psychol Hum Percept Perform* **33**(3), 716–729.

Werner, G. and Todd, P. (1997), Too many love songs: Sexual selection and the evolution of communication, *in* P. Husbands and I. Harvey, eds, 'ECAL97', MIT Press, Cambridge, MA, pp. 434–443.

Wilson, S. (2008), Spatial swarm granulation, *in* 'Proceedings of the 2008 International Computer Music Conference', SARC, ICMA, pp. 4–7.

Windsor, L. and Desain, P. (2000), Introduction: Multiple perspectives on rhythm perception and production., *in* P. D. . L.Windsor, ed., 'Rhythm: Perception and production', Swets & Zeitlinger., pp. xi–xvi.

Wittgenstein, L. (1958), *Philosophical Investigations*, Basil Blackwell. Translated by G. E. M. Anscombe.

Wundt, W. (1874), *Grundz
"uge der physiologischen Psychologie*, Engelmann, Leipzig.

Xenakis, I. (1989), 'Concerning time', *Perspectives of New Music* **27**(1), 84Ű92.

Yeston, M. (1976), *The Stratification of Musical Rhythm*, Yale University Press.

Zahavi, A. and Zahavi, A. (1997), *The handicap principle: A Missing Piece of Darwin's puzzle*, Oxford University Press.

Zicarelli, D. (1987), 'M and jam factory', *Computer Music Journal* **11**(4), 13–29.

Bound copies of published papers

ENGINEERING - THE ROLE OF SOCIAL PRESSURE : A NEW ARTIFICIAL LIFE APPROACH TO SOFTWARE FOR GENERATIVE MUSIC

By

JOAO M. MARTINS* EDUARDO R. MIRANDA**

ABSTRACT

Most current research into computers and music focus on the development of media technology for delivering music to consumers (e.g., MP3 format, Internet search engines, and so on). This research focuses on the development of technology for musical creativity. This paper focuses on a particular technology currently being developed, based on Artificial Life (A-Life). The Artificial Life (A-life) approach to the development of software for music is a promising new development. However, the vast majority of existing A-life-based systems for musical composition employ Genetic Algorithms (GA) to produce musical melodies, rhythms and so on. In these systems, music parameters are represented as genotypes and GA operators are applied on these representations to produce music according to the given fitness criteria. It is suggested that strictly GA-based methods suffer from the fact that musical composition should not be constrained by a definite set of fitness criteria. Moreover music is largely a cultural phenomenon driven by social pressure and this is cumbersome to model with standard GA alone. An alternative approach is proposed to using strictly GA-based methods: the design of evolutionary algorithms that consider music as a cultural phenomenon whereby social pressure plays an important role in the development of musical conventions. This paper introduces three algorithms of the authors' own design: popularity, transformation and complexity algorithms, respectively. Tools are also devised for extracting information about the behaviour of the algorithms in many different ways, providing the means to study the outcomes systematically.

Keywords: Evolutionary Computer Music, Artificial Life systems for musical composition, Computer models of music.

INTRODUCTION

From the time of discovery almost three thousand years ago, the direct relationship between the pitch of a note and the length of a string or pipe, to the latest computer models of human music cognition and intelligence, musicians have always looked at science to provide new and challenging paradigms to study and create music. The field of Computer Music is as old as Computer Science. Computers have been programmed to play music as early as the early 1950's when Geoff Hill programmed the CSIR Mk1 computer, in Sydney, Australia, to play the popular musical melodies (Doornbusch, 2005). Nowadays, the computer is becoming increasingly ubiquitous in all aspects of music. Applications of computer technology to music ranges from systems for musical composition to systems for distribution of music on the Internet. The implementation of such applications often demands the skiliful combination of Software Engineering and artistic creativity. Whereas most current research into computers and music focuses on the development of media technology for delivering music to consumers (e.g., MP3 format, Internet search engines, and so on). This research focuses on the development of technology for musical creativity. That is, technology to aid musicians to create content for the media. This paper focuses on a particular technology that the authors are developing, which is based on Artificial Life (A-Life).

The A-Life approach to music is a promising new development for composers. It provides an innovative and natural means for generating musical ideas from a specifiable set of primitive components and processes reflecting the compositional process of generating a variety of ideas by brainstorming followed by selecting the most promising one for further iterated refinement (Kim and Cho, 2006). We are interested in implementing systems for composition using A-Life-based models of cultural transmission; for example, models of the development and maintenance of musical styles within a particular cultural contexts and their reorganization and adaptation in response to cultural exchange.

Existing A-Life-based systems for musical composition normally employ Genetic Algorithms (GA) to produce musical melodies, rhythms and so on. In these systems music parameters are represented as "genotypes" and GA operators are applied on these representations to produce music according to the given fitness criteria. Because of the highly symbolic nature of Western music notation, music parameters are suitable for GA-based processing and a number of musicians, including the authors, have used such systems to compose music. However, two problems have been identified with the GA-based approach to generative music. Firstly, a musical composition should not be driven by a constant set of fitness criteria. And secondly, music is largely a cultural phenomenon driven by social pressure and this is cumbersome to model with standard GA alone.

The first problem emerges from the fact that 'music is not an exact science'. For example, it differs from Engineering, whereas the success of a piece of Engineering would normally be measured by its ability to match a number of functional requirements effectively, the success of a piece of music cannot be measured so objectively. Indeed, whereas good engineers are praised for following the rules of their métier strictly, good composers (at least in the Western music tradition) are praised for clever violations of musical conventions. Moreover, in most cases, composers do not explicitly know a priority how a new piece of music will sound like, until it is completed and indeed performed. Therefore, rather than tools to generate efficient solutions to problems automatically, composers need tools to explore a vast space of possible outcomes. Biles (1994) proposed an interesting approach to implement GA-based systems for the exploration of a space of musical possibilities, which takes into consideration the evaluation of the user; that is, the user evaluates the fitness of each generation of "solutions". This is surely a very interesting idea, but this slows down the compositional process enormously. Biles is aware of this problem, which he refers to as the "fitness bottleneck" problem.

The second problem is largely related to a problem that is endemic in the field of Computer Music, which is the tendency to design systems to generate music from algorithms that were not designed for music in the first instance. For example, in the late 1980s it became fashionable to implement systems that generated music from fractals (Mandelbrot, 1982). There was a tendency at that time to overstate the adequacy of fractals as generators of music. Nowadays, we may be witnessing a similar case of overstatement on the adequacy of GA as generators of music. Although we acknowledge that there have been a few rather successful stories (Biles, 1994), we believe that additional A-Life-based methods need to be developed in order to move the exciting field of *Evolutionary Computer Music* (Miranda and Biles, 2007; Miranda, 2004) forward.

One way forward is to build systems with A-Life algorithms designed or suitably modified to address musical issues. The work presented in this paper contributes to this line of thought by looking into the design of algorithms that consider music as a cultural phenomenon whereby social pressure plays an important role in the development of musical conventions. A plausible method to embed social dynamics in such algorithms is to design them within the framework of interacting autonomous software agents.

This paper introduces three algorithms of the authors' own design, referred to as *popularity*, *transformation* and *complexity* algorithms, respectively. These algorithms are used to implement a system for the composition of rhythms where the user can generate rhythmic sequences and also monitor the behaviour of the system. The system offers the ability to extract information about the behaviour of the agents and the evolving rhythms in many different ways, providing the composers with means to explore the outcomes systematically. This paper will focus on the algorithms themselves, the information that one can extract about their behaviours and the analyses of the behaviours. An in-depth discussion on how the algorithms are used artistically to compose pieces of music falls beyond the scope of this paper.

By way of related research, the work by De Boer (De Boer, 1999) on modeling the emergence of vowel systems by means of imitations games is cited. Also, Miranda (Miranda, 2002) has developed a model of the emergence of intonation systems using imitation games. This research is inspired by the work developed by research into gaining a better understanding of the evolution of language with computer models (Kirby, 2002; Vogt, 2000), particularly the work of Steels (Steels, 1995) on language imitation games with software agents. Basically an imitation game consists of one agent picking a random sound from its repertoire and the other agent trying to imitate it. Then, a feedback is given about the success of the imitation. On the basis of this feedback, the agents update their memories.

1. The Agents

In this section, the agents and their "cognitive ability" are introduced; that is, the operations that they are able to perform on rhythms.

The agents are identical to each other and the number of agents in a group may vary. The agents move in a virtual 2D space (Figure 1) and they normally interact in pairs. Essentially, the agents interact by playing rhythmic sequences to each other, with the objective of developing repertoires of rhythms collectively. At each round, each of the agents in a pair plays one of the two different roles: the *player* and the *listener*. The agents may perform operations on the rhythms that they play to each other, depending on the interaction algorithm being used and on the status of the emerging repertoire. The agents are provided with a memory to store the emerging rhythms and other associated information.

The fundamental characteristic of human beings is that we are able to perceive, and more importantly, to produce an isochronous pulse (Handel, 1989). Moreover, humans show a preference for rhythms composed of

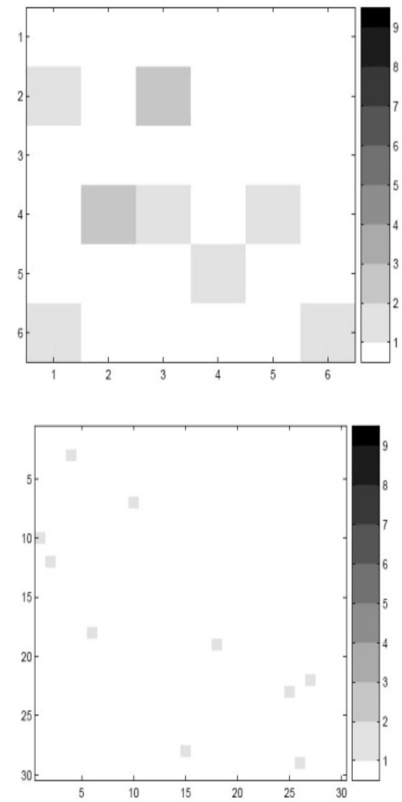


Figure 1. 2D virtual worlds with different sizes holding 10 agents. A darker colour indicates a cluster. (This will be clarified in due course)

integer ratios of the basic isochronous pulse (Drake and Bertrand, 2001). Therefore, rhythms are represented here as interonset intervals in terms of small integer ratios of an isochronous pulse (Figure 2).

1.1 Transformations of Rhythms

At the core of the mechanism by which the agents develop rhythmic sequences is a set of basic transformation operations. These operations enable the agents to generate new rhythmic sequences and change the rhythmic sequences that they learn as the result of the interactions with other agents. The transformation operations are as follows:

- Divide a rhythmic figure by two (Figure 3a).

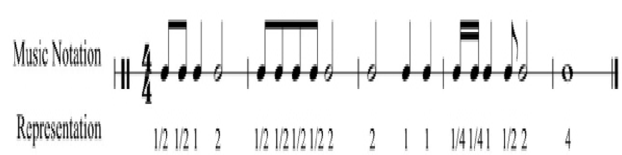


Figure 2. Standard music notation of a rhythmic sequence and its corresponding interonset representation

- Merge two rhythmic figures (Figure 3b).
- Add one element to the sequence (Figure 3c).
- Remove one element from the sequence (Figure 3d).

The definition of these transformations were inspired by the dynamical systems approach to study human bimanual coordination (Kelso, 1984) and is based on the notion that two coupled oscillators will converge to stability points at frequencies related by integer ratios (Beek et al., 2000). Furthermore, common music notation facilitates these types of transformations. Other transformations that divide a figure into three, five and other prime numbers are defined, but the impact of these additional transformations on the model is beyond the scope of this paper. Addition and removal transformations were introduced to increase the diversity in the pool of rhythms and to produce rhythms of different lengths.

1.2 Measurement of Similarity of Rhythms

The agents are programmed with the ability to measure the degree of similarity of two rhythmic sequences. This measurement is used when they need to assess the similarity of the rhythms that they play to each other. Also, this algorithm is used to measure the similarity between the repertoires of rhythms from different agents.

In the previous paper (Martins et al., 2005), a method to measure the degree of similarity between two sequences of symbols was introduced by comparing various subsequences at various levels. The result is a vector,

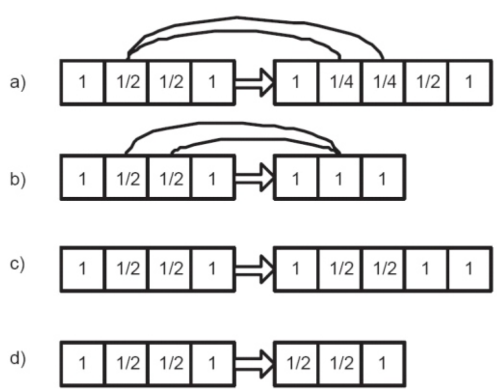


Figure 3(a - d). Examples of rhythmic transformations

referred to as the *Similarity Coefficients Vector* (SCV), which contains the interim results of the comparisons between subsequences. For the present work, a version of the SCV method that deals with rhythmic sequences is devised, which is introduced below.

Let us define the block distance between two sequences containing the same number of elements as follows (1):

$$d\left(\mathbf{v},\mathbf{w}\right) = \sum_{i=1}^{n} |v_i - w_i| \qquad \text{...............(1)}$$

Where v and w are the two sequences (vectors) that are being compared, and v_i and w_i are the individual components of these vectors.

After obtaining the resulting evaluation of the block distances on a given level (length of a subsequence), we can write a matrix for the *k*-level, corresponding to the comparison of all the subsequences with length *k* between the two main sequences (2):

$$\mathbf{D}^{(k)} = \begin{bmatrix} d(\mathbf{v}_1^{(k)},\mathbf{w}_1^{(k)}) & \dots & d(\mathbf{v}_1^{(k)},\mathbf{w}_{(m-k+1)}^{(k)}) \\ d(\mathbf{v}_2^{(k)},\mathbf{w}_1^{(k)}) & \dots & d(\mathbf{v}_2^{(k)},\mathbf{w}_{(m-k+1)}^{(k)}) \\ \vdots & \vdots & \vdots \\ d(\mathbf{v}_{(n-k+1)}^{(k)},\mathbf{w}_1^{(k)}) & \dots & d(\mathbf{v}_{(n-k+1)}^{(k)},\mathbf{w}_{(m-k+1)}^{(k)}) \end{bmatrix} \qquad \text{...............(2)}$$

Where *d* is the distance defined by Equation 1 between all the subsequences $v^{(k)}$ of v and all the subsequences $w^{(k)}$ of w. Next, let us define the *k*-level *Similarity Coefficient* as follows (3):

$$c^{(k)}(\mathbf{v},\mathbf{w}) = \frac{z(k)}{(n-k+1)(m-k+1)} \qquad \text{...............(3)}$$

Where, *z(k)* is the number of zeros in the matrix $D^{(k)}$. Roughly speaking, the similarity coefficient measures the sparsity of the matrix $D^{(k)}$. The higher the coefficient *c(k)*, the higher is the similarity between the subsequences of level *k*.

Next, we can collect all the *k*-level coefficients in a vector referred to as *Similarity Coefficient Vector* (SCV). This is defined as follows (4):

$$\mathbf{C} = \left[c^{(1)}, c^{(2)}, \dots, c^{(min(m,n))}\right] \qquad \text{...............(4)}$$

Figure 4 shows an example of building a 3-level Distance Matrix and its respective SCV is $SCV = [0.4167\ 0.1333\ 0.1250\ 0]$.

From this vector we can obtain a scalar value in order to establish a comparative analysis between larger sets of rhythms, such as the repertoires of two agents. We can take the rightmost nonzero value from the SCV, which corresponds to the higher level where two matching sequences can be found. We can either take a weighted sum of the SCV values or the average of all values, as follows (5):

$$SCV_{av} = \frac{1}{min(m,n)} \sum_{k=1}^{min(m,n)} SCV(k) \qquad \text{..............(5)}$$

where $SCV(k)$ are the coefficients of similarity for each of the k levels.

The next step is to compare the repertoire of the agents in order to observe the development of relationships amongst the agents in a group of agents; for instance, to observe if the agents form distinct sub-groupings.

The similarity of the repertoire of rhythms amongst the agents in a group is computed by creating a matrix of SCV_{av} values of the repertoires of all pairs of agents. Matrices with the columns and rows corresponding to the number of rhythms in the memory of each agent reveals the similarity between their repertoires (Figure 5).

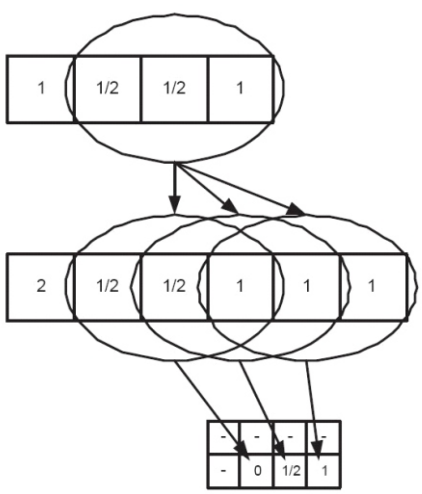


Figure 4. Example of building a 3-level Distance Matrix

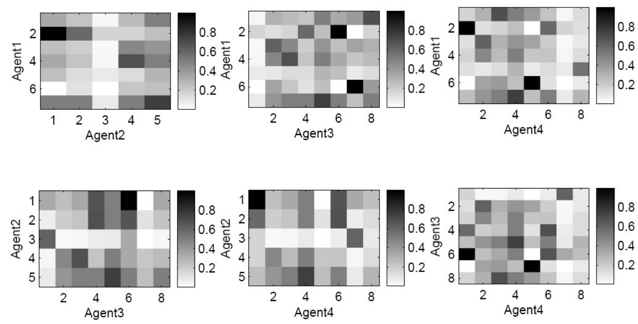


Figure 5. Similarity matrices between the repertoires of 4 agents. The darker the colour, the more similar the rhythms

By collapsing both the rows and the columns of the matrices, and taking the maximum values for each of them and an averaged sum,the scalar of similarity between repertoires is obtained, as follows (6):

$$SimRep_{k,l} = \frac{1}{nR_{Ak} + nR_{Al}} \left[\sum_{i=1}^{nR_{Ak}} max(SCV_{av})_{rows} + \sum_{j=1}^{nR_{Al}} max(SCV_{av})_{cols} \right] \qquad \text{..............(6)}$$

Where the first term corresponds to the sum of the maximum values of the SCV_{av}, defined in Equation 5, for every row, and the second term is the correspondent for every column; nR_{Ak} and nR_{Al} are the number of rhythms in the repertoire of the compared agents.

In the case shown in Figure 6, it would be $SimRep_{k,l} = 0.7$ or, conversely, the distance between the repertoires of both the agents k and l would be $DistRep_{k,l} = 1 - SimRep_{k,l} = 0.3$. The values of 0.65 and 0.8 in Figure 6 corresponds to the similarity of the repertoires from the point of view of each agent, which is used to

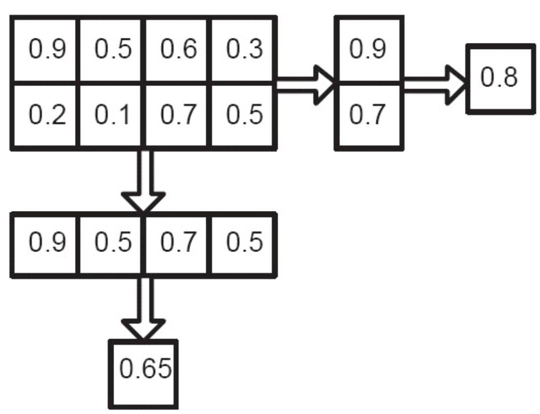


Figure 6. Scalar value of similarity between repertoires

generate proximity matrices and graphs to monitor the behaviour of the evolving repertoire of rhythms and the agents.

Finally, the development of repertoires of rhythms for a group of agents as a whole can be observed by conducting a hierarchical cluster analysis of all distance measures between the agents (*DistRep*). This cluster analysis produces a dendrogram using a linkage method based on an unweighted average distance, also known as group average in which the distance between two clusters *A* and *B*, D_{AB}, is given by the following equation (7)

$$D_{AB} = \frac{1}{N_A . N_B} \sum_i d_i \qquad \text{..............(7)}$$

where N_A and N_B are the number of elements in *A* and *B*, and d_i are pairwise distances between the elements of clusters *A* and *B*. The hierarchical cluster analysis produces a dendrogram of the type shown in Figure 11. This dendrogram is drawn through an iterative process until all the individuals or clusters are linked.

1.3 Measurement of Complexity of Rhythms

The complexity of a rhythmic sequence is measured as follov

$$Complexity = \frac{nF + \sum_{i=1}^{nF} n_i}{\sum_{i=1}^{nF} T_i} \qquad \text{...............(8)}$$

Where *nF* is the number of rhythmic figures contained in the sequence, n_i is the value of the numerator of a rhythmic figure, and T_i is the relative length of a rhythmic figure, considering that each rhythmic figure is a fraction of the pulse.

This is a computationally cost effective method to measure the complexity of a rhythmic sequence. It is important to bear in mind that our implementation ensures that there are no reducible fractions included in the sequence, meaning that there is always a single numerical representation for a given rhythm. Figure 7 shows an example of a graph plotting the complexity of a sequence of relative interonset intervals [1, 1] as it is transformed thirty times recurrently.

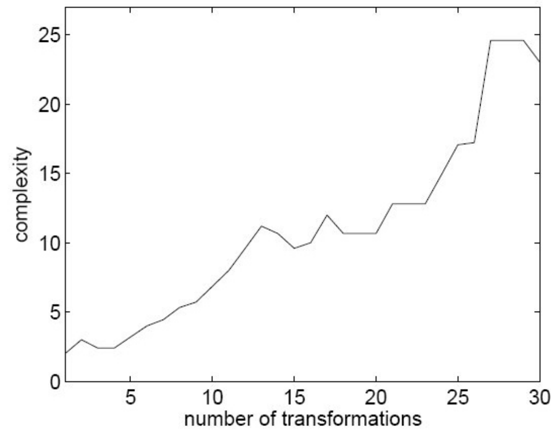


Figure 7. The complexity of a given rhythmic sequence increases in function of the number of transformations applied to it

2. The Algorithms

This section introduces the three proposed algorithms and the respective analysis methods that are implemented. Each algorithm is introduced in the context of illustrative experiments aimed at studying the development of repertoires of rhythmic sequences from three different perspectives:

- From the perspective of an individual agent.
- From the perspective of a group of agents, referred to as *the society*.
- From the perspective of the developed rhythms.

From the perspective of an individual agent, the development of the size and the complexity of the repertoire of individual agents is analyzed. From the perspective of the society, the values of the corresponding individual measures from the agents, as well as the similarity between the agents and how they are clustered in terms of the rhythms they share are analyzed. Finally, from the perspective of the developed rhythms, their lifetime, the amount of rhythmic sequences that the society develops and the degree to which the agents share similar rhythms are measured. The lifetime of a rhythmic sequence is traced by counting the number of agents that hold the sequence in their memories during the interactions. Figure 8 shows graphs illustrating these various types of analyses.

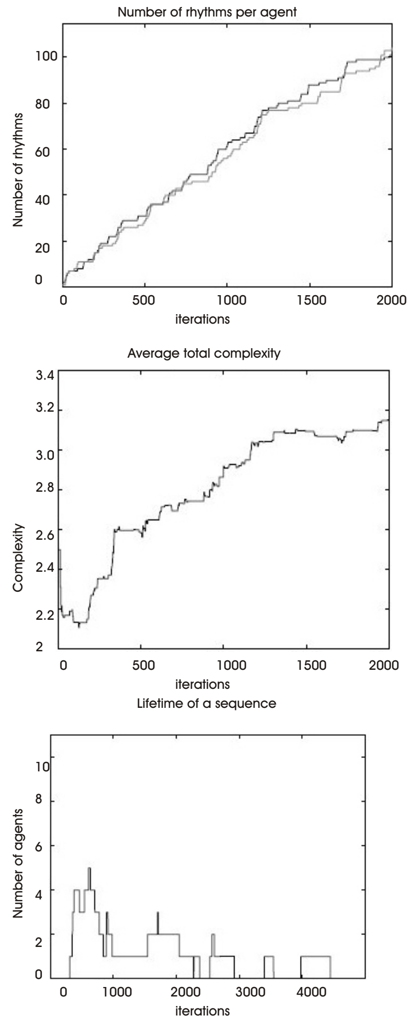


Figure 8. a) Development of the size of the repertoire for different agents; b) Complexity of the rhythms of the society; c)Number of agents sharing a particular rhythm

The experiments were run for 5000 iterations each for a number of times, with the objective of observing the behaviour of the agents, the society and the evolving rhythms, under different conditions. The experiments were run with societies of 3, 10 and 50 agents. For some of the experiments the lifetime of the agents were limited to 1000 iterations; when an agent dies, another is born. Sometimes the algorithms take into account the movement of the agents in the 2D space, which may or may not influence the nature of the iterations.

2.1 The Popularity Algorithm

Popularity is a numerical parameter that each agent attributes to a rhythm in its repertoire. This parameter is modified both by the (agent-) listener and by the (agent-) player when they interact with each other. If the listener recognises a rhythm (that is, if it holds the "perceived" rhythm in its repertoire), then it will increase the popularity index of this rhythm and will give a positive feedback to the player. A positive feedback is an acknowledgment signal, which will prompt the player to increase the popularity index of the rhythm in question in its repertoire. Conversely, if the listener does not recognize the rhythm, then it will add it to its repertoire and will give a negative feedback to the player. This negative feedback will cause the player to decrease the popularity index of this rhythm. Furthermore, there is a memory loss mechanism whereby after each iteration all the rhythms have their popularity index decreased by a small value of 0.05. This accounts for a natural drop in the popularity index due to ageing. A diagram summarising the popularity algorithm is displayed in Figure 9.

Figure 10 shows the results after 5000 iterations of the popularity algorithm without population renewal.

Figure 10a displays the development of the repertoire of individual agents and Figure 10b displays the corresponding average across all the agents. Here the repertoires of the agents grow steadily upto approximately 1000 iterations and subsequently oscillates around a stable point. Figure 10c displays the

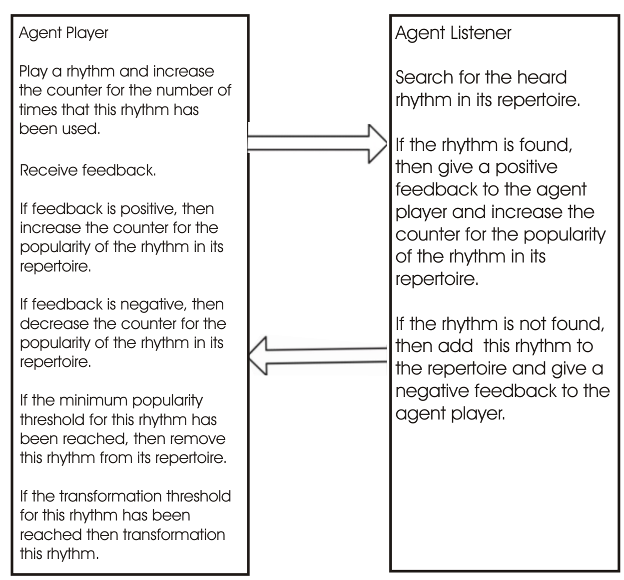


Figure 9. The popularity algorithm

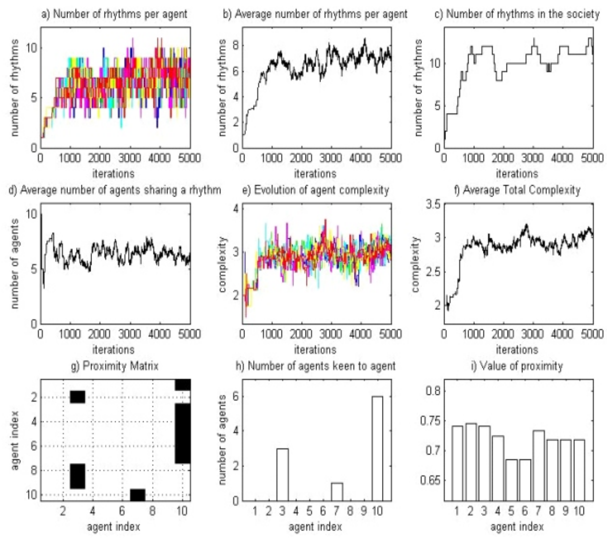


Figure 10(a-i). Results from a typical run of the popularity algorithm with 10 agents

development of the repertoire of the whole society being a direct consequence of the lifetime of each rhythm. The average number of agents sharing a rhythm (Figure 10d) is calculated by summing the instant number of agents sharing a rhythm for all rhythms, and dividing the result by the number of rhythms currently present in the society (Figure 10c). This graph (Figure 10d) provides the means to assess the global behaviour of the society; for instance, if it develops coherently in terms of the popularity of the existing rhythms. Figure 10e represents the development of complexity of the individual agents and Figure 10f gives the corresponding average. Initially, the size and complexity of the repertoire of individual agents are very close to average, but this trend is replaced quickly by the repertoires of different sizes amongst the agents.

The last three graphs show the degree of similarity between the repertoires of the agents according to the similarity measure defined earlier. Figure 10g displays the information about the identity of the agent with whom each agent relates most; i.e., one which has the highest similarity value. The graph in Figure 10h shows the agents that are regarded by others as being most similar to them. In this case, it shows that agent 3 has three agents with similar repertoires, and agent 10 is the one that concentrates the highest number of keen agents, having six agents considering its repertoire to be more similar to theirs.

Hierarchical cluster analysis is conducted in order to observe the groupings of agents according to the similarity of their repertoires. Figure 11 shows the dendrogram containing elements of three societies of 10 agents each: Society 1 comprises agents 1 to 10, Society 2 comprises agents 11 to 20, and Society 3 the remaining 21 to 30. By comparing the three societies we can observe the three clearly independent clusters, which were developed separately in three separate runs with the same set of parameters. In addition to the previous observations, this suggests that the repertoires that emerged from the popularity algorithm display diversity, are stable in terms of size, and are coherent within their respective societies. The differences in the clusters within a given society can also be observed.

By allowing the agents to move in their environment, it is also been investigated whether the popularity algorithm (an others) could influence the movement of the agents and whether this process would influence the development of their repertoires. In this case, if a listening agent "recognises" the rhythm played by the other agent, then it will follow the player agent in the space in the next iteration of the algorithm.

Figure 12 shows periodic clustering of one or more groups of agents that move together and keep interacting until the cluster is scattered due to an unsuccessful interaction.

In Figure 13, we can observe two behaviours that are typical of the popularity algorithm with movement taken into account. The first being that there are many more

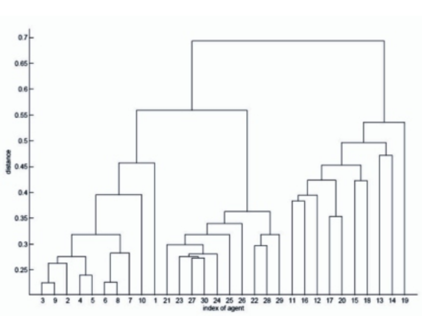


Figure 11. Dendrogram resulting from the hierarchical cluster analysis conducted in the context of the popularity algorithm containing three independent societies with 10 agents each

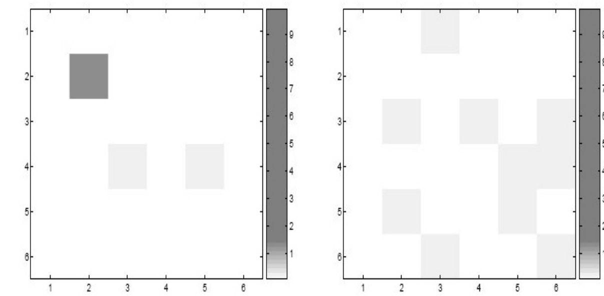


Figure 12. Clustering can take place (figure on the left) followed by scattering at a later stage (figure on the right). A cluster is indicated by a darker colour

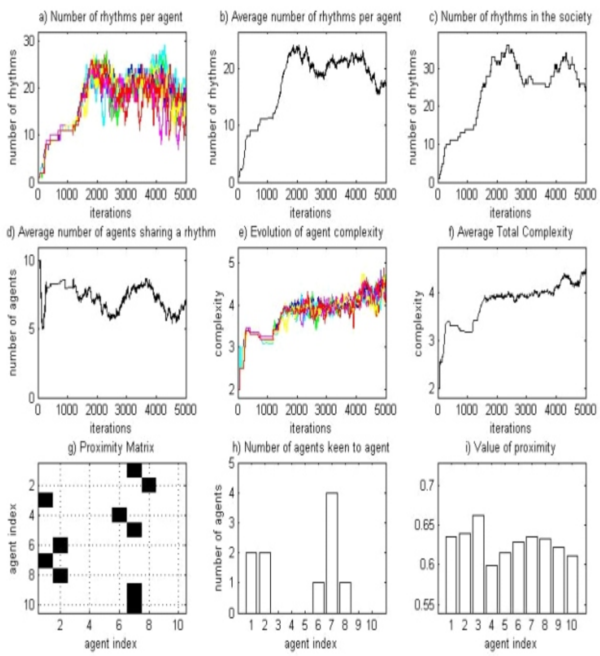
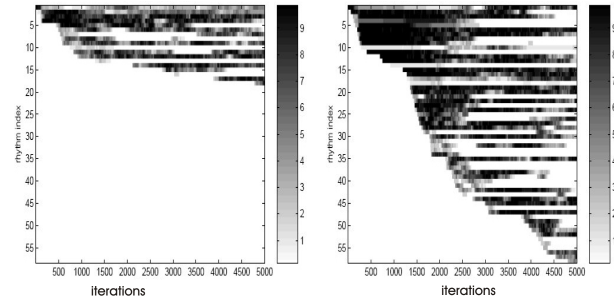


Figure 13. Results from a typical run of the popularity algorithm taking into account the movement of the agents as an influencing factor in the evolution of the repertoire

rhythms affecting the interactions than in the case without movement; this is due to the fact that every time a positive feedback occurs, two or more agents will form a group. This increases the number of interactions between them and consequently the number of rhythms in their repertoires. The second being that there is an initial overshoot of the size of the repertoire before reaching a level of stability. This is possibly caused by the initial clustering of agents when individual repertoires grow consistently among very closely related agents.

Figure 14 shows the lifetime of sequences that emerged during the typical runs of the popularity algorithm.



a) Popularity algorithm

b) Popularity algorithm with movement influencing the repertoire

Figure 14. The lifetime of all rhythms that emerged using the popularity algorithm in cases where: a) movement does not influence the developments, and b) movement influences the developments. The number of agents that share a particular rhythm is represented by tones of grey; that is, the darker the colour the higher the number of agents sharing a particular rhythm

2.2 The Transformation Algorithm

A diagram summarising the transformation algorithm is given in Figure 15. As its name suggests, the transformation algorithm applies transformations on a rhythm whenever it is communicated between agents. The motivation behind this algorithm is to foster novelty in the repertoires of the agents. The transformation algorithm allowed for experiments aimed at assessing the degree to which the transformations occurring during the interactions have an impact on the organisation of the emerging repertoire as time progresses.

It is possible to observe in Figure 16 that due to the rise of the amount of transformations, the repertoires are much larger than in the popularity algorithm. For instance, compare Figure 10b with Figure 16b. Figure 16f shows the development of the average complexity of the society, where we can observe two clearly differentiated growing rates before and after 200 iterations. When this algorithm is run with 50 agents we can also observe similar growing rates, although the initial rate is not as steep as it is with 10 agents, and the transition is smoothed (Figure 17).

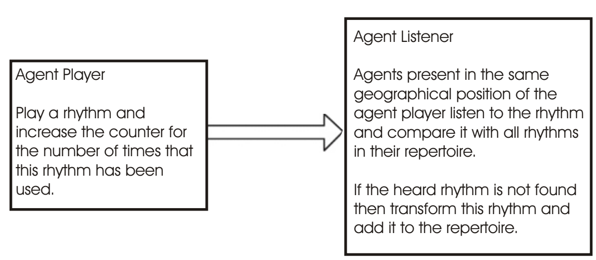


Figure 15. The transformation algorithm

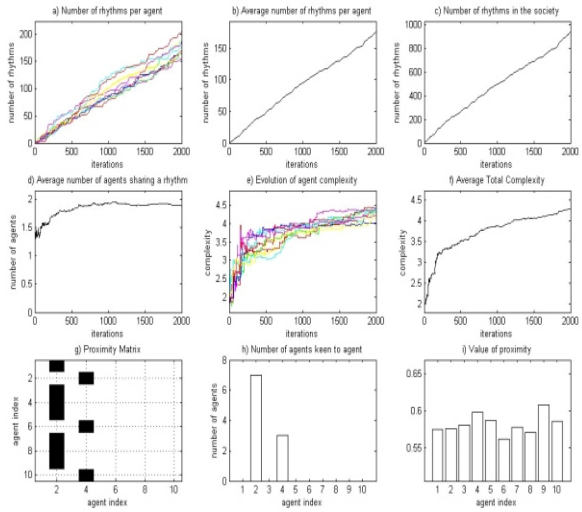


Figure 16. Results from a typical run of the transformation algorithm with 10 agents

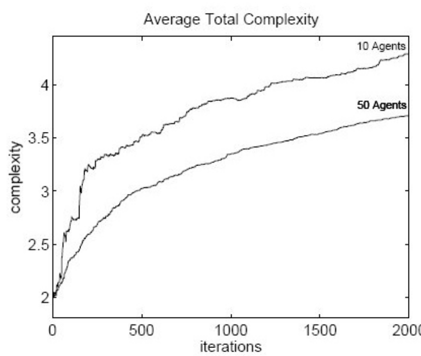


Figure 17. Average complexity evolution curves resulting from the transformation algorithm with 10 and 50 agents

2.3 The Complexity Algorithm

A diagram summarising the complexity algorithm is given in Figure 18. With the complexity algorithm one can study the effect of preference for particular types of rhythms. For instance, we can observe whether the agents would show preference for rhythms with identical complexity.

With the complexity algorithm, the agents include in their repertoire only those listened rhythms that fall within a window of complexity centred in the average complexity of the rhythms of the listening agent. That is, all listened rhythms that fall within the interval of [*AvComplexity* - 1, *AvComplexity*+1] will be included in the repertoire of the agent.

Figure 19 displays the results from a run of the complexity algorithm with the same parameters as for the run of the

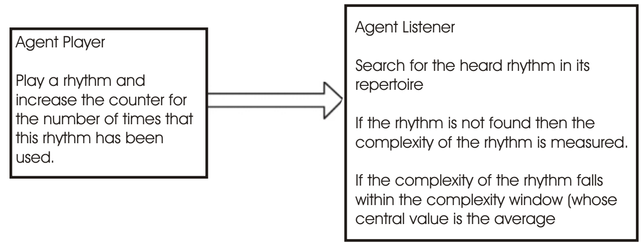


Figure 18. The complexity algorithm

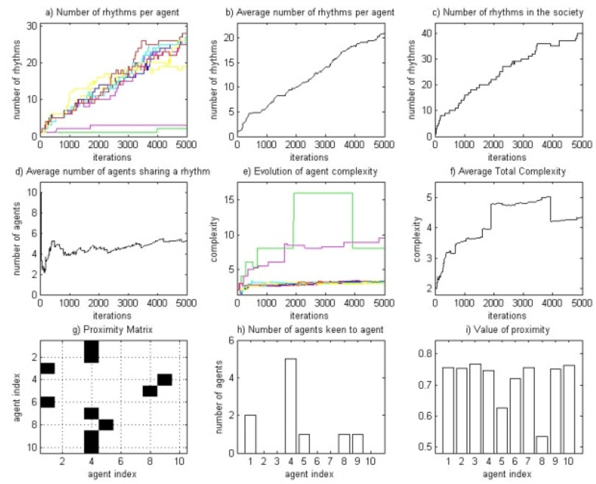


Figure 19. Results from a typical run of the complexity algorithm with 10 agents

popularity algorithm shown in Figure 10. The most interesting emergent behaviour that can be observed from these graphs is the distinct repertoires developed by the agents 5 and 8; they are distinct in terms of the complexity (represented by two distinct plots in Figure 19e) and the number of developed rhythms (represented by distinct plots at the bottom on the graph in Figure 19a). Although they are considered to have the smaller values of proximity in relation to the closer agent (Figure 19i), their development seems to be tightly connected. It is seen here that initial small changes in complexity due to transformations can actually result in completely different developments between the agents.

The cluster tree for the results shown in Figure 19 is given in Figure 20. Two main clusters appear, separated by a value of *DistRep* = 0.8. Furthermore, the two agents that at an early stage of the simulation were able to perform transformations leading to sequences of higher complexity remain more apart than the agents of the other cluster.

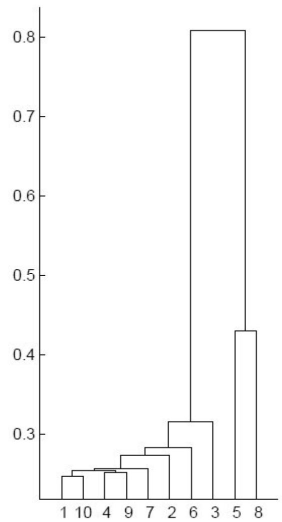


Figure 20. Dendrogram resulting from the hierarchical cluster analysis conducted in the context of the complexity algorithm with 10 agents

Conclusion

Most current approaches using A-Life in software for generating music entail the application of a GA. It was suggested that a strictly GA-based approach to generate music is questionable because they were not designed to address musical problems in the first place, but Engineering problems. The act of composing music seldom involves an automated selective procedure towards an ideal outcome, based on a set of definite fitness criteria.

As a way forward, the authors suggested that A-Life-based systems for generating music should employ algorithms that consider music as a cultural phenomenon whereby social pressure plays an important role in the development of musical conventions. To this end, three algorithms inspired by A-Life were proposed: the popularity, transformation and complexity algorithms, respectively. In addition, a number of methods were developed to monitor the behaviour of the algorithms.

In all runs of the three algorithms, the emergence of coherent repertoires were observed across the agents in the society. Clustering of agents according to their repertoires could also be observed on various occasions.

Whereas the size of the repertoire is controlled by a popularity parameter in the first algorithm, it tends to grow constantly in the other two algorithms. It is understood that this behaviour would change if the lifetime of the agents is limited, which would imply some form of population renewal. This might increase the role of the memory loss mechanism and therefore constrain the growth of the repertoire. Also a small subset of agents that tend to concentrate the preference of most of the population was observed. This trend tended to appear in many runs with different settings, in all the three algorithms. In the third algorithm we observed that large clusters of agents tended to appear, grouped according to the complexity and average number of rhythms.

As mentioned earlier, a system for the composition of rhythms was implemented using the three algorithms introduced in this paper. The analysis methods shown above offer the ability to extract information about the behaviour of the agents and the evolving rhythms in many different ways, providing composers the means to explore the outcomes systematically. However, an in-depth discussion on how the system is used artistically to compose pieces of music falls beyond the scope of this paper, and shall be reported in the future.

Acknowledgement

This project was developed thanks to a grant from The Leverhulme Trust. The authors are most grateful for this support.

References

[1]. Beek, P. J., Peper, C. E., and Daffertshofer, A. (2000). "Timekeepers versus nonlinear oscillators: how the approaches differ". In P. Desain and L. Windsor, Eds., *Rhythm Perception and Production*, pp. 9-34. London: Swets and Zeitlinger (now Taylor & Francis).

[2]. Biles, J. A. (1994). "Genjam: A genetic algorithm for generating jazz solos". *Proceedings of the International Computer Music Conference*, Aarhus(Denmark). San Francisco, USA: International Computer Music Association.

[3]. De Boer, B. (1999). *Self-Organisation in Vowel Systems*. PhD thesis, Vrije Universiteit Brussel.

[4]. Doornbusch, P. (2005). *The Music of the CSIRAC: Australia's First Computer Music*. Victoria, Australia:

Common Ground.

[5]. **Drake, C. and Bertrand, D. (2001).** "The Quest for Universals in Temporal Processing in Music". *Annals of the New York Academy of Sciences*, 930(1):17-27.

[6]. **Handel, S. (1989).** *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, USA: The MIT Press.

[7]. **Kelso, J. A. (1984).** "Phase transitions and critical behavior in human bimanual coordination". *American Journal of Physiology Regulatory, Integrative and Comparative Physiolpgy*, 246(6):1000-1004.

[8]. **Kim, K.J. and Cho, S.B. (2006).** "A comprehensive overview of the applications of artificial life". *Artificial Life*, 12(1):153-182.

[9]. **Kirby, S. (2002).** "Natural language from artificial life". *Artificial Life*, 8(2):185-215.

[10]. **Mandelbrot, B. B. (1982).** *The fractal geometry of nature*. New York, USA: W.H. Freeman and Co.

[11]. **Martins, J. M., Gimenes, M., Manzolli, J., and Maia Jr., A. (2005).** "Similarity measures for rhythmic sequences". *Proceedings of the 10th Brazilian Symposium on Computer Music (SBCM)*, Belo Horizonte (Brazil).

[12]. **Miranda, E. and Biles, A., Eds. (2007).** *Evolutionary Computer Music*. London, UK: Springer-Verlag.

[13]. **Miranda, E. R. (2004).** "At the Crossroads of Evolutionary Computation and Music: Self-Programming Synthesizers, Swarm Orchestras and the Origins of Melody", *Evolutionary Computation* 12(2):137-158.

[14]. **Miranda, E. R. (2002).** "Mimetic development of intonation", *Proceedings of the 2nd International Conference on Music and Artificial Intelligence (ICMAI 2002)*. Springer Verlag - Lecture Notes on Artificial Intelligence.

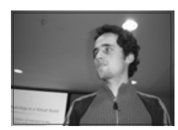
[15]. **Steels, L. (1995).** "A self-organizing spatial vocabulary". *Artificial Life*, 2(3):319-332.

[16]. **Vogt, P. (2000).** *Lexicon Grounding on Mobile Robots*. PhD thesis, Vrije Universiteit Brussel.

ABOUT THE AUTHORS

** Ph.D Research Student, Interdisciplinary Centre for Computer Music Research (ICCMR), University of Plymouth, Plymouth, UK.*

*** Professor in Computer Music, Interdisciplinary Centre for Computer Music Research (ICCMR), University of Plymouth, Plymouth UK.*



Joao M. Martins received his BEng in Electrical Engineering from the University of Coimbra, Portugal. He is currently pursuing his Ph.D on the topic of Computer Music at the Interdisciplinary Centre for Computer Music Research (ICCMR) under the guidance of Prof. Miranda. His research is concerned with Artificial Life modeling of music evolution. He was one of the organizers of the Music-AL workshop on music and Artificial Life, as part of the European Conference on Artificial Life 2007.

Prof. Eduardo R Miranda received his M.Sc in Music Technology from the University of York, UK and Ph.D in Music from the University of Edinburgh, UK. He served as a Research Scientist at SONY Computer Science Laboratory in Paris, France, before moving to the University of Plymouth in 2003 to establish the Interdisciplinary Centre for Computer Music Research (ICCMR) in the Faculty of Technology. He is currently the chair of the School of Computing, Communications and Electronics Research Committee and is the head of ICCMR.

Emergent Rhythmic Phrases in an A-Life Environment

João M. Martins and Eduardo R. Miranda

Interdisciplinary Centre for Computer Music Research
University of Plymouth, Plymouth, Devon PL4 8AA
United Kingdom.
{joao.martins,eduardo.miranda}@plymouth.ac.uk

Abstract. The A-Life approach to Music is a promising new development. The vast majority of existing A-Life systems for musical composition employ a Genetic Algorithm (GA) to produce musical melodies, rhythms, and so on. In these systems, music parameters are represented as genotypes and GA operators are applied on these representations to produce music according to given fitness criteria. We have identified two methodological limitations of such GA-based systems: one relates to the fact that composition should not be driven by a constant set of fitness criteria and the other is to do with the fact that music is largely a cultural phenomenon driven by social pressure and this is cumbersome to model with standard GA alone. An approach improve this scenario is to build systems with A-Life algorithms designed primarily to address musical issues, rather than using algorithms that were not designed for music in the first place. The work presented in this paper contributes to this line of thought by proposing the design of algorithms that consider music as a cultural phenomenon whereby social pressure plays an important role in the development of musical conventions. We introduce three algorithms: popularity, transformation and complexity algorithms, respectively. The algorithms were implemented in the context of a system for composition of rhythms, where the user can explore their potential to generate rhythmic sequences and also monitor their behavior. Finally, we explore the composition capabilities of the system by using the rhythms developed by the agents during the simulations in a collective performance environment. This bottom-up approach automatically defines an implicit metric structure.

Keywords
Music and A-Life, emergence of rhythms, interactive autonomous agents, social pressure in cultural evolution

1 Introduction

A comprehensive overview of applications of A-Life published recently mentioned that A-Life has been applied in the field of Music and indicated a few examples of A-Life systems for musical composition [1]. Why should musicians be interested in A-Life?

From the discovery almost three thousand years ago of the direct relationship between the pitch of a note and the length of a string or pipe, to the latest computer models of human musical cognition and intelligence, musicians have always looked at science to provide new and challenging paradigms to study and compose music.

The A-Life approach to Music is a promising new development for composers and musicologists alike. For composers, A-Life provides an innovative and natural means for generating musical ideas from a specifiable set of primitive components and processes reflecting the compositional process of generating a variety of ideas by brainstorming followed by selecting the most promising ones for further iterated refinement. For musicologists, A-Life techniques might be used to model the cultural transmission and change of a population's body of musical ideas over time; e.g., to model the development and maintenance of musical styles within particular cultural contexts and their reorganization and adaptation in response to cultural exchange. In both cases, the musical evolution can be influenced by a variety of constraints and tendencies built into the system, such as realistic cognitive and environmental factors that might influence the way in which music is experienced, learned, stored, modified, and passed on between individuals.

The vast majority of existing A-Life systems for musical composition employ a standard Genetic Algorithm (GA) to produce musical melodies, rhythms, and so on. Normally, music parameters are represented in these systems as genotypes and GA operators are applied on these representations to produce music according to given fitness criteria; for a review, please refer to [2].

Because of the highly symbolic nature of Western music notation, music parameters are suitable for GA-based processing and a number of composers, including ourselves, have used such systems to compose music. However, we have identified two methodological limitations of such GA-based systems which may jeopardise further developments in this area: a) fitness criteria are not easy to define when dealing with musical composition and b) music is largely a cultural phenomenon driven by social pressure and this is cumbersome to model with standard GA alone.

The first limitation emerges from the fact that music is not an exact science. For example, it differs from engineering. Whereas the success of a piece of engineering would normally be measured by its ability to match a number of functional requirements effectively, the success of a piece of music cannot be measured so objectively. Indeed, whereas good engineers are praised for following the rules of their métier strictly, good composers (at least in the Western music tradition) are praised for clever violations of musical conventions. Moreover, in most cases, composers do not explicitly know a priory how a new piece of music will sound like until it is completed and indeed performed. Therefore, rather than tools to generate efficient solutions to problems automatically, composers need tools to explore a vast space of possible outcomes.

Biles [3] proposed an interesting approach to implement GA-based systems for the exploration of a space of musical possibilities, which takes into consideration the evaluation of the user; that is, the fitness of each generation is evaluated by the user. This is surely a very interesting idea, but this slows down the compositional process enormously. Biles is aware of this problem, which he refers to as the “fitness bottleneck” problem.

The second limitation is largely related to a problem that is endemic in the field of Computer Music, which is the tendency to design systems to generate music from algorithms that were not designed for music in the first instance. For example, in the late 1980s it became fashionable to implement systems that generated music from fractals [4], but such systems seldom produced significant pieces of music. There was a tendency at the time to overstate the adequacy of fractals for algorithmic composition. In reality, fractals are not appropriate to convey musical information, but appealing images: the eye can grasp an entire image at a fraction of the time needed to grasp even a short sound sequence.

Nowadays, we may be witnessing a similar case of overstatement on the adequacy of GA for algorithmic composition. Although we acknowledge that there have been rather successful stories (e.g. [3, 5]), we believe that additional evolutionary computation methods need to be developed in order to move the field of evolutionary computer music forward.

One way forward is to build systems with A-Life algorithms designed or suitably modified to address musical issues. A-Life methods have been previously used for music composition [6, 7] or to study the evolution of bird songs [8–10].

The work presented in this paper contributes to this line of thought by looking into the design of algorithms that consider music as a cultural phenomenon whereby social pressure plays an important role in the development of musical conventions. A plausible method to embed social dynamics in such algorithms is to design them within the framework of interacting autonomous agents.

In this paper we introduce three algorithms, referred to as popularity, transformation and complexity algorithms, respectively. These algorithms were implemented in the context of a system for composition of rhythms. In this system the user can explore the potential of these algorithms to generate rhythmic sequences and also monitor the behavior of the system. The system offers the ability to extract information about its behavior in many different ways, providing composers the means to explore the outcomes systematically.

Our research is greatly inspired by the work developed by research into gaining a better understanding of the evolution of language with computational models [11–15], particularly the work of Steels [11] on language imitation games with software agents and robots. Basically an imitation game consists of one agent picking a random sound from its repertoire and the other agent trying to imitate it. Then a feedback is given about the success of the imitation. On the basis of this feedback, the agents update their memories.

By way of related research, we cite the work by de Boer [12] on modeling the emergence of vowel systems by means of imitations games. Also, Miranda [16] has developed a variant of de Boer’s games in order to model the emergence of intonation systems.

In a previous paper [17] we provided a detailed explanation on the algorithms of interaction that enable repertoires of rhythms to develop. We have also studied the development of complexity of the repertoires and similarity between agents as a result of their behaviours. In this paper we intend to go a step further placing the agents in a group performance, letting the structure of the rhythmic phrases be defined collectively.

2 The Agents

The agents are identical to each other and the number of agents in a group may vary. The agents move in a virtual 2D space and they normally interact in pairs (Figure 1). Essentially, the agents interact by playing rhythmic sequences to each other, with the objective of developing repertoires of rhythms collectively. At each round, each of the agents in a pair plays one of two different roles: the player and the listener. At each interaction, the agents may perform operations on the rhythms that they play to each other, depending on the interaction algorithm and on the status of the emerging repertoire. The agents are provided with a memory to store the emerging rhythms and other associated information.

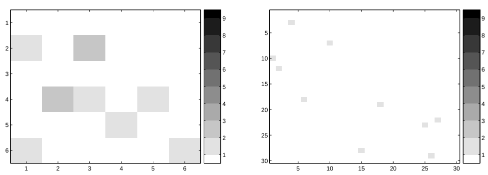


Fig. 1. 2D virtual worlds with different sizes holding 10 agents. A darker color indicates a cluster. (This will be clarified in due course.)

One interesting ability of human beings is that we are able to perceive, and more importantly, to produce an isochronous pulse [18]. Moreover, humans show a preference for rhythms composed of integer ratios of the the basic isochronous pulse [19]. Therefore, rhythms are represented here as interonset intervals in terms of small integer ratios of an isochronous pulse (Fig. 2).

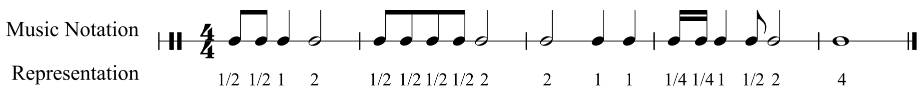


Fig. 2. Standard music notation of a rhythmic sequence and its corresponding interonset representation.

2.1 Transformations of Rhythms

At the core of the mechanism by which the agents develop rhythmic sequences is a set of basic transformation operations. These operations enable the agents to generate new rhythmic sequences and change the rhythmic sequences that they learn as the result of the interactions with other agents. The transformation operations are as follows:

– Divide a rhythmic figure by two (see Fig. 3a)
– Merge two rhythmic figures (see Fig. 3b)
– Add one element to the sequence (see Fig. 3c)
– Remove one element from the sequence (see Fig. 3d)

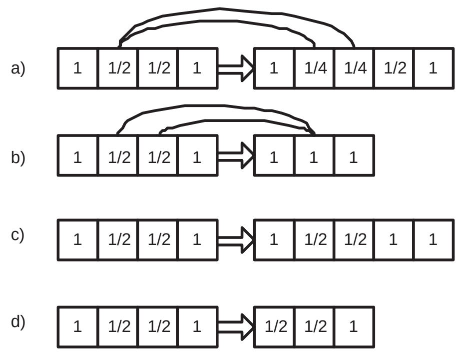


Fig. 3. Examples of rhythmic transformations.

The definition of these transformations were inspired by the dynamical systems approach to study human bimanual coordination [20] and is based on the notion that two coupled oscillators will converge to stability points at frequencies related by integer ratios [21]. Furthermore, common music notation facilitates these types of transformations. We have defined other transformations that divide a figure into three, five, and other prime numbers, but the impact of these additional transformations on the model is beyond the scope of this paper. Addition and removal transformations were introduced to increase diversity in the pool of rhythms and to produce rhythms of different lengths.

3 The Interaction Algorithms and the Experiments

The interaction algorithms and the analysis methods that we have implemented in our system are introduced below. Each algorithm is introduced in the context of illustrative experiments aimed at studying the development of repertoires of rhythmic sequences from three different perspectives:

- From the perspective of an individual agent
- From the perspective of a group of agents, referred to as the society
- From the perspective of the developed rhythms

From the perspective of an individual agent, we studied the development of the size and the complexity of the repertoire of individual agents. From the perspective of the society we averaged values of the corresponding individual measures from the agents, as well as similarity between agents and how they were clustered in terms of the rhythms that they shared. Finally, from the perspective of the developed rhythms, we measured their lifetime, the amount of rhythmic sequences that the society developed and the degree to which the agents shared similar rhythms. We traced the lifetime of a rhythmic sequence by counting the number of agents that possessed this sequence at each iteration. Fig. 4 shows graphs illustrating these various types of analyses.

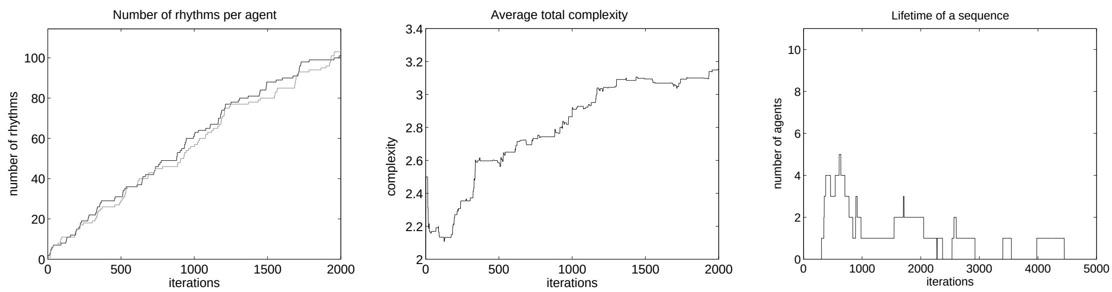


Fig. 4. a) Development of the size of the repertoire for different agents.; b) Complexity of the rhythms of the whole society; c)Number of agents sharing a particular rhythm.

The experiments were run for 5000 iterations each for a number of times, with the objective of observing their behavior under different conditions. We have run experiments with societies of 3, 10 and 50 agents. On some of the experiments we limited the lifetime of the agents to 1000 iterations; when an agent dies, another is born. Sometimes the algorithm considers the movement of the agents in the 2D space, which may or may not influence the nature of the interactions.

In this paper we focus only the results of the popularity algorithm. For a detailed exposition of the results of the other two algorithms please refer to [17].

3.1 The Popularity Algorithm

Popularity is a numerical parameter that each agent attributes to a rhythm in its repertoire. The parameter is modified both by the listener and by the player during an interaction. If the listener recognises the rhythm (that is, if it holds this rhythm in its repertoire), then it will increase the popularity index of this rhythm and will give a positive feedback to the player. A positive feedback is an acknowledgment signal, which will prompt the player to increase the popularity index of this rhythm in its repertoire as well. Conversely, if the listener does not recognize the rhythm, then it will add this rhythm to its repertoire and will give a negative feedback to the player, which will cause the player to decrease the popularity index of this rhythm. Furthermore, there is a memory loss mechanism whereby after each interaction all the rhythms have their popularity index decreased by a small value of 0.05. This accounts for a natural drop in popularity due to ageing of the rhythm. The diagram of this interaction is displayed in Fig. 6a.

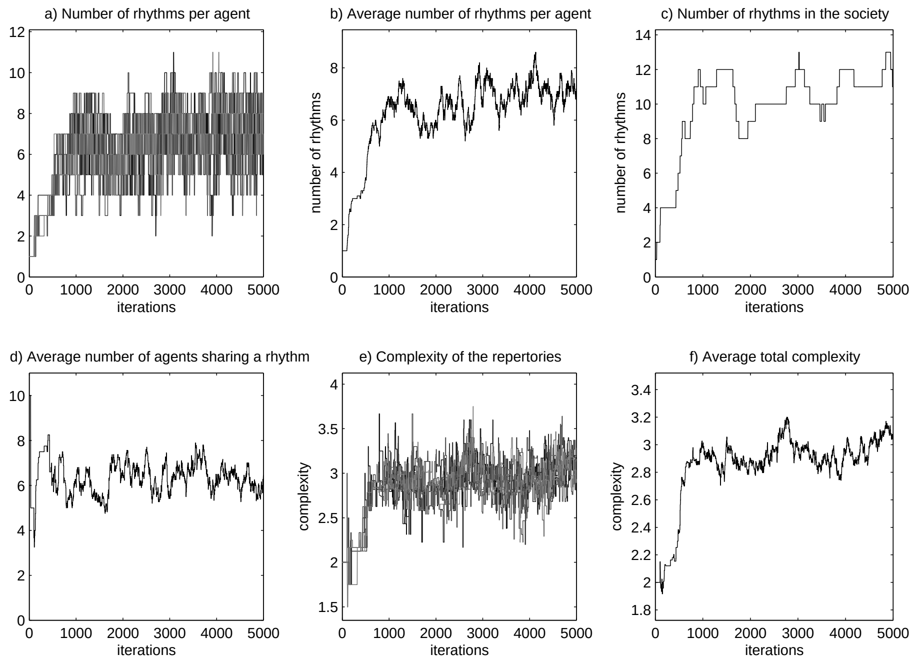


Fig. 5. Results from a typical run of the popularity algorithm with 10 agents.

Fig. 5 shows the results after 5000 iterations of the popularity algorithm without population renewal. Fig. 5a displays the development of the repertoire from the individual agents and the graph in Fig. 5b displays the corresponding average across the agents. Here the repertoire of each agent grows monotonously during 500 iterations and subsequently oscillates around a stable point.

Fig. 5c displays the development of the repertoire of the whole society being a direct consequence of the lifetime of each rhythm. The average number of agents sharing a rhythm (Fig. 5d) is calculated by summing the instant number of agents sharing a rhythm (Fig. 4c) for all rhythms, and dividing the result by the number of rhythms currently present in the society (Fig. 5c). Fig. 5d) provides the means to assess the global behavior of the society; for instance, if it develops coherently in terms of the popularity of existing rhythms.

Fig. 5e represents the development of complexity of the individual agents and Fig. 5f gives the corresponding average. Initially, the size and complexity of the repertoire of individual agents are very close to the average, but this trend is replaced quickly by repertoires of different sizes amongst the agents.

3.2 The Transformation and Complexity Algorithms

As its name suggest, the transformation algorithm (Fig. 6b) applies transformations on a rhythm whenever it is communicated between agents. The motivation behind this algorithm is to foster novelty. We conducted experiments to evaluate the degree to which transformations occurring during the interactions have an impact on the organisation of the emerging repertoire as time progresses.

The diagram of the complexity algorithm is shown in Fig. 6c. With the complexity algorithm we studied the effect of preference for particular types of rhythm; in this case, we wanted to establish whether the agents would show preference for rhythms with identical complexity; we have developed methods to measure this complexity. Here the agents include in their repertoire only those listened rhythms that fall within a window of complexity centered in the average complexity of the rhythms of the listening agent.

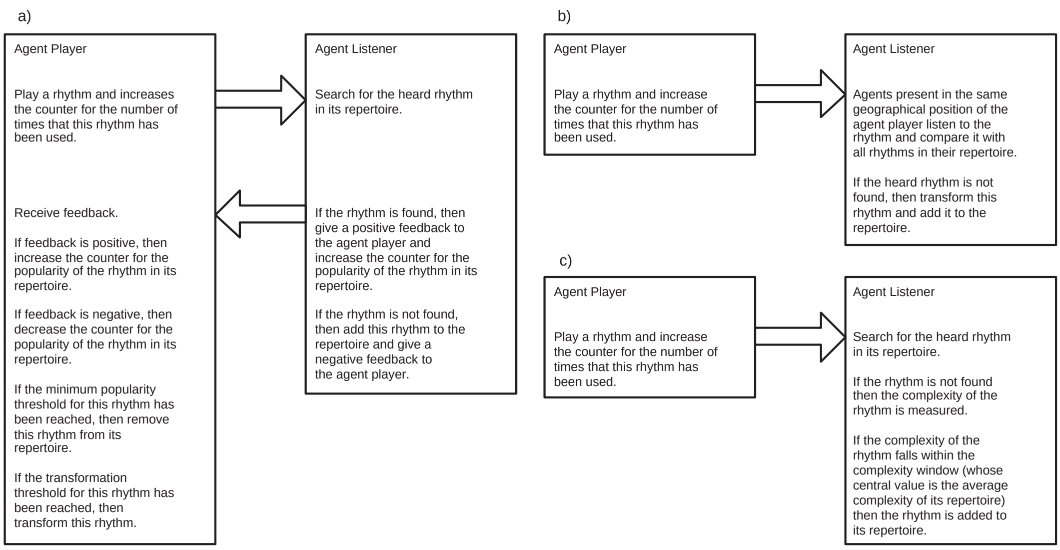


Fig. 6. a) Popularity algorithm; b) Transformation algorithm; c) Complexity algorithm.

4 Rhythmic Phrases in the Social Context

The interaction processes introduced above developed into some interesting behaviours, revealing the dynamics of the representations of the rhythmic units. From a musical point of view, however, the output of the system is not yet satisfactory. This is not surprising because our intention is that these acquired units should be considered as the basic material the agents will use when playing in a synchronised mode. In this section we will demonstrate how the agents can create a rhythmic background texture that will also establish the metric structure of a longer piece.

4.1 Emergent Phrase Length

Most often the rhythmic sections of musical pieces consist of repetitions of small rhythmic units. This fact may have the function of either reinforce or contradict a metric structure. We decided to conduct an experiment where all the agents present in a given geographic position would play their rhythms simultaneously, as opposed to the interaction algorithms presented in Sec. 3. Instead of focusing on the learning process we observed how the agents would play a collective rhythmic piece.

If each agent plays one of the rhythms from its repertoire and repeats it, there will be a strong metrical cue associated with this repetition. The rhythms that belong to the repertoires of the agents may or may not have different lengths due to the transformations (Sec.2.1). When played together there will be an instant where all the agents will hit the initial beat of their basic rhythm at the same time. The difference between two such consecutive instants defines the length of the music phrase.

When the lengths of the basic rhythms are divisible in relation to each other then the length of the longest will define the size of the phrase. In case the length values are not divisible (3:2, 4:3, 5:3,...) the repetitions will generate an interesting polyrhythmic effect.

In a polyrhythm, two or more independent rhythms are played simultaneously. Polyrhythms are particularly abundant in African music, Indian classical music, Cuban music and Jazz. For a more detailed explanation on polyrhythms please refer to Handel [22].

Algorithmically, this can be achieved by finding the least common multiple of the lengths of all the basic rhythms. As an example we let the main rhythmic phrase be composed and each agent will have an assigned rhythmic phrase to compose other rhythmic units from the repertoire. The algorithm is defined as follows:

– Select a basic rhythm from the repertoire.
– Calculate the least common multiple between the lengths of the basic rhythms of all the agents.
– Repeat the basic rhythm across the entire composition, except for its assigned phrase.

– Select from a series of rhythms contained in the repertoire to compose an individual rhythmic phrase.

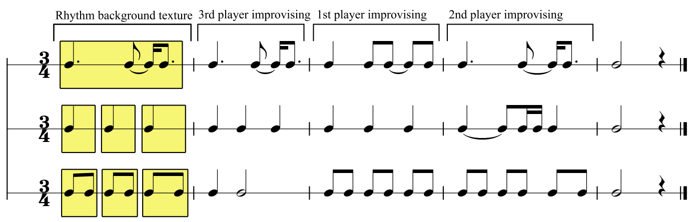


Fig. 7. Rhythmic composition resulting from the performance when three agents meet and play simultaneously.

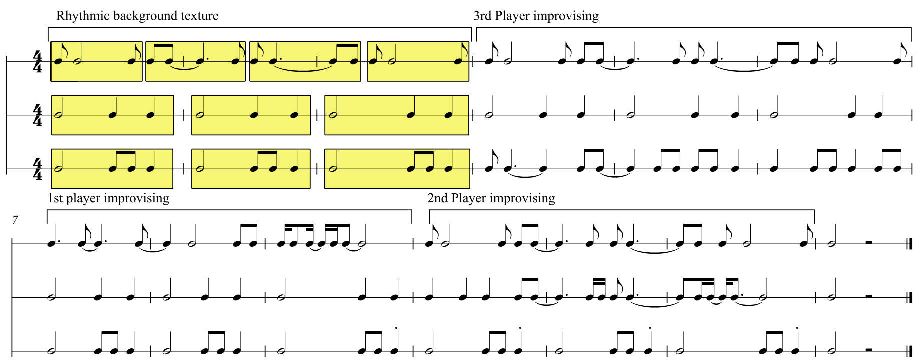


Fig. 8. Rhythmic composition where the phrase length is the result of one repetition of the rhythm from the first agent and two repetitions of the basic rhythms of the other two (polyrhythmic underlying structure).

In Figs. 7 and 8 it is possible to observe the generated score for a group of 3 agents in different stages of the experiments.

5 Conclusion

Most current approaches to musical composition with A-Life entail the application of a standard GA to produce streams of symbols representing musical parameters, such as musical notes. We suggested that one of the main limitations of this approach is that GAs are not entirely adequate for musical composition because they were not designed to address musical problems in the first place; the act of composing music seldom involves an automated selective procedure towards an ideal outcome based on a set of definite fitness criteria. As a way forward, we suggested that musical composition systems may best benefit from A-Life if the algorithms were designed to address specific musical issues. In addition to providing a more realistic music systems, such algorithms may also be useful for building models to study the evolution of music, which would follow up on the research work being conducted in the field of evolution of language [11–15].

In this paper we introduced a few algorithms, which address music as a cultural phenomenon whereby social pressure steers the development of musical conventions (in this case, repertoires of rhythmic sequences).

We also propose an algorithm that enables the agents to create longer rhythmic structures by composition of the rhythmic units that they exchange during the interactions. The algorithm suggests a bottom-up approach to rhythm structure generation. Longer phrases emerge from the usage and repetition of the rhythmic units in a collective context.

While the system is able to produce a great variety of rhythms and coherent rhythmic variations, which is what we had expected to observe in the first instance, the system also displayed a number of interesting and surprising behaviors that beg further scrutiny. We are currently studying the behaviors of these algorithms in order to ascertain whether they could be used to model the way in which rhythms emerge and develop in real societies, e.g. tribal music in Africa.

We are currently experimenting with runs involving agents with different behaviors and with agents that change their behavior during the interactions. We are also conducting experiments where the agents learn from the collective performance environment in order to observe the emergence of composition grammars and new behaviours.

Examples of the rhythms generated by the system, accompanied by a brief explanations of the behaviors that generated them are available at:

http://cmr.soc.plymouth.ac.uk/members/jmartins/research.htm

6 Acknowledgements

The authors would like to thank Simon Durrant and Steve Shaw for their valuable comments on technical aspects of the model. This work is financially supported by The Leverhulme Trust.

References

1. Kim, K.J., Cho, S.B.: A comprehensive overview of the applications of artificial life. Artificial Life **12**(1) (2006) 153–182
2. Miranda, E., Biles, A., eds.: Evolutionary Computer Music, Springer-Verlag London (2007)

3. Biles, J.A.: Genjam: A genetic algorithm for generating jazz solos. In Anderson, P., Warwick, K., eds.: Proceedings of the International Computer Music Conference, Aarhus(Denmark), International Computer Music Association (1994)
4. Mandelbrot, B.B.: The fractal geometry of nature. New York: W.H. Freeman and Co. (1982)
5. Towsey, M., Brown, A., Wright, S., Diederich, J.: Towards melodic extension using genetic algorithms. Educational Technology & Society **4**(2) (2001) 54–65
6. Brown, A.R.: Exploring rhythmic automata. In Rothlauf, F.e.a., ed.: Proceedings of the 3rd European Workshop on Evolutionary Music and Art, Lausanne(Swizerland), Springer Verlag (2005) 551–556
7. Miranda, E., Todd, P.: A-life and musical composition: A brief survey. In: Proceedings of the IX Brazilian Symposium on Computer Music, Campinas,(Brazil) (2005)
8. Todd, P.M., Werner, G.: Frankensteinian methods for evolutionary music composition. In Griffith, N., Todd, P.M., eds.: Musical networks: Parallel distributed perception and performance. MIT Press/Bradford Books, Cambridge (MA), USA (1999) 313–339
9. van den Broek, E., Todd, P.M.: Piep piep piep - ich hab' dich lieb: Rhythm as an indicator of mate quality. In: ECAL03. (2003) 425–433
10. Ritchie, G., Kirby, S.: A possible role for selective masking in the evolution of complex, learned communication systems. In Lyon, C., N.C., Cangelosi, A., eds.: Emergence of Communication and Language, Springer (2007) 387–401
11. Steels, L.: A self-organizing spatial vocabulary. Artificial Life **2**(3) (1995) 319–332
12. de Boer, B.: Self-Organisation in Vowel Systems. PhD thesis, Vrije Universiteit Brussel AI-lab (1999)
13. Kirby, S.: Natural language from artificial life. Artificial Life **8**(2) (2002) 185–215
14. Cangelosi, A., Parisi, D., eds.: Simulating the Evolution of Language. Springer Verlag, London (2002)
15. Vogt, P.: Lexicon Grounding on Mobile Robots. PhD thesis, Vrije Universiteit Brussel (2000)
16. Miranda, E.R.: Mimetic development of intonation. In: Proceedings of the 2nd International Conference on Music and Artificial Intelligence (ICMAI 2002), Springer Verlag - Lecture Notes on Artificial Intelligence (2002)
17. Martins, J.M., Miranda, E.R.: Breeding rhythms with artificial life. (Submitted to Journal) (2007)
18. Handel, S.: Listening: An Introduction to the Perception of Auditory Events. The MIT Press (1989)
19. Drake, C., Bertrand, D.: The Quest for Universals in Temporal Processing in Music. Ann NY Acad Sci **930**(1) (20) 17–27
20. Kelso, J.A.: Phase transitions and critical behavior in human bimanual coordination. Am J Physiol Regul Integr Comp Physiol **246**(6) (1984) R1000–1004
21. Beek, P.J., Peper, C.E., Daffertshofer, A..: Timekeepers versus nonlinear oscillators: how the approaches differ. In: Rhythm Perception and Production. Swets and Zeitlinger (2000) 9–34
22. Handel, S.: Using polyrhythms to study rhythm. Music Perception **1**(4) (1984) 465–484

A Connectionist Architecture for the Evolution of Rhythms

João M. Martins and Eduardo R. Miranda

Interdisciplinary Centre for Computer Music Research
University of Plymouth, Plymouth, Devon PL4 8AA
United Kingdom.
{joao.martins,eduardo.miranda}@plymouth.ac.uk

Abstract. In this paper we propose the use of an interactive multi-agent system for the study of rhythm evolution. The aim of the model proposed here is to show to what extent new rhythms emerge from both the interaction between autonomous agents, and self-organisation of internal rhythmic representations. The agents' architecture includes connectionist models to process rhythmic information, by extracting, representing and classifying their compositional patterns. The internal models of the agents are then explained and tested. This architecture was developed to explore the evolution of rhythms in a society of virtual agents based upon imitation games, inspired by research on Language evolution.

1 Introduction

The early applications of evolutionary computation to music go back to 1991 with the works of Horner and Goldberg by applying genetic algorithms to thematic bridging [1]. Since then there have been many successful attempts to apply these techniques to music. For a discussion on the history and achievements genetic algorithms please refer to Gartland-Jones and Copley [2].

Neural Networks have also been used extensively in the context of music. There have been connectionist models for pitch perception, rhythm and metre perception, melody conduction and composition, many of them collected in Griffith and Todd's book [3].

Memetic theory, the cultural counterpart of biological evolution, was invented by Dawkins in 1979 [4], and postulates that culture is an evolutionary process evolving through the exchange, mutation and recombination of units of information that can be observed in different scales. Although the definition of a meme is still quite obscure, there have been some computational attempts to model the evolution of musical style according to this theory [5].

In the specific case of rhythm composition, we can find applications of evolutionary computation such as the Interactive Genetic Algorithm (IGA)from Horowitz [6] to breed drum measure loops, the CONGA system from Tokui and Iba [7] using genetic algorithms and genetic programming to generate rhythms which are evaluated by the user, and the creation of rhythms with cellular automata by Brown [8].

All these methods have been developed mainly with three applications in mind: Sound synthesis, composition, and musicology [9]. This paper focuses on the later; i.e., a framework for the study the evolution of music.

2 Imitation Games: Language and Music

Agent based modelling is a technique frequently seen in the A-Life context to study complex systems. The emergent behaviour of the system is observed when autonomous elements self-organise as a consequence of the interactions between each other and the environment. Regarding music, the applications of A-life models are described by Miranda and Todd [10]. The scope of the work presented on this paper considers a society of agents where rhythms are exchanged, processed and categorised with neural networks.

In the real world, there is no direct transposition of the knowledge between individuals, this meaning that it is not possible to copy all the information inside a person's brain and present it to another. In the case of language or a musical performance, this features get more accentuated as there is a strangulation in the channel and consequently in the amount of information that you are able to process. Although, is easy to exchange information in the computer without loss of data, for the purpose of simulation we need to find processing mechanisms and interaction schemes that can cope with this human limitation.

While some defend the innateness of Language and thus the role of genetic mutations in its evolution, Steels [11] defends that language corresponds to a Self-organising phenomena like the ones observed in chemical and biological processes. Furthermore language develops subject to big pressures of the environment, such as limited time for articulation of words, and acoustically adverse environments.

The same duality of opinions can arise on the musical side. The transmission media is the same as language, and music is also subject to the same kind of pressures, although not constrained to meanings and concepts. Werner and Todd [12] put emphasis on the role of mate selecting pressures for the evolution of repertoires, and the evaluation of the specimen fitness is made the according to the musical material. Miranda [13] explored the self-organising potential of agents' societies by furnishing the agents with motor and auditory skills and letting them evolve a shared repertoire of short sound sequences through imitation games .

Originally inspired by Wittgenstein [14], Luc Steels [15] proposed a model of imitation games for artificial agents. Bart de Boer [16] applied this game methodology to study the emergence of a coherent vowel system handling phono-articulatory parameters. Miranda [17] applied a slightly different version of the algorithm to develop intonations. Basically the game consists of one agent picking a random sound from its repertoire and the other agent trying to imitate it. Then feedback is given about the success of the imitation. On the basis of this feedback, the agents update their vowel repertoires.

Our approach differs from the applications previously presented in the sense that the judgement is made upon a system of internal categories of each of the agents and how the repertoire evolves in the continuous search to generate music that the other agent will recognise in his internal categories system.

In this paper we introduce the groundwork that characterises our approach; i.e., the connectionist nature of the agent's mechanism for representing rhythms.

3 Agents Architecture

We will present the architecture an agent containing two neural networks in cascade that receive a stream of rhythmic events as input and contain three output neurons that map these rhythms into a tridimensional space. For a comprehensive foundation on neural network theory please refer to Haykin's book [18].

Each agent is provided with a set of two neural networks: a SARDNET and a one layer Perceptron (Figs 2 and 5). The first one receives the stimulus sequentially from an input, encoded as a MIDI stream of rhythmic events, and generates an activation pattern corresponding to the agents perception of the type of event and its place in the sequence. The dynamics of this network is fully explained in Sec. 3.1. The pattern of activation from the Sardnet then becomes the input of the later network, the Perceptron, which generates three output values that enable the categorisation of the received sequences. The architecture and learning rules of the Perceptron are explained in Sec. 3.2.

The events are represented as vectors with three components. The first component defines the musical instrument (timbre), the second defines the loudness (velocity), and the third defines the value in milliseconds that the sound lasts (Inter-onset interval). These three dimensions correspond to human perceptual attributes with different scales in sensitivity and range. Modelling these differences in the learning algorithm was not part of the scope of this paper.

3.1 Sardnet

The SARDNET [19] is a self-organising neural network for sequence classification that was applied in phonology and recently it was also applied to simulations for evolving melodies [20]. This network is an extension of the original Self Organised Map (SOM) which is a neural network used for unsupervised learning developed by Kohonen [21]. The SOM has proven to be a powerful tool for many engineering applications and some of its variations have provided explanations for the organisation and development of the visual cortex [22].

The SOM is also called a competitive network or "winner-takes-all" net, since the node with largest input "wins" all the activation, which reflects on the possibility of updating that unit in order to become more similar to the input. The neighbouring units of the winning neuron are also updated according to a neighbourhood function that organises representations of similar stimuli in a topographically close manner.

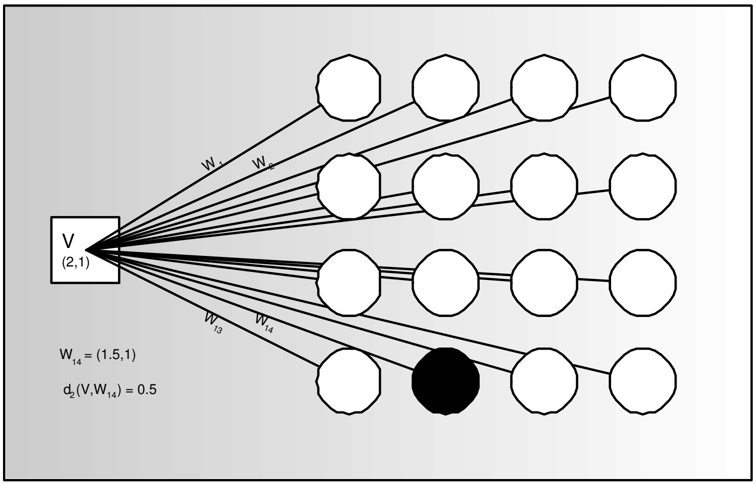


Fig. 1. Kohonen Self Organising Feature Map (SOM)

In Fig. 1 we can see a diagram of a SOM with 16 units and one input. The dimension of the input vector determines the dimension of the weights vector of each unit. To determine which weight vector is the closest one to the input unit, the euclidean distance is calculated:

$$d_2(\mathbf{v}, \mathbf{w}) = \sqrt{\sum_{i=1}^{n} |v_i - w_i|^2} \tag{1}$$

The SARDNET keeps some essential features from the SOM, but adds two important features that enables us to deal with sequences of events. The first diverging characteristic is that the winning neuron is removed from subsequent competitions, and the second difference corresponds to holding the previous activations with a decay in each time step. The dynamics of SARDNET is shown on Fig. 2 where we can observe a the stream of events passing through the input and activating three units in sequence (W_{14}, W_7, W_2). The training algorithm for the SARDNET is shown on Tab. 1.

Like the SOM, the SARDNET uses the Euclidean distance $d_2(w, v)$ from Eq. 1 to evaluate which is the weight that better matches the input. On step 3 of the algorithm the weight of the winning and the neighbourhood units are changed according to the standard rule of adaptation:

$$\Delta w_{jk} = \alpha(w_{jk,i} - v_i) \tag{2}$$

where α depends also on the distance to the winning unit, meaning its position in the neighbourhood. The neighbourhood function decreases as the map becomes more organised.

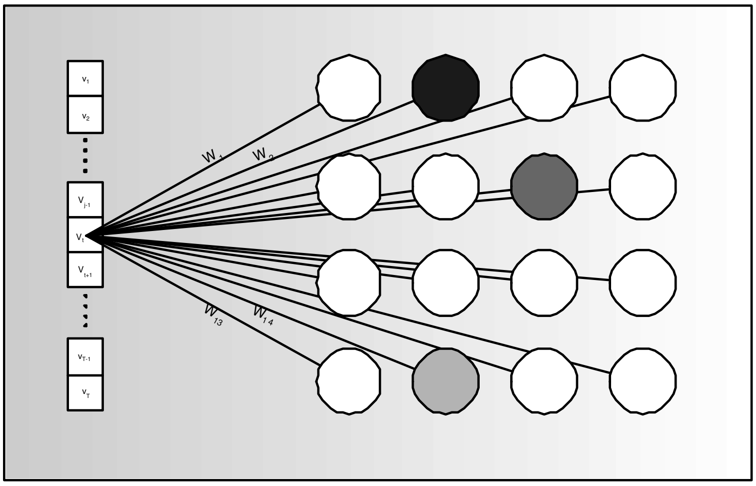


Fig. 2. SARDNET - Self-organising activation, retention and decay network

As in step 5 of the algorithm, all the active units are decayed proportionally to the decay parameter d,

$$\eta_{jk}(t+1) = d\eta_{jk}(t), \qquad 0 < d < 1 \tag{3}$$

	0	1	2	3	4
0	0.9	.	.	.	.
1	.	.	.	.	.
2	.	.	.	.	.
3	.	.	.	0.66	.
4	.	0.73	.	.	0.81

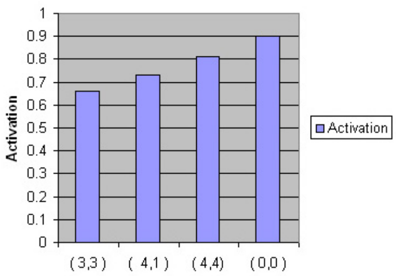


Fig. 3. Activations from the output layer on in two different views

In the following section we present the details of the Perceptron, the network that receives the activation patterns from the SARDNET, keeping the relevant information about this activation patterns across several sequences.

INITIALIZATION: Clear all map nodes to zero
MAIN LOOP: While not end of sequence 1. Find inactive weight vector that best matches the input. 2. Assign 1.0 activation to that unit. 3. Adjust weight vectors of the nodes in the neighbourhood. 4. Exclude the winning unit from subsequent competitions. 5. Decrement activation values for all other active nodes.
RESULT: Sequence representation = activated nodes ordered by activation values.

Table 1. The Sardnet training algorithm

3.2 Perceptron

The Perceptron is a neuron-like learning network developed by Rosenblatt [23] which is a one layer feed-forward neural network with a set of inputs that are fully connected to an output layer. The outputs of Perceptrons are explicit functions of the inputs. Fig. 5 shows its architecture.

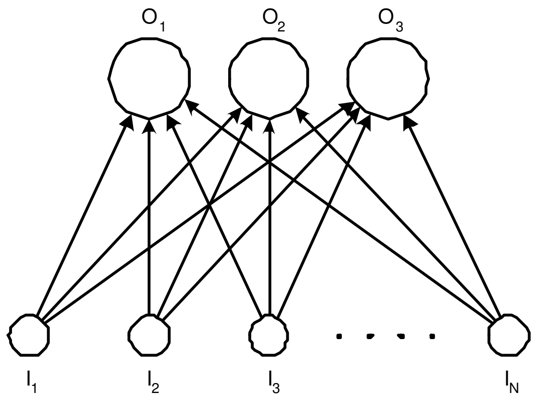


Fig. 4. Perceptron network

$$O_i = g(h_i) = g\left(\sum_k w'_{ik} I_k\right) \tag{4}$$

Eq. 4 is the propagation function of the Perceptron and $g(h)$ in Eq. 5 is the activation function computed by the units. In this case this function is a sigmoidal function,

$$O_i = g(h_i) = \frac{1}{1 + \exp(-h_i)} \tag{5}$$

The Perceptron uses the gradient descendant method to change the weights in order to adjust the test input to a given target.

$$\Delta w_{jk} = \eta * (T_k - O_k) I_j; \tag{6}$$

where η is the learning rate, T is the target value and $T_k - O_k$ is the corresponding error during the training phase.

The number of inputs of the Perceptron is the number of units of the SARDNET. The number of output neurons is arbitrarily defined as being 3 to be able to visualise the results in a tridimensional grid. This output grid enables the categorisation of the input sequences.

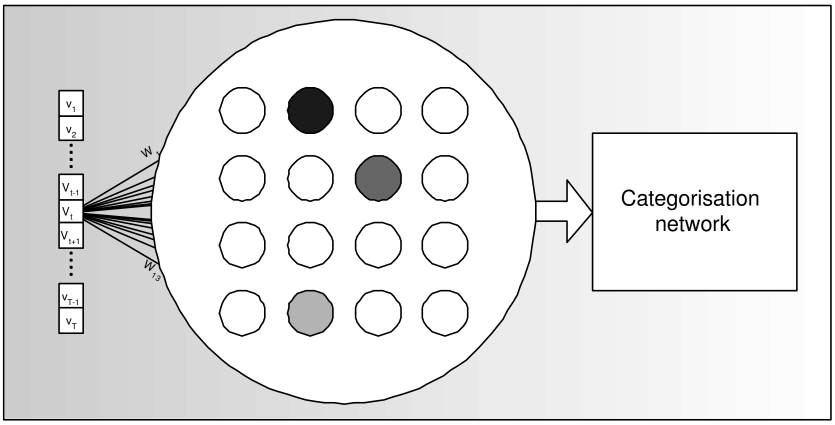


Fig. 5. Interaction diagram for the imitation game proposed

4 Analysis of the Agent

4.1 Sardnet

First we trained the Sardnet solely with prerecorded rhythms. We used a map with 50 elements, 10 in the length and 5 in the breadth, a learning rate of 0.1. The map was initialised with random weights in the range of -1 to 1. To perform the first organisation tasks the map was fed with 5 sequences of rhythms of latin music, each of them containing one or two instruments, very much like it would be if these were performed by other agents. After a couple of iterations a pattern of organisation could already be observed in the network, but the correspondent sequences extracted sounded extremely chaotic. After 50 iterations the rhythms start to sound organised as well, and the changes to the timbre of the instrument have the largest perceptual impact. This was expected to be so, as there is no discrimination in the organisation algorithm regarding the different weight components. Nevertheless, the organisation process is fine tuned enough to adapt perceptually perfectly to the incoming sequence after 80 iterations, and a learning musician is also expected to make timbre mistakes.

The graphs from Fig. 6 show the evolution of the third component of the weights (Inter-onset Intervals). The first graph shows the initial value of the weights, as explained above, the second shows the organisation process after 20 iterations, and the third shows the weights stabilised after 80 iterations. Fig. 6 d) shows the difference between the sums of the weights in two consecutive iterations, this being a measure of the stabilisation of the weights.

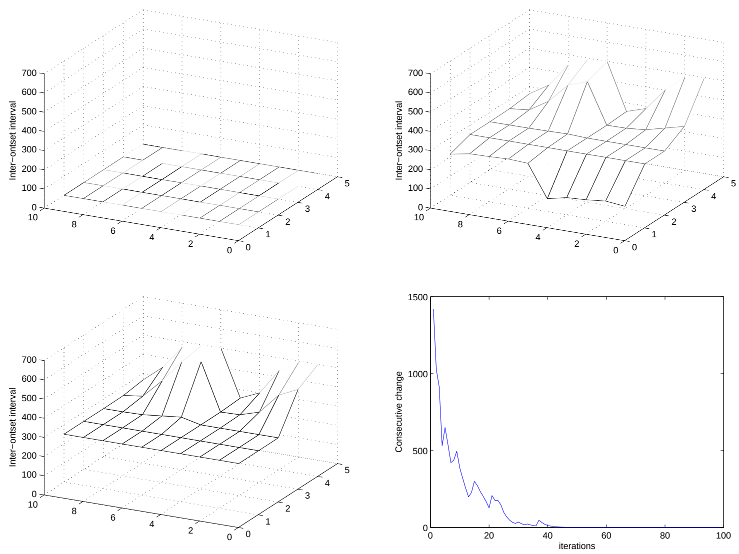


Fig. 6. Sarnet weight evolution without change of neighbourhood: a) Weight initialisation; b) After 20 iterations; c)After 80 iterations; d) Difference between the weights' sum in consecutive iterations

Previously it was stated that the SOM adapts its weights, not only for the winning elements, but also in its neighbourhood. In Fig. 7 it is shown the same organisation process but considering the neighbourhood change. The parameter σ controls the range of the the gaussian that changes the neighbourhood. By using an initial value of $\sigma = 2.97$ we can more rapidly capture the global characteristics of the input. It is necessary to reduce gradually this value in order not to destroy the representations of the events that occur less frequently. Comparing Figs. 6d) and 7d) we see that this procedure accelerates the convergence process.

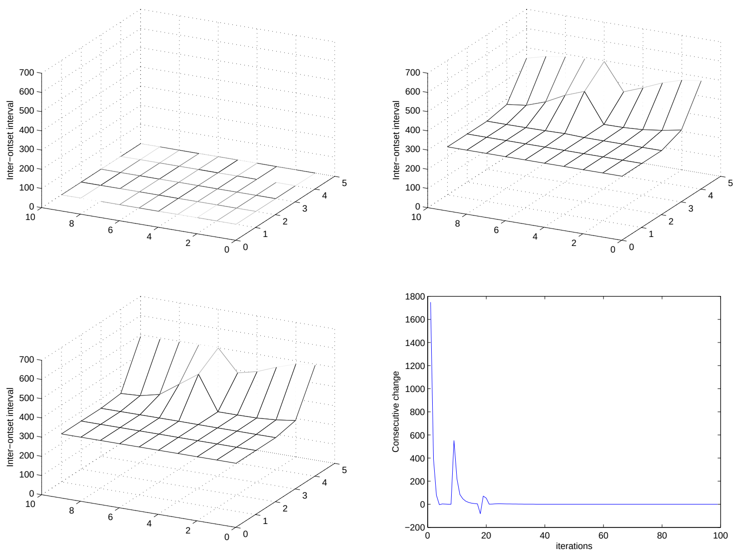


Fig. 7. Sarnet weight evolution with change of neighbourhood: a) Weight initialisation; b) After 20 iterations; c)After 80 iterations; d) Difference between the weights' sum in consecutive iterations

One of the most important conclusions is that although it is possible to extract very similar sequences from both maps, the internal representation can be quite different, as can be seen from both Figs. 6 and 7 both trained with the same sequences.

4.2 Perceptron

The Perceptron's architecture is explained in Sec. 3.2. The Perceptron used for these experiments had 50 input units, that receive their values directly from the activations of the output layer of the Sardnet. These input units are fully connected to 3 output neurons enabling the mapping and categorisation of the input sequences into a tridimensional space of straightforward visualisation. We chose the first three activation layers of 50 elements corresponding to three rhythms fed previously to the Sardnet, and trained the Perceptron to respond to these patterns with three different targets, namely $[1, 0, 0]$, $[0, 1, 0]$ $[0, 0, 1]$. This process took 434 epochs to reach an error of categorisation of 10^{-3} as can be seen in Fig. 8 a). Each training patterns is marked with an (o) in the categorisation space (Fig. 8 b)). Later, we fed the perceptron with the last two rhythms and observed its activation marked with an (x). These were found to be much closer to the $[0, 1, 0]$ target, which interestingly correspond to the most similar pattern regarding the IOIs.

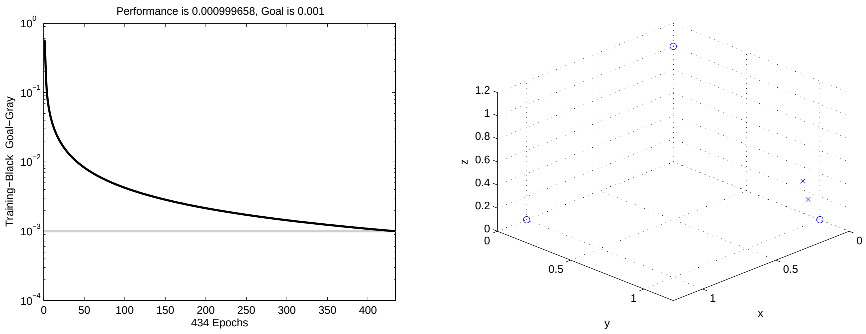


Fig. 8. a) Perceptron error in learning process; b) Categorisation space

5 Conclusion

With this paper we presented the architecture of an interactive virtual agent that is able to learn rhythms. The agent is composed of two neural networks that are able to learn the rhythms representation through self-organising processes. As it happens with humans, the agents always have different internal representations for the rhythms they listen to. Furthermore, the output of the networks categorises the incoming sequences and provides a measurement for the agents to judge how related are the listened rhythms. The rhythm representation allows for all types of rhythms to be encoded, considering event variables of Inter-onset interval, timbre and intensity. Several tests to the individual networks were made to show the potential to evolving rhythms and categories. We are now studying the results of number of simulations of imitation games where different rhythmic repertoires were evolved from scratch under a variety of different scenarios.

References

1. Horner, A., Goldberg, D.: Genetic algorithms and computer-assisted music composition. In: Proceedings of the Fourth International Conference on Genetic Algorithms, San Mateo, CA, Morgan Kauffman (1991)
2. Gratland-Jones, A., Copley, P.: The suitability of genetic algorithms for musical composition. Contemporary Music Review **22**(3) (2003) 43–55
3. Griffith, N., Todd, P.: Musical Networks. MIT-Press, Cambridge:USA (1999)
4. Blackmore, S.: The Meme Machine. Oxford University Press (1999)
5. Gimenes, M., Miranda, E.R., Johnson, C.: A memetic approach to the evolution of rhythms in a society of software agents. In: Proceedings of the 10th Brazilian Symposium of Musical Computation (SBCM), Belo Horizonte (Brazil) (2005)
6. Horowitz, D.: Generating rhythms with genetic algorithms. In Anderson, P., Warwick, K., eds.: Proceedings of the International Computer Music Conference, Aarhus(Denmark), International Computer Music Association (1994)
7. Tokui, N., Iba, H.: Music composition with interactive evolutionary computation. In: Proc. 3rd International Conf. on Generative Art, Milan, Italy (2000)

8. Brown, A.R.: Exploring rhythmic automata. In Rothlauf, F., et al., eds.: Proceedings of the 3rd European Workshop on Evolutionary Music and Art, Lausanne(Swizerland), Springer Verlag (2005)
9. Coutinho, E., Gimenes, M., Martins, J., Miranda, E.R.: Computational musicology: An artificial life approach. In: Proceedings of the 2nd Portuguese Workshop on Artificial Life and Evolutionary Algorithms Workshop, Covilhã(Portugal), Springer Verlag (2005)
10. Miranda, E., Todd, P.: A-life and musical composition: A brief survey. In: Proceedings of the IX Brazilian Symposium on Computer Music, Campinas,(Brazil) (2003)
11. Steels, L.: The synthetic modeling of language origins. Evolution of Communication **1**(1) (1997) 1–34
12. Werner, G., Todd, P.: Too many love songs: Sexual selection and the evolution of communication. In Husbands, P., Harvey, I., eds.: ECAL97, Cambridge, MA, MIT Press (1997) 434–443
13. Miranda, E.R.: Emergent sound repertoires in virtual societies. Computer Music Journal (MIT Press) **26**(2) (2002) 77–90
14. Wittgenstein, L.: Philosophical Investigations. Blackwell Publishers (1979)
15. Steels, L.: A self-organizing spatial vocabulary. Artificial Life **2**(3) (1995) 319–332
16. de Boer, B.: Self-Organisation in Vowel Systems. PhD thesis, Vrije Universiteit Brussel AI-lab (1999)
17. Miranda, E.R.: Mimetic development of intonation. In: Proceedings of the 2nd International Conference on Music and Artificial Intelligence (ICMAI 2002), Springer Verlag - Lecture Notes on Artificial Intelligence (2002)
18. Haykin, S.: Neural Networks. Prentice Hall, New Jersey:USA (1999)
19. James, D.L., Miikkulainen, R.: SARDNET: a self-organizing feature map for sequences. In Tesauro, G., Touretzky, D., Leen, T., eds.: Advances in Neural Information Processing Systems 7, Cambridge, MA, USA, MIT Press (1995) 577–84
20. Bosma, M.: Musicology in a virtual world: A bottom up approach to the study of musical evolution. Master's thesis, University of Groningen (2005)
21. Kohonen, T.: Self-Organizing Maps. Springer Series in Information Sciences. Springer-Verlag Berlin and Heidelberg GmbH & Co. K (1997)
22. Bednar, J.A., Miikkulainen, R.: Joint maps for orientation, eye, and direction preference in a self-organizing model of V1. Neurocomputing (in press) (2006)
23. Rosenblatt, F.: Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. Spartan Books, Washington (1962)

Similarity Measures for Rhythmic Sequences

João M. Martins [1] , Marcelo Gimenes [1] , Jônatas Manzolli [2] , Adolfo Maia Jr. [2]

[1]Computer Music Research, Faculty of Technology
University of Plymouth, Plymouth, Devon PL4 8AA
United Kingdom.

[2]Departamento de Matemática Aplicada, IMECC and
Núcleo Interdisciplinar de Comunicação Sonora (NICS), UNICAMP,
13.081-970, Campinas, SP, Brazil.

`joao.martins@plymouth.ac.uk, marcelo.gimenes@plymouth.ac.uk`

`jonatas@nics.unicamp.br, adolfo@nics.unicamp.br`

Abstract. *This paper presents a new model for measuring similarity in a general Rhythm Space. Similarity is measured by establishing a comparison between subsequences of a given rhythm. We introduce the hierarchical subdivision of rhythm sequences in several levels, and compute a Distance Matrix for each level using the "block distance". The information about the similarity of the rhythmic substructures is retrieved from the matrices and coded into a Similarity Coefficient Vector (SCV). We also present possibilities for the reduction to single values of similarity derived from the SCV. In addition, two applications of the formal model are presented, showing the potential for development using this approach.*

1. Introduction

"To study rhythm is to study all of music. Rhythm both organises, and is itself organised by, all the elements which create and shape musical processes" and this is how Meyer and Cooper [Cooper G., 1963] emphasize the importance of rhythm in the overall structure of music. In the 20^{th} century, rhythm in music, has finally been put in the focus of the attention of composers, such as Igor Stravinsky, Olivier Messiaen, Iannis Xenakis. Since then, an increasing interest to automatically compare music styles and composing techniques. Rhythm, as the most fundamental aspect of music plays a decisive role in this task.

The rhythm organization of music, as well as speech sounds and environmental events, are highly dependent on the perception of human beings. Even when subjects are presented with equal pulses at equally spaced intervals, the pulses are perceived as being grouped in a regular metric structure [Handel, 1989]. It is argued that this implicit metrical organization improves attention and memorisation of sequential tasks.

Perception influenced a great number of researchers in music [Gabrielsson, 1973, Povel and Essens, 1985]. Cooper and Meyer [Cooper G., 1963], developed an auditory theory based on *Gestalt* theories of perception, where rhythm groups in the basic level are seen as units that are categorised according to the position of the accentuated notes. Also according to them, one strong cue in rhythm organization is the one of pattern repetition. When a rhythmic motive is repeated, the brain integrates it creating a unit that is memorised and categorised accordingly.

In our approach we provide a measure of the occurrence of these repetitive patterns in a stream of rhythmic events. We leave out of this study the implications of accentuation, melody, harmony, timbre, and articulation to the perception of rhythm, and we do so for two reasons: Firstly, we are able to extract interesting and meaningful information solely from the position were the events take place, and secondly, we can find repertoire for percussion that does not contemplate any of the former musical characteristics apart from accentuation. We strongly believe, though, that our measure can be extended to incorporate some of these characteristics. Furthermore, we can have rhythms which do not obey the marks of bars or any metrical structure. This enables to compare and distinguish rhythmic sequences with different subdivisions, and possibly to provide some insight on situations that metric is difficult to extract.

Computers find it simple to discriminate if something is equal or different, but the problem rises when there is the need to evaluate if something is similar [Minsky, 1988]. The necessity of similarity measures concerns many areas of music research, specially music information retrieval systems [Hewlett and Selfridge-Field, 2005], automatic rhythm transcription of human-performed music to MIDI protocol [Takeda et al., 2003], evaluation of copyright issues, and evolutionary music [Miranda, 2004].

On the side of the abstract models, interesting results were achieved using the Levenshtein distance, also called edit distance. This a popular method for measuring similarity between strings of text of arbitrary length. This method counts the number of insertions, deletions and substitutions necessary to change one string into another other, being this number the measure of similarity between the sequences. Orpen and Huron have applied this distance to measure melodic, rhythmic and harmonic similarity in Bach chorales [Orpen and Huron, 1992]. Mongueau and Sankoff provided a method which can be seen as an extension of the previous [Mongeau and Sankoff, 1990]. Instead of considering that each transformation to the sequence contributes with the value of one to the distance, each transformation contributes with a weighted value sensitive to the kind of musical differences who are to be measured.

In this work we are most interested in constructing an abstract and formal model which can be able to compare rhythm patterns in the most general way, capturing information in several layers of detail. In addition we intend our model to be able to manipulate rhythm sequences in order to create new ones, which could be used in music composition. In the future we will extend this work by comparing it with the existing formal models and we hope to establish a closer relation between our model and perception by testing rhythmic similarity with human subjects.

In the next section we formally introduce the concepts of Rhythm Space and Similarity Measure. In section 3 we describe our algorithm implementation. In section 4 we present two applications of our model. In the last section we conclude with some comments about the model and list some interesting topics for further research.

2. Rhythm Space and Similarity Measure

In this work, rhythm sequences are thought as elements (or vectors) of a finite dimension vector space. Formally we have coded rhythms as sequences of numbers $(b_1, b_2, \ldots, b_r,)$, where the entries b_i can be any number of the set $B = \{-1, 0, 1, 2, \ldots, J\}$ which we named *Beat Set*. A positive number in a sequence indicates first that one have a beat and its magnitude indicates the level of acentuation, such as strong beat, weak beat, half strong beat, etc. The number of sequencial 0s indicates the duration of the beat. The number -1 indicates pauses or, in MIDI protocol, a note off. We associate the positive numbers to accentuation. For example, taking $J = 3$ we get the Beat Set $B = \{-1, 0, 1, 2, 3\}$, the accentuation should be read as: 1 means a *weak beat*, 2 a *half strong* and 3 a *strong beat*. Then we can construct rhythm sequences like, for example, $(3, 0, 1, 0, -1, 0, 0, 0, 2, 0, 0, 0, 3)$. Clearly we can have as many accentuation we wish, just extending the Beat Set. However we must introduce a prescription in order to avoid some ambiguities.

Rule: *If in a rhythm sequence a value -1 occurs, it can only occur again after a positive number had occurred first.*

This rule avoids ambiguities, for example, if we compare the sequences like $a = (1, -1, -1, 0, 0)$ and $b = (1, -1, 0, 0, 0)$. Since they are different one from another, the distance (see below) between them is positive. Nevertheless they represent the same events (in this case, pause) which intuitively suggests the distance must be zero. Our rule says that only the second sequence b is a valid one, that is, it is an element of our Rhythm Space *R* defined below.

Now, given a Beat Set *B*, we define its associated n-rhythm vector space $R_n(B)$ as the set of all n-vectors $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in which each entry is an element of the Beat Set *B*. On $R_n(B)$ we can define a distance. Let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ be two vectors in $R_n(B)$. The p-distance between them is defined as

$$d_p\left(\mathbf{v}, \mathbf{w}\right) = \left(\sum_{i=1}^{n} |v_i - w_i|^p\right)^{1/p}. \tag{1}$$

The value of p can be chosen according to the application or the kind of music considered. In our examples and applications we take, for the sake of simplicity, $p = 1$, the so called *block distance*.

The above p-distance is defined only for rhythm sequences which have the same length. Obviously, in most of applications, we must compare rhythm sequences of different lengths. Although it is possible to define a distance between vectors with different sizes (the so called Hausdorff Distance) we prefer to use for comparison of arbitrary rhythm sequences the concept of similarity in a particular way. So, the next logical step is to put together all the possible rhythm sequences into a same *Rhythm Space* and define a *Similarity Measure* on it. This is as follows.

Firstly, we define the **Rhythm Space**, denoted here by *R*, as the union of all $R_n(B)$, that is, $R = \bigcup_{i=0}^{\infty} R_n(B)$. We name, alternatively, the elements of *R* as *Rhythm Vectors*. Note that for each given Beat Set we have an associated Rhythm Space. We introduce similarity measure on *R* as follows. Given an arbitrary rhythm vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, we define a k-level subsequence $\mathbf{v}^{(k)}$ of $\mathbf{v}$ as any subsequence with k elements extracted from $\mathbf{v}$, preserving the original order of $\mathbf{v}$. For example, if $\mathbf{v} = (2, 0, 0, 1, -1, 0, 1, 1)$ we can extract five ordered four-levels sequences, namely, $\{(2, 0, 0, 1), (0, 0, 1, -1), (0, 1, -1, 0), (1, -1, 0, 1), (-1, 0, 1, 1)\}$. It is easy to see that a vector with n elements has $n - k + 1$ k-level subsequences. Now, given two rhythm vectors $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_m)$ in *R* consider all k-level sequences of both vectors, that is, the sets $S_{\mathbf{v}}^{(k)} = \{\mathbf{v}_i^{(k)}, i = 1, 2, \ldots, n - k + 1)\}$ and $S_{\mathbf{w}}^{(k)} = \{\mathbf{w}_j^{(k)}, j = 1, 2, \ldots, m - k + 1)\}$. If, for example, $m \leq n$ we only can consider sequences with length smaller than m, that is, we must take $1 \leq k \leq min(m, n)$.

Formally, for each k-level, we define the (i, j)-elements of the *k-level Distance Matrix* $\mathbf{D}^{(k)}$of two vectors $\mathbf{v}$ and $\mathbf{w}$ as:

$$\left[\mathbf{D}^{(k)}(\mathbf{v}, \mathbf{w})\right](i, j) = d_p(\mathbf{v}_i^{(k)}, \mathbf{w}_j^{(k)}) \tag{2}$$

where $i = 1, 2, \ldots, n-k+1$ and $j = 1, 2, \ldots, m-k+1$. Below we show a visualization of a general k-level Distance Matrix.

$$\mathbf{D}^{(k)} = \begin{bmatrix} d_p(\mathbf{v}_1^{(k)}, \mathbf{w}_1^{(k)}) & d_p(\mathbf{v}_1^{(k)}, \mathbf{w}_2^{(k)}) & \dots & d_p(\mathbf{v}_1^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \\ d_p(\mathbf{v}_2^{(k)}, \mathbf{w}_1^{(k)}) & d_p(\mathbf{v}_2^{(k)}, \mathbf{w}_2^{(k)}) & \dots & d_p(\mathbf{v}_2^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \\ \vdots & \vdots & \vdots & \vdots \\ d_p(\mathbf{v}_{(n-k+1)}^{(k)}, \mathbf{w}_1^{(k)}) & d_p(\mathbf{v}_{(n-k+1)}^{(k)}, \mathbf{w}_2^{(k)}) & \dots & d_p(\mathbf{v}_{(n-k+1)}^{(k)}, \mathbf{w}_{(m-k+1)}^{(k)}) \end{bmatrix}$$

Although there exist many different measures we restrict our analysis, as mentioned above, to the block distance ($p = 1$) and the Beat Set to $\mathbf{B} = \{0, 1\}$ (Fig. 1).

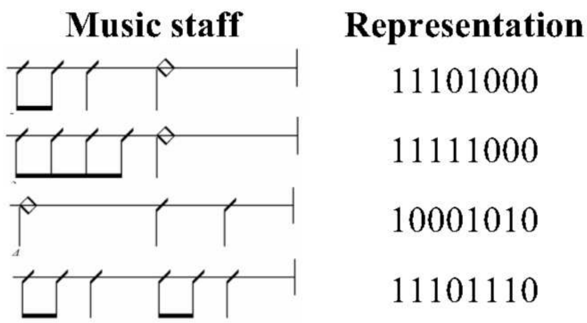


Figure 1: Musical notation and correspondent coding

We get, then, a k-level Distance Matrix whose elements are non negative integers. Now, we define the *k-level Similarity Coefficient* as the

$$c^{(k)}(\mathbf{v}, \mathbf{w}) = \frac{z(k)}{(n-k+1)(m-k+1)} \tag{3}$$

where $z(k)$ is the number of zeros in the matrix $\mathbf{D}^{(k)}$. Roughly speaking the similarity coefficient measures the *sparsity* of the matrix $\mathbf{D}^{(k)}$. Greater the coefficient $c^{(k)}$, greater is the similarity between the subsequences of level k. In the extreme case a matrix with all coefficients equal to 1, it means that one of the sequences has a perfect copy of it contained in other one.

Now we can collect all the k-levels coefficients in a vector we name *Similarity Coefficient Vector* (SCV). It reads like

$$\mathbf{C} = \left[c^{(1)}, c^{(2)}, \dots, c^{min(m,n))}\right] \tag{4}$$

In Fig. 2 we show an example of the 3-level Distance Matrix and its respective SCV. Bellow we provide an example of this approach.

Example:
Take the Beat Set as $B = \{0, 1\}$. Let $\mathbf{v} = (1, 0, 1, 1)$ and $\mathbf{w} = (1, 0, 1, 1, 0, 1)$ be two rhythm sequences in *R*. The possible 3-level sequences for $\mathbf{v}$ and $\mathbf{w}$ are:

- $S_{\mathbf{v}}^{(3)} = \{(1, 0, 1), (0, 1, 1)\}$
- $S_{\mathbf{w}}^{(3)} = \{(1, 0, 1), (0, 1, 1), (1, 1, 0), (1, 0, 1)\}$

Let us take, for the sake of simplicity, $p = 1$ on the Rhythm Space. Since we must take the distance between all elements of each level up to sixth level, one can guess that a large number of evaluations is needed. We show below only the distance between the

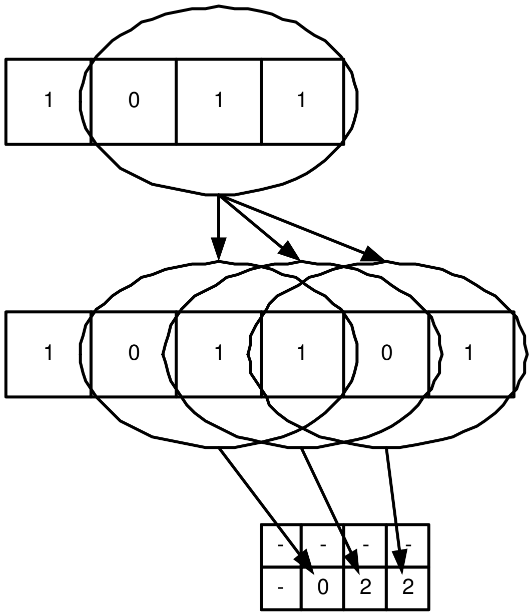


Figure 2: Building the Distances Matrix for the 3-level

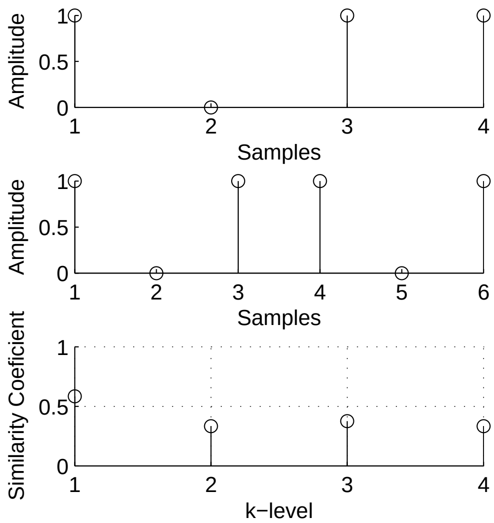


Figure 3: Example of two vectors and their Similarity Coefficients Vector (SCV)

combinations in the 3-level. According to the procedure describe above and shown in Fig. 2, we obtain the (2×4) 3-level Distance Matrix:

$$D^{(3)}(\mathbf{v}, \mathbf{w}) = \begin{bmatrix} 0 & 2 & 2 & 0 \\ 2 & 0 & 2 & 2 \end{bmatrix} \tag{5}$$

The 3-level coefficient can be read easily from this matrix and it is $c^{(3)} = 3/8 = 0.375$.

The complete SCV (Fig. 3) for the above example is given by

$$\mathbf{C} = [0.5833, 0.3333, 0.3750, 0.3333] \tag{6}$$

In the next section we show the algorithmic implementation and make some additional comments on our model.

3. Algorithm Implementation

We have implemented an algorithm in MATLAB which is able to construct, manipulate, and play the rhythm sequences defined by our model. We restrict our analysis below to rhythms without accentuation and articulation and also without pauses. This is a crude approximation to real rhythms and, in our model, it is accomplished by taking the simplest Beat Set, that is, $B = \{0, 1\}$. These aspects will be added in a further implementation of our formal modal described above.

We also devised a function to play back the input sequences, to do a subjective evaluation of the result of the measure. The events correspond to sinusoidal functions with exponential decay and we introduced a short tone at the starting point as a reference for the beginning of the sequence.

3.1. Similarity Coefficient Vector

The algorithm picks two sequences of elements extracted from the Beat Set and computes, for each k-level, the matrix $\mathbf{D}^{(k)}$. The meaning of a zero in a matrix element, corresponds to a perfect match between sub-sequences of the two input vectors. At this point we have as many matrices D as the length of the shortest input vector.

The sparsity of the matrix, which means the number of zeros in each $\mathbf{D}^{(k)}$ matrix, will give information on how similar are the subsequences of that particular k-level. The algorithm computes the ratio between the number of zeros and the product of the dimensions of $\mathbf{D}^{(k)}$ for each k-level and stores those values in the SCV.

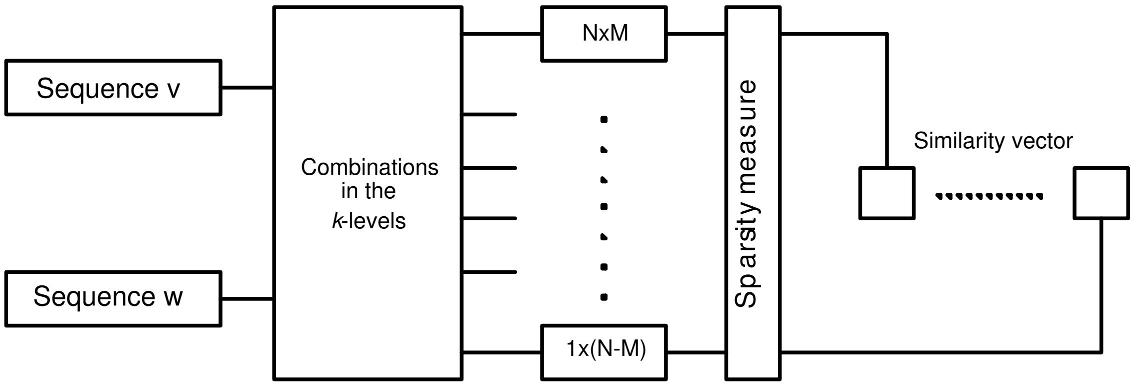


Figure 4: Model for the creation of the SCV between the two vectors

3.2. Analysis by Single Values

In addition to the SCV, we thought that it would also be interesting to reduce the quest for measuring similarity to a single value, enabling an easier comparison between the sequences. We offer three different solutions for this problem.

It is clear to us that the content of the first element provides a low level of information, as it only tells us wether the distribution of events (1s) and no-events (0s) in both of the vectors is even, or it is polarized towards having more events or no-events. On the other hand, the content of the last element of the SCV gives us the highest information. Finding a non-zero value in this position implies that the shortest input sequence exists at least once in the longest input sequence. In most of the comparisons this element will be zero. So the last non-zero element will tell us that what is the size of the longest sub-sequence that is common to both input sequences.

Another parameter that may be useful is the sum of the elements of SCV, which will take into account the coefficients from all the k-levels. For example, by considering

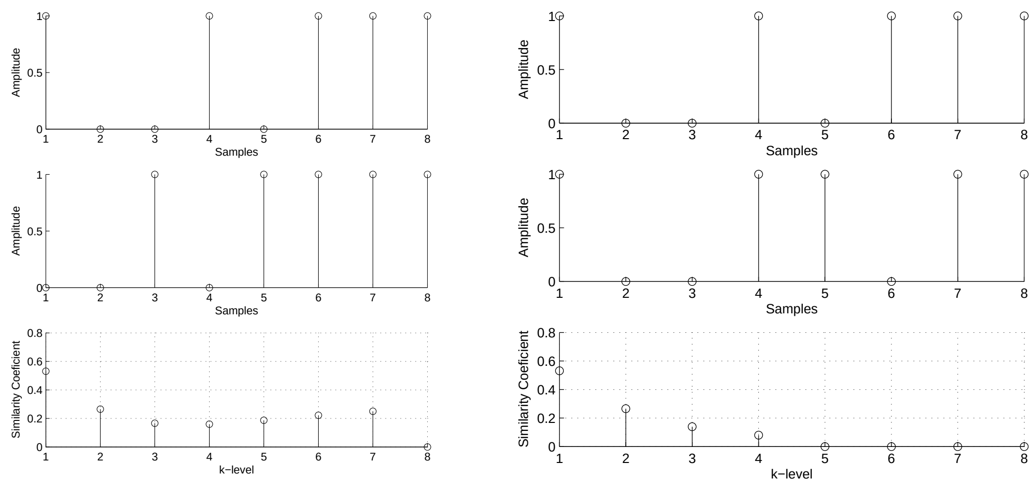


Figure 5: Vector with high values of the Similarity Coefficients values

the input sequence $\mathbf{v} = (1,0,0,1,0,1,1,1)$ we ran the distance for all possible $\mathbf{w}$ vectors of length 8. In Fig. 5 we present two vectors that show high similarity with the presented sequence. By maximizing the sum of the Similarity Coefficients vector and removing the input vector from the competition, we arrive to the value $\sum C_i = 1.7829$ with the most similar vector being $\mathbf{w} = (0,0,1,0,1,1,1,1)$ as can be seen in Fig. 5 (left). If instead we use the same vector and minimize the sum of the SCV elements, we get the least similar vector. For the example above the resultant value is $\sum C_i = 0.5179$, and the resultant $\mathbf{w}$ vector will be the no-event vector.

However, the sum above, does not consider, that there is greater importance in the rightmost elements of the vector. This can be achieved by taking a weighted sum of the elements, with an increasing profile of the weights.

4. Computer Applications

The formal model presented in this work is flexible enough to be used in several applications from rhythm analysis to creation of new rhythms, etc. Below we describe just two applications, namely, *Net-Rhythms* related to Neural Networks, developed by one of the authors(JM), and *RGeme*developed by another author (MG) in which AI agents learn rhythmic sequences from one another.

4.1. Neural Networks and Rhythms

Net-rhythms is a tool developed to classify and store rhythmic representations in a neural network. The framework used by this tool is constituted by a Neural Network called the SARDNET [James and Miikkulainen, 1995], an extended Kohonen self-organising feature map [Kohonen, 1985]. This network was developed to study the study of sequences and organization of phonemes in the context of language. We decided to explore its potential in the representation of rhythmic sequences, and new problems arose particularly related to the measurement of the distance between two vectors. The diagram on Fig. 6 explains how the network works.

The rhythms are coded according to the representation depicted in Fig. 1. Whenever a small rhythmic sequence $\mathbf{v}_t$ in time step t reaches the input, the distance from that sequence to all weight vectors w_j is computed. The neuron corresponding to weight more similar to the input, according to the defined distance, is activated and removed from further testing. As time progresses all activations are decayed, implying that after some time

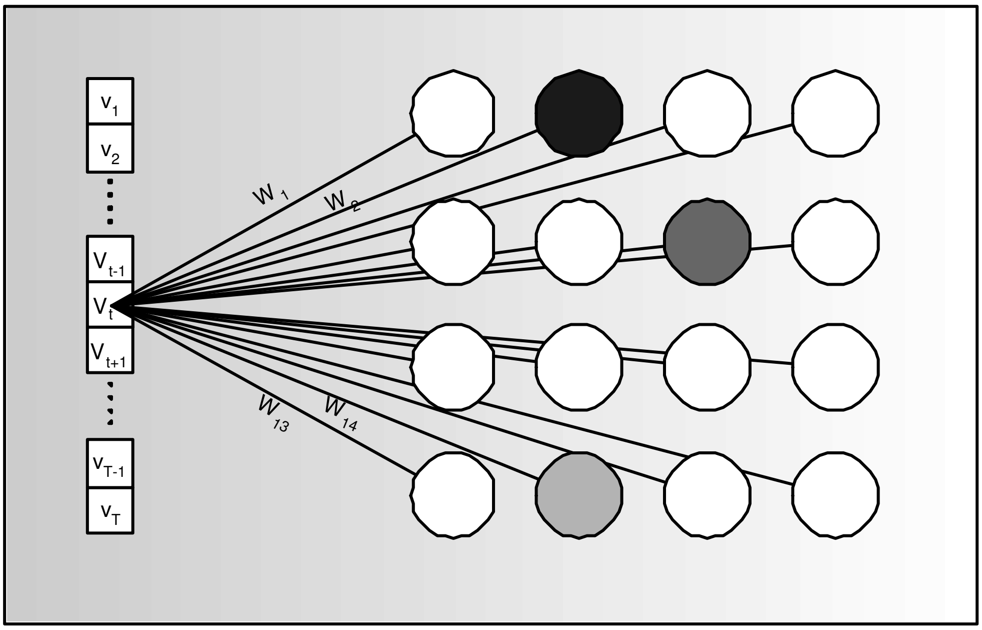


Figure 6: Diagram of the Sardnet with three activated neurons

steps there will be a ladder of activations in the network. In Fig. 6 there are three activated neurons represented by grey tonalities, corresponding the $\mathbf{w}_{14}$ weight to the first activated neuron, and $\mathbf{w}_2$ to the last one. Finally, after a complete sequence on time T, the weights from the activated neurons are slightly adapted in order to decrease their distances to input vectors.

This network can represent in a bidimensional space the rhythms that arrive sequentially to the input, and self-organize simulating a learning procedure.

As stated before, the choice of the wining neuron implies the measurement of the distance from the input to each of the neurons from the Sardnet. The distance proposed by the original creators of the network was the Euclidean distance, however there are some problems as this measure. The Euclidean distance does not allow sequences with different lengths and does not capture the similarity between equal sub-sequences that have their position shifted in time. The use of the SCV, and the other measures presented in Sec. 3.2, help solving the problems presented above.

4.2. Agents and Memes: A Rhythm Imitation Society

RGeme is an artificial intelligence system for the composition of rhythmic passages inspired by Richard Dawkin's theory of memes that is being presently developed in the Future Music Lab at the University of Plymouth. According to Dawkins [Dawkins, 1989, Dawkins, 1991], memes are basic units of cultural transmission in the same way that genes, in biology, are units of genetic information. Other researchers have already studied some applications of this concept in music [Cox, 2001, Gabora, 1997]

In *RGeme* a rhythmic composition is understood as the process of interconnecting sequences of basic elements (or ”rhythmic memes”) that have varied roles in the stream. Intelligent agents learn these roles from examples of musical pieces in order to evolve a ”musical worldview” which consists of a ”style matrix” of basic rhythms. During the learning stage, agents parse examples of pieces of music in search for rhythmic memes. These (candidates) memes are then compared with the agent's database of memes which we named Style Matrix (see Table 1), and are stored or transformed accordingly. For example, if the candidate meme is not already present in the agent's Style Matrix, it is copied and it's weight is set to 1. Now the model of distance of rhythm patterns is used in this application in order to upgrade the weight of each meme in Style Matrix. In this way the style of the memes evolves in time. In *RGeme* it was used the block distance. So all memes in the agent's style matrix have their weight upgraded according to their distance to the candidate meme.

Meme	dFL	dLL	nL	W
01011101	1	1	6	1.0385
11011101	1	1	31	1.0424
10001000	1	1	1	1.0181
10010101	1	1	1	1.0171
11011010	1	1	1	1.0159
10011010	1	1	4	1.0090
10011001	1	1	4	1.0075
11111111	1	1	1	1.0040
10000000	1	1	1	1.0000

Table 1: Extract from an Agent Style Matrix

where

- dFL: date the meme was first listened to
- dLL: date the meme was last listened to
- nL: number of times the meme was listened to
- W: upgraded weight

In the second stage, the system creates new rhythmic sequences (Production Phase) according to the musical structures and rules that were previously extracted from the styles of the pieces that were used in the learning stage. At this stage, agents are able to learn from each other's ”compositions” and capable of evolving new rhythmic styles by adapting to each other's rhythms. Clearly, new distances and similarity measures as shown above can be implemented in *RGeme*, which, of course could result in a different evolution of the memes society. This is presently under investigation.

5. Conclusion and Perspectives

We presented a model for measuring similarity in a general Rhythm Space, which include all the possible rhythm sequences. The key issue and innovative contribution of this work is the hierarchical subdivision of rhythm sequences in several levels and the construction of a Distance Matrix for each one of them. The information is coded in a Similarity Coefficient Vector (SCV), whose entries estimate the similarity between rhythm sequences in different k-levels. These coefficients are related to the sparsity the k-levels Distance Matrices. We also provided a easier to read single value measure for similarity. In addition we presented two applications for our formal model. Clearly, it can be applied in many other areas of music analysis and composition. It can be also applied on musical learning devices, such as self evaluation systems to relate played sequences to previously defined ones.

There is plenty of room to extend this work. For example, is yet to be done the comparison between this distance with other well established similarity measurements, such as the Levenshtein and Hausdorff distances. In addition to the block distance, we could use new basic p-distances and check which of them is better to the applications the user has in mind.

A further and also important problem is to link the formal similarity in this work to the rhythmic perception of human beings.

In this paper we have only shown the potential of our methods to construct similarity measures. The problem of perception of rhythm sequences deserves a deeper study by itself. This, as well comparisons with other methods, will be done in a future work.

6. Acknowledgments

This work was supported by the Leverhulme Trust (JM), the São Paulo State Research Foundation (FAPESP) (AMJ), and CAPES (Brazil) (MG, AMJ). A.M.J is grateful to the Computer Music Research group from the University of Plymouth, for the hospitality during the making of this work.

References

Cooper G., Meyer, L. (1963). *The Rhythmic Structure of Music*. University Of Chicago Press.

Cox, A. (2001). The mimetic hypothesis and embodied musical meaning. *MusicæScientiæ*, 2:195–212.

Dawkins, R. (1989). *The The Selfish Gene*. Oxford University Press, London.

Dawkins, R. (1991). *The Blind Watchmaker*. Penguin Books, London.

Gabora, L. M. (1997). The origin and evolution of culture and creativity. *Journal of Memetics -Evolutionary Models of Information Transmission*.

Gabrielsson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns i. *Scand. J. Psychol.*, 14:138–160.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events*. The MIT Press.

Hewlett, W. B. and Selfridge-Field, E., editors (2005). *Music Query: Methods, Models, and User Studies*. MIT Press.

James, D. L. and Miikkulainen, R. (1995). SARDNET: a self-organizing feature map for sequences. In Tesauro, G., Touretzky, D., and Leen, T., editors, *Advances in Neural Information Processing Systems 7*, pages 577–84, Cambridge, MA, USA. MIT Press.

Kohonen, T. (1985). The self-organizing map. In *Proceedings of IEEE*, volume 78, pages 1464–1480.

Minsky, M. (1988). *Society of Mind*. Simon & Schuster.

Miranda, E. R. (2004). At the crossroads of evolutionary computation and music: Self-programming synthesizers, swarm orchestras and the origins of melody. *Evolutionary Computation*, 12(2):137–158.

Mongeau, M. and Sankoff, D. (1990). Comparison of musical sequences. *Computers and the Humanities*, 24(3):161–175.

Orpen, K. S. and Huron, D. (1992). Measurement of similarity in music: A quantitative approach for non-parametric representations. *Computers in Music Research*, 4:1–44.

Povel, D. J. and Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4):411–441.

Takeda, H., Nishimoto, T., and Sagayama, S. (2003). Automatic rhythm transcription from multiphonic midi signals. In *ISMIR 2003*.

Computational Musicology: An Artificial Life Approach

Eduardo Coutinho, Marcelo Gimenes, João M. Martins and Eduardo R. Miranda
Future Music Lab
School of Computing, Communications & Electronics
University of Plymouth
Drake Circus
Plymouth PL4 8AA
UK
Email: {eduardo.coutinho,marcelo.gimenes,joao.martins,eduardo.miranda}@plymouth.ac.uk

***Abstract*— Artificial Life (A-Life) and Evolutionary Algorithms (EA) provide a variety of new techniques for making and studying music. EA have been used in different musical applications, ranging from new systems for composition and performance, to models for studying musical evolution in artificial societies. This paper starts with a brief introduction to three main fields of application of EA in Music, namely *sound design*, *creativity* and *computational musicology*. Then it presents our work in the field of computational musicology. Computational musicology is broadly defined as the study of Music with computational modelling and simulation. We are interested in developing A-Life-based models to study the evolution of musical cognition in an artificial society of agents. In this paper we present the main components of a model that we are developing to study the evolution of musical ontogenies, focusing on the evolution of rhythms and emotional systems. The paper concludes by suggesting that A-Life and EA provide a powerful paradigm for computational musicology.**

I. Introduction

Acoustics, Psychoacoustics and Artificial Intelligence (AI) have greatly enhanced our understanding of Music. We believe that A-Life and EA have the potential to reveal new understandings of Music that are just waiting to be unveiled.

EA have varied applications in Music, with great potential for the study of the artificial evolution of music in the context of the cultural conventions that may emerge under a number of constraints, including psychological, physiological and ecological constraints.

We identify three main fields of application of EA in Music: *sound design*, *creativity* and *computational musicology*. The following sections briefly survey these three main fields of application. Then we introduce our work in the field of computational musicology, inspired on A-Life techniques and EA.

A. Sound Design

The production of sound faced a revolution in the middle of the 20^{th} century with the appearance of the digital computer [1]. Computers were given instructions to synthesise new sounds algorithmically. Synthesisers (or *software synthesisers*) soon became organized as a network of functional elements (signal generators and processors) implemented in software. Comprehensive descriptions of techniques for computer sound synthesis and programming can be found in the literature [2].

The vast space of parameter values that one needs to manage in order to synthesise sounds with computers led many engineers to cooperate with musicians in order find effective ways to navigate in this space. Genetic algorithms (GA) have been successfully used for this purpose [3]. EA have also been used to develop topological organizations of the functional elements of a software synthesiser, using Genetic Programming (GP) [4].

The use of extremely brief time-scales gave rise to granular synthesis [5], a technique that suits the creation of complex sounds [6], adding more control problems to the existing techniques. One of the earliest applications of EA to granular synthesis is *Chaosynth*, a software designed by Miranda [7] that uses Cellular Automata (CA) to control the production of sound grains. *Chaosynth* demonstrates the potential of CA for the evolution of oscillatory patterns in a two-dimensional space. In most CA implementations, CA variables (or cells) placed on a 2D matrix are often associated with colours, creating visual patterns as the algorithm evolves in time. However, in *Chaosynth* the CA cells are associated with frequency and amplitude values for oscillators. The amplitude and frequency values are averaged within a region of the 2D CA matrix, corresponding to an oscillator. Each oscillator contributes a partial to the overall spectrum of a grain. The spectra of the grains are generated according to the evolution of the CA in time (Fig. 1).

More recently, Mandelis and Husbands [8] developed *Genophone*, a system that uses genetic operators to create new generations of sounds from two sets of preset synthesis parameters. Some parameters are left free to be manipulated with a data-glove by an external user, who also evaluates the fitness of the resulting sounds. Offspring sounds that are ranked best by the user will become parents of a new population of sounds. This process is repeated until satisfactory sounds are found.

B. Creativity

One interesting question with respect to the use of computers for aiding musical creativity is whether computers can

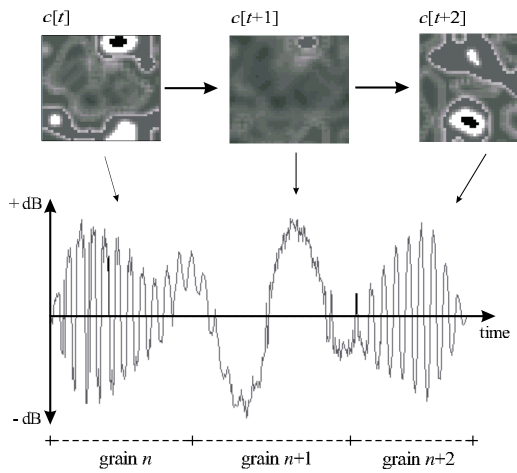


Fig. 1. Each snapshot of the CA produces correspond to a sound-grain. (Note, however, that this is only a schematic representation, as the grains displayed here do not actually correspond to these particular snapshots.)

create new kinds of musical compositions. In this case, the computer should neither be embedded with particular well-known compositional models at the outset nor learn from selected examples, which is not the case with most Artificial Intelligence-based systems for generating musical compositions.

Composers have used a number of mathematical models such as combinatorial systems, grammars, probabilities and fractals [9][10][11] to compose music that does not imitate well-known styles. Some of these composers created very interesting pieces of new music with these models and opened innovative grounds in compositional practices, e.g., the techniques created by Xenakis [12].

The use of the emergent behaviour of EA, on the other hand, is a new trend that is becoming very popular for its potential to generate new music of relatively good quality. A great number of experimental systems have been used to compose new music using EA: Cellular Automata Music [13], CA Music Workstation [14], CAMUS [15], MOE [16], GenDash [17], CAMUS 3D [18], Living Melodies [19] and Genophone [20], to cite but a few.

For example, *CAMUS* [15] takes the emergent behaviour of Cellular Automata (CA) to generate musical compositions. This system, however, goes beyond the standard use of CA in music in the sense that it uses a two-dimensional Cartesian representation of musical forms. In this representation the coordinates of a cell in the CA space correspond to the distances between the notes of a set of three musical notes.

As for GA-based generative music systems, they generally follow the standard GA procedures for evolving musical materials such as melodies, rhythms, chords, and so on. One example of such system is Vox Populi [21], which evolves populations of chords of four notes, through the operations of crossover and mutation.

EA have also been used in systems that allow for interaction in real-time; i.e., while the composition is being generated. In fact, most GA-based systems allow for this feature by letting the user to control GA operators and fitness values while the system is running. For example, Impett proposed an interesting swarm-like approach to interactive generative musical composition [22]. Musical composition is modelled here as an agent system consisting of interacting embodied behaviours. These behaviours can be physical or virtual and they can be emergent or preset. All behaviours co-exist and interact in the same world, and are adaptive to the changing environment to which they belong. Such behaviours are autonomous, and prone to aggregation and generation of dynamic hierarchic structures.

C. Computational Musicology

Computational musicology is broadly defined as the study of Music by means of computer modelling and simulation. A-Life models and EA are particularly suitable to study the origins and evolution of music. This is an innovative approach to a puzzling old problem: if in Biology the fossils can be studied to understand the past and evolution of species, these "fossils" do not exist in Music; musical notation is a relatively recent phenomenon and is most prominent only in the Western world. We are aware that Musicology does not necessarily need computer modelling and simulation to make sense. Nevertheless, we do think that "*in silico*" simulation can be useful to develop and demonstrate specific musical theories. These theories have the advantage that they can be objective and scientifically sound.

Todd and Werner [23] proposed a system for studying the evolution of musical tunes in a community of virtual composers and critics. Inspired by the notion that some species of birds use tunes to attract a partner for mating, the model employs mating selective pressure to foster the evolution of fit composers of courting tunes. The model can co-evolve male composers who play tunes (i.e., sequences of notes) along with female critics who judge those songs and decide with whom to mate in order to produce the next generation of composers and critics. This model is remarkable in the sense that it demonstrates how a Darwinian model with a pressure for survival mechanism can sustain the evolution of coherent repertoires of melodies in a community of software agents. Miranda [24] [25] proposed a mimetic model to demonstrate that a small community of interactive distributed agents furnished with appropriate motor, auditory and cognitive skills can evolve from scratch a shared repertoire of melodies (or tunes) after a period of spontaneous creation, adjustment and memory reinforcement. One interesting aspect of this model is the fact that it allows us to track the development of the repertoire of each agent of the community. Metaphorically, one could say that such models enable us to trace the musical development (or "education") of an agent as it gets older.

From this perspective we identify three important components of an Artificial Musical Society: agents synchronization, knowledge evolution, and emotional content in performance. The first presents itself as the basis for musical communication between agents. The second, rooted on the first, allows musical information exchange, towards the creation of a cultural environment. Finally we incorporate the indispensable influence of emotions in the performance of the acquired music knowledge. The following sections present this three aspects separately. Even though they are parts of the same model, experiments were run separately. We are working towards the complete integration of the model, and co-evolution of the musical forms: from motor response to compositional processes and performances.

II. Emergent Beat Synchronisation

A. *Inspiration: Natural Timing*

Agents interacting with one another by means of rhythm need mechanisms to achieve beat synchronisation.

In his book *Listening*, Handel [26] argues that humans have a biological constrain referred to as *Natural Timing* or *Spontaneous Tempo*. This means that when a person is asked to tap an arbitrary tempo, they will have a preference. Furthermore, if the person is asked to tap along an external beat that is faster or slower, and if the beat suddenly stops, then they will tend to approximate to their preferred tempo. The tap interval normally falls between 200 msec and 1.4 sec, but most of the tested subjects were in the range of 200 - 900 msec [27]. The claim that this phenomenon is biologically coded rises from the extreme proximity of these values when observed in identical twins. The same disparity observed for unrelated subjects is observed in fraternal twins. The time interval between two events is called Inter-Onset Interval (IOI).

In our model, the agents "are born" with different natural timings by default. As they interact with each other, each agent adapts its beat to the beats of the other agents.

B. *Synchronisation Algorithm*

Computational modeling of beat synchronisation has been tackled in different ways. Large and Kolen devised a program that could tap according to a rhythmic stimulus with nonlinear-oscillators [28], using the gradient descendant method to update their frequency and phase. Another approach, by Scheirer, consisted of modelling the perception of meter using banks of filters [29]. We propose an algorithm based on Adaptive Delta Pulse Code Modulation (ADPCM) that enables the adaptation of different agents to a common ground pulse, instead of tracking a given steady pulse. Our algorithm proved to be more compatible with Handel's notion of natural timing, as discussed in the previous section. As in ADPCM for audio, where a variable time step tracks the level of an audio signal, the agent in our model uses a variable time step to adjust its IOI to an external beat. The agent counts how many beats from the other agents fit into its cycle and it determines its state based on one of the following conditions: SLOW (listened to more than one beat), FAST (no beats were listened), or POSSIBLY_SYNCHRONISED (listened to one beat). Depending on whether the agent finds itself in one of the first two states, it increases or decreases the size of their IOIs. Delta corresponds to the amount by which the value of an IOI is changed. If the agent is in the POSSIBLY_SYNCHRONISED state and the IOIs do not match, then there will be a change of state after some cycles, and further adjustments will be made until the IOIs match. However, the problem is not solved simply by matching the IOI of the other agent. Fig. 2(b) illustrates a case where the IOIs of two agents are the same but they are out of phase. An agent solves this problem by delaying its beat until it produces a beat that is close to the beat of the other agent (Fig. 2(c)).

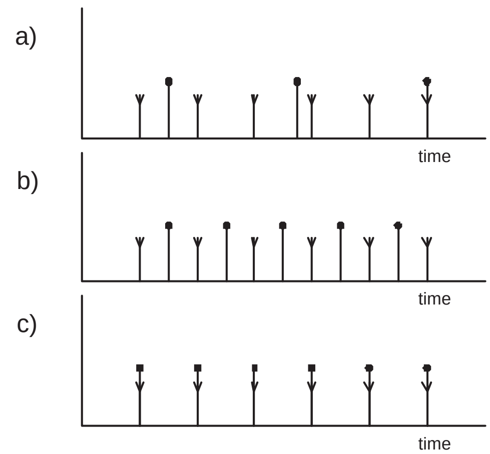


Fig. 2. (a) The agents have different IOIs; (b) The agents have the same IOI but they are out of phase; (c) The IOIs are synchronised.

C. *Experiment and Result*

In this section we present the result of an experiment with two agents adapting to each other's beats. Fig. 3 shows the temporal evolution of the IOIs of the agents. The minimum value for Delta, which is also the initial value of the time step, is different for the two agents. If the agent recognises that it is taking too long to change its state, the former value of Delta is multiplied by 2. Oscillatory patterns were observed when they were close to finding a common beat, due to the fact that both agents changed their IOIs and phases when they realised that they were not synchronised. The solution to this problem was solved by letting only one of the agents to change the phase after hearing one beat from the other agent.

Agent 1 started with an IOI equal to 270 ms and it had an initial adaptation step of 1 ms. Agent 2 started with an initial IOI equal to 750 ms and it had an initial adaptation step of 3 ms. Fig. 3 shows that the agents were able to find a matching IOI of 433 ms and synchronise after 26 beats. Notice that they found a common IOI after 21 beats, but they needed 5 more beats to synchronise their phases.

One interesting alternative that requires further study is the interaction between the agents and a human player. In the present case study the system requires many beats to reach synchronisation, but it is expected that the ability that humans have to find a common beat quickly may introduce a shortcut into the whole process.

In this experiment, the "spontaneous tempo" and the Delta values of the agents were initialised by hand. But once the synchronisation algorithm is embedded in a model to study

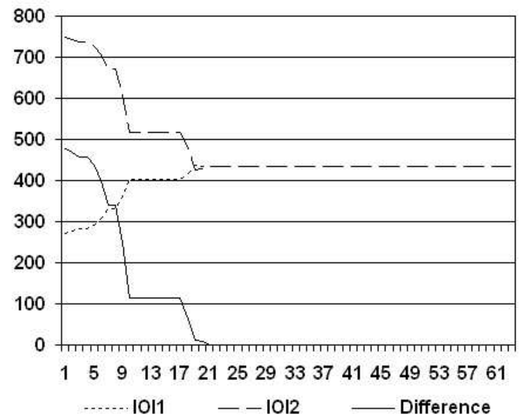


Fig. 3. Evolution of IOIs and their difference.

the evolution of musical rhythm one needs to implement a realistic way to initialise these values. Different agents can be implemented with different default Delta value but it would be more realistic to devise a method to modulate such value in function of some form of musical expression, or semantics. In order to do this, we are looking into ways in which we could program the agents to express emotions. In this case, the agents should be given the ability to modulate Delta coefficients and initial deviations from their "spontaneous tempo" in function of their emotional state. In section IV we present the first phase of an emotional system that we are developing to implement this.

III. Musical Ontogenesis in an Artificial Society

In Philosophy of Science, ontogenesis refers to the sequence of events involved in the development of an individual organism from its birth to its death. We therefore use the term musical ontogenesis to refer to the sequence of events involved in the development of the musicality of an individual. Intuitively, it should be possible to predict the music style of future musicians according to restrained music material that is absorbed during their formative stages. But would it be possible to objectively study the way in which composers or improvisers create music according to their educational background? Although it may be too difficult to approach this subject with real human musicians, we suggest that it should be possible to develop such studies with artificial musicians. A model of musical ontogenesis is therefore useful to study the influence of the musical material learned during the formative years of artificial musicians, especially in systems for musical composition and improvisation. A growing number of researchers are developing computer models to study cultural evolution, including musical evolution ([30] [31] [32] [33]). Gimenes [34] presents RGeme, an artificial intelligence system for the composition of rhythmic passages inspired by Richard Dawkin's theory of memes. Influenced by the notion that genes are units of genetic information in Biology, memes are defined as basic units of cultural transmission. A rhythmic composition would be understood as a process of interconnecting ("composition maps") sequences of basic elements ("rhythmic memes"). Different "rhythmic memes" have varied roles in the stream. These roles are learned from the analysis of musical examples given to train the system.

A. RGeme

The overall design of the system consists of two broad stages: the learning stage and the production stage. In the learning stage, software agents are trained with examples of musical pieces in order to evolve a "musical worldview". The dynamics of this evolution is studied by analysing the behaviour of the memes logged during the interaction processes.

At the beginning of a simulation a number of Agents is created. They sense the existence of music compositions in the environment and choose the ones with which they will interact, according to some previously given parameters such as the composer's name and the date of composition.

Agents then parse the chosen compositions to extract rhythmic memes (Candidate Memes) and composition maps. The new information is compared with the information that was previously learned and stored in a matrix of musical elements (Style Matrix). All the elements in the Style Matrix possess a weight that represents their relevance over the others at any given moment. This weight is constantly changing according to a transformation algorithm that takes into account variables such as the date the meme was first listened to, the date it was last listened to and a measure of distance that compares the memes stored in the Style Matrix and the Candidate Memes. These features can be seen in more detail in [34].

At last, in the production phase the Agents execute composition tasks mainly through the reassignment of the various Composition Maps according to the information previously stored in the learning phase.

B. Experiment and Result

The different Style Matrices that are evolved in an agent's lifetime represent the evolution of its musical worldview. One can establish the importance of the diversity of the raw material (in terms of developing different musical worldviews) based on the data stored in the Style Matrix's log files. It is possible to directly control the evolution of an agent's worldview, for instance, by experimenting with different sets of compositions originated from different composers.

In Fig. 4 we show the results obtained from an experiment involving a simple learning scenario. During a given period of time an agent only interacted with a particular set of compositions by Brazilian composer Ernesto Nazareth. Afterwards, the agent interacted with a different set of compositions by another Brazilian composer, Jacob do Bandolim. In the same figure, each line represents the evolution of the relative importance (weight) of a small selection of memes that the agent learned during the simulation. Fig. 5 shows the musical notation for each one of these memes. We can observe different behaviours in the learning curves, which means that the agent was exposed to each one of these memes in different ways.

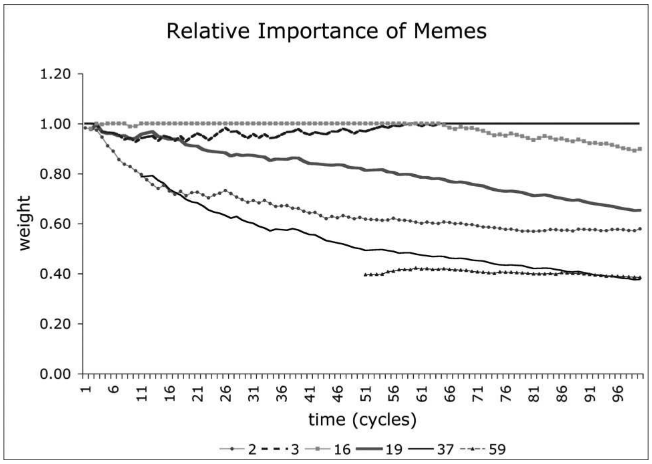


Fig. 4. Relative importance of memes in time.

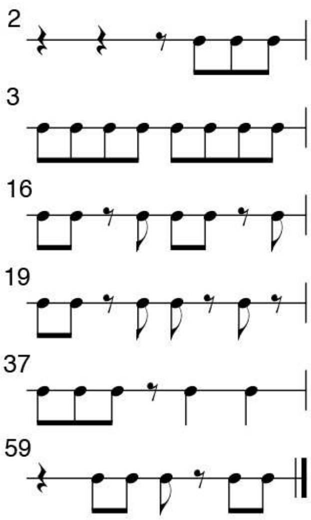


Fig. 5. Musical representation of rhythmic memes.

RGeme has the potential to execute intricate simulations with several Agents learning at the same time from rhythms by composers from inside and outside the system's environment. We believe that this model will allow for the objective establishment of a sophisticated musical ontogenesis through which one will be able to control and predict the musical culture of the inhabitants of artificial communities.

There is however a number of problems that needs to be addressed in order to increase the complexity of this model. One such problem is beat synchronisation, which has been discussed in the previous section. It is possible to observe the behaviour of thousands of male fireflies flashing synchronously during their mating season. Each insect has its own preferred pulse but they gradually adjust their pulses to a single global beat by observing each other [35]. Different humans also have their own preferred pulses, which are driven towards synchrony when engaged in collective musical performance with other humans, non-humans or both. As with fireflies, this mechanism is believed to be biologically coded in humans.

Nonetheless, music is mostly the result of a cultural context [36]. Specially in our research, the rules for composition and performance should emerge from social interactions of agents.

IV. Modelling Emotions

A. Expressivity

The use of expressive marks by Western composers documents well the common assumption that emotions play an important role in music performance.

Expressive marks are performance indications, typically represented as a word or a short sentence written at the beginning of a movement, and placed above the music staff. They describe to the performer the intended musical character, mood, or emotion as an attribute of time, as for example, *andante con molto sentimento*, where *andante* represents the tempo marking, and *con molto sentimento* its emotional attribute.

Before the invention of the metronome by Dietrich Nikolaus Winkel in 1812, composers resorted to words to describe the tempo (the rate of speed) in a composition: Adagio (slowly), Andante (walking pace), Moderato (moderate tempo), Allegretto (not as fast as allegro), Allegro (quickly), Presto (fast). The metronome's invention provided a mechanical discretization of musical time by a user chosen value (beat-unit), represented in music scores as the rate of beats per minute (quarter-note = 120). However, after the metronome's invention, words continued to be used to indicate tempo, but now often associated with expressive marks. In some instances, expressive marks are used in lieu of tempo markings, as previous associations indicate the tempo being implied (e.g. *funebre* implies a slow tempo).

The core "repertoire" of emotional attributes in music remains short. Expressions such as *con sentimento*, *con bravura*, *con affetto*, *agitato*, *appassionato*, *affetuoso*, *grave*, *piangendo*, *lamentoso*, *furioso*, and so forth, permeate different works by different composers since Ludwig van Beethoven (1770-1827) (for an example see Fig. 6) . But what exactly do these expressions mean?

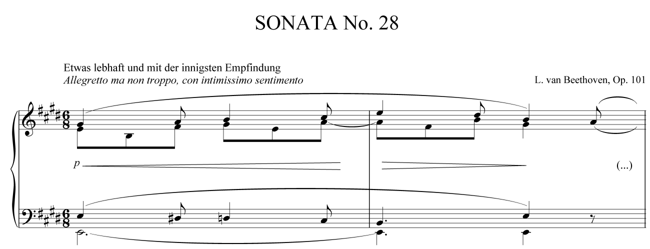


Fig. 6. Beethoven score: example of use of emotional attributes.

Each performer holds a different system of beliefs of what expressions such as *con sentimento* represent, as our understanding of emotions has not yet reduced them to a lawful behaviour. Without consensus on the individual meaning of such marks, a performance *con sofrimento* is indistinguishable from one *con sentimento*, since both expressions presume an equally slow tempo. Although we have no agreement on the meaning of expressive marks and their direct musical consequences, musicians have intuitively linked expressivity with irregularity within certain boundaries. Celebrated Polish pianist and composer Ignacy Jan Paderewski (1860-1941) stated: "every composer, when using such words as *espressivo*, *con molto sentimento*, *con passione*, and so on, demands (...) a certain amount of emotion, and emotion excludes regularity... to play Chopin's G major Nocturne with rhythmic rigidity and pious respect for the indicated rate of movement would be (...) intolerably monotonous (...). Our human metronome, the heart, under the influence of emotion, ceases to beat regularly - physiology calls it arrhythmic, Chopin played from his heart. His playing was not rational, it was *emotional*" [37].

Composers are well aware that a clear representation of the musical idea reduces ambiguity in the interpretation of the message (the music score). However, the wealth of shadings, accents, and tempo fluctuations found in human performances are, at large, left unaccounted by the composer as the amount of information required to represent these type of nuances carries, in practice, no linear bearing in the detail human performers can faithfully reproduce.

While the electronic and computer music mediums provide composers the power to discretize loudness and time related values in very small increments (for example, MIDI systems [38] use 128 degrees of loudness, and time measured in milliseconds), we note that music scores for human performances use eight approximate levels of loudness (ppp, pp, p, mp, mf, f, ff, fff), and time is discretized in values hundreds of milliseconds long. If we compare any two "faithful" human performances of a work, we conclude that, from performance to performance, only the order of notes remains strictly identical.

Expression marks operate as synesthesia, that is, the stimulation of one sense modality to rise to a sensation in another sense modality [39]. Although their direct musical consequences remain unclear, we can deduce which musical levels are susceptible of being influenced: time and loudness.

These are structural levels where small value changes produce significantly different results. The amount of information needed to describe such detail in fine resolution falls outside the precision limits with which human performers process a music score to control time and the mechanics of traditional music instruments.

> "*Look at these trees!*" Liszt told one of his pupils, "*the wind plays in the leaves, stirs up life among them, the tree remains the same. That is Chopinesque rubato*[1]."

B. Emotions

We go back to the 19^{th} century to find the earliest scientific studies: Darwin's observations about bodily expression of emotions [40], James's studies on the meaning of emotion [41], and Wundt's work on the importance of emotions for Psychology [42]. But studies on behaviour focused for many years only on higher level cognitive processes, discarding emotions [43]. Still, emotions were occasionally discussed, and the ideas changed considerably within the last decade or so. Research connecting mind and body, and the role of emotions in rational thinking gained prominence after the work of Cannon and Bard [44]. In short, they suggested that there are parallel neural paths from our senses to the experience of an emotion and to its respective physiological manifestation. Later Tomkins [45][46], Plutchik [47][48] and Izard [49][50][51][52] developed similar theories. They suggested that emotions are a group of processes of specific brain structures and that each of these structures has a unique concrete emotional content, reinforcing their importance. Ekman proposed a set of basic (and universal) emotions [53], based on cross-cultural studies [54]. These ideas were widely accepted in evolutionary, behavioural and cross-cultural studies, by their proven ability to facilitate adaptive responses.

Important insights come from Antonio Damasio [55][56][57], who brought to the discussion some strong neurobiological evidence, mainly exploring the connectivity between body and mind. He suggested that, the process of emotion and feeling are part of the neural machinery for biological regulation, whose core is formed by homeostatic controls, drives and instincts. Survival mechanisms are related this way to emotions and feelings, in the sense that they are regulated by the same mechanisms. Emotions are complicated collections of chemical and neural responses, forming a pattern; all emotions have some regulatory role to play, leading in one way or another to the creation of circumstances advantageous to the organism exhibiting the phenomenon. The biological function of emotions can be divided in two: the production of a specific reaction to the inducing situation (e.g. run away in the presence of danger), and the regulation of the internal state of the organism such that it can be prepared for the specific reaction (e.g. increased blood flow to the arteries in the legs so that muscles receive extra oxygen and glucose, in order to escape faster). Emotions are inseparable from the idea of reward or punishment, of pleasure or pain, of approach or withdrawal, or personal advantage or disadvantage.

Our approach to the interplay between music and emotions follows the work of these researchers, and the relation between physiological variables and different musical characteristics [58]. Our objective is to develop a sophisticated model to study music performance related to an evolved emotional system. The following section introduces the first result of this development.

C. The Model

The current version of our model consists of an agent with complex cognitive, emotional and behavioural abilities. The agent lives in an environment where it interacts with several objects related to its behavioural and emotional states. The agent's cognitive system can be described as consisting

[1] *Rubbato*: from the Italian "robbed", used to denote flexibility of tempo to achieve expressiveness.

Physiological Data	Drives	Variation
Adrenaline	Explore	neural activity (arousal)
Blood Sugar	Hunger	metabolism, food
Endorphine	Boredom	metabolism, toys
Energy	Fatigue	metabolism, bed
Vascular Volume	Thirst	metabolism, water
Pain	Withdraw	metabolism, obstacles
Heart Rate	-	metabolism, all objects

TABLE I
PHYSIOLOGICAL DATA, DRIVES, AND THEIR DYNAMICS.

of three main parts: Perceptual, Behavioural, and Emotional systems.

The Perceptual system (inspired in LIVIA [59] and GAIA [60]) receives information from the environment through a retina modelled as close as possible to a biological retina in functional terms. It senses a bitmap world through a *ray tracing* algorithm, inspired by the notion that photons travel from the light-emitting objects to the retina. The Behavioural system is divided into two sub-systems: Motivational and Motor Control. These sub-systems define the interaction of the agents with their environment. While the agents interact with objects and explore the world, the Motivational sub-system uses a feed-forward neural network to integrate visual input and information about their internal and physiological states. The network learns through a reinforcement learning algorithm. As for the Motor Control sub-system, the agents control their motor system by means of linear and angular speed signals, allowing them to navigate in their world; this navigation includes obstacle avoidance and object interaction. The Emotional system considers the role of emotions as part of an homeostatic mechanism [56]. The internal body state of an agent is defined by a set of physiological variables that vary according to their interaction with the world and a set of internal drives. The physiological variables and the internal drives in the current version of the model are listed in Table I. The agents explore the world and receive the stimuli from it. Motor Control signals are also controlled by the neural network. There are several types of objects: food, water, toys, beds, and obstacles. Each of them is related to one or more physiological variables. Interacting with objects causes changes in their internal body state. For instance, the Vascular Volume (refer to Table 1) of an agent will be increased if it encounters water and manifests the desire to drink it. The agent's own metabolism can also change physiological data; e.g. moving around the world decrease the energy level of an agent. An emotional state reflects the agent's well-being, and influences its behaviour through an amplification of its body alarms. For further details on the model, refer to [61].

We propose that these emotional states affect music performance, reflecting the agent emotional state in the music. There are a few studies regarding the communication of emotions through music; for further details please refer to [58]. We simulated different musical performance scenarios inspired by these studies, and the next section presents the outcome from

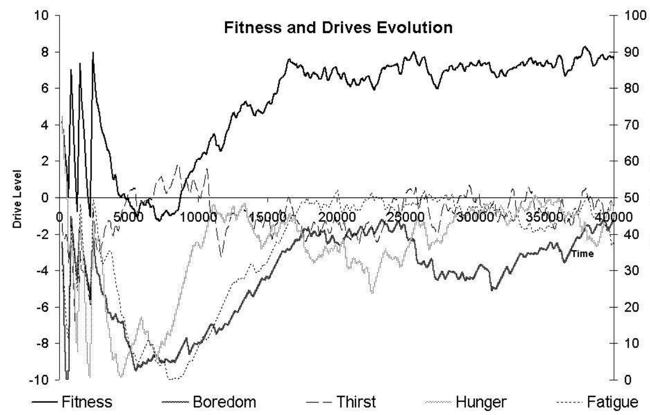


Fig. 7. Fitness vs Drives Evolution.

running the Emotional part of our model.

D. Experiment and Result

The objective of this experiment was to analyze the ability of an agent to regulate its homeostasis. To achieve this task we studied the emergence of associations between world stimuli with internal needs; in other words, an implicit world/body map. Fig. 7 shows the relation between fitness function (reflecting the agent's well-being) and the evolution of the agent's drives. The values are averages for each 200 iterations intervals. An overall increase of fitness is shown, suggesting that the agent is capable to adapt itself to new environments. Fig. 7 also shows a decrease of the amplitude of the drives as time evolves. By looking at the evolution of the drives in time we can observe that they were maintained within a certain range. This reflects the ability of the agent to respond to its "body needs". Apparently the agent not only learned how to adapt to the environment, but also did it effectively, maintaining a "healthy behaviour" by self-regulation of the homeostatic process.

A complete analysis of the system is presented in [61] .

E. Performance

Two physiological variables, selected for their influence in actual human performances [58], *Heart Rate* and *Adrenaline*, control *tempo* and *velocity* (loudness) in the performance of a piece of music [62], reflecting neural activity and emotions valence (whether positive or negative), mirroring the agent's emotional state. *Heart Rate* values modulate the on-times of events within each measure (bar), in this case 4000 ms, with a maximum deviation of +/- 640 ms. *Adrenaline* values modulate events' velocity (loudness) between user chosen limits, in this case, 80 and 127. The results can be heard at http://cmr.soc.plymouth.ac.uk/ecoutinho/ (link *Polymnia*).

We collected the data from the simulation in the previous section to "perform" a piece of music [62]; in this case to playback a MIDI recording of a piece. In Fig. 8 we present the first measure of the piece. The anatomy of each note here represented by three parameters (MIDI messages): note-number, note-duration (measured in ms), and velocity

Fig. 8. Score: J.S.Bach - *Prelude no I, BWV 846, from the Well Tempered Klavier I*

Heart Beat	**Note duration**	**Adrenaline**	**Amplitude**	**MIDI**
71.000	226.387	5.207	92	[145 60 92]
69.263	239.539	9.228	94	[145 64 94]
69.649	246.098	9.652	97	[145 67 97]
69.991	257.054	9.538	97	[145 72 97]
69.770	233.929	9.466	86	[145 76 86]
70.269	264.449	9.962	112	[145 67 112]
70.659	249.279	9.962	98	[145 72 98]
70.548	255.966	10.239	121	[145 76 121]
70.890	237.331	10.335	99	[145 60 99]
70.902	278.871	10.334	102	[145 64 102]
70.908	256.56	10.363	105	[145 67 105]
70.950	250.47	10.288	109	[145 72 109]
71.301	265.184	10.392	85	[145 76 85]
71.323	237.109	10.084	107	[145 67 107]
71.322	261.263	10.078	90	[145 72 90]
71.655	240.51	10.169	112	[145 76 112]

TABLE II

PERFORMANCE DATA (MIDI MESSAGES: [INSTRUMENT PITCH VELOCITY] - PIANO.

(loudness). In the original MIDI file notes are played every 250 ms. In our piece their duration varies according to *Heart Rate* value(see Fig. II). Velocity (or loudness) is controlled by the level of *Adrenaline*. The system related *Heart Rate* onto music by mirroring stable or unstable situations, relaxation or anxiety with deviations from original rhythmic structure of each measure of music, and *Adrenaline*, by, on the one hand, mirroring excitement, tension, intensity, or, on the other hand, boredom, low arousal, by changes in note-velocity (loudness); refer to Table II.

We are currently testing the model with different conditions and metabolism, specifically the amount of resources needed to satisfy drives and the way in which these drives decrease and increase in time. A deep analysis of the behaviour of the model may reveal that performance in different environments and with different agent metabolisms can play a strong role in the affective states.

V. CONCLUDING REMARKS

At the introduction of this paper we indicated that EA has been used in a number of musical applications, ranging from sound synthesis and composition to computational musicology. An increasing number of musicians have been using EA for artistic purposes since the early 1980s. However, the potential of EA for computational musicology started to be explored only recently, after the works by researchers such as Todd, Kirby and Miranda [23] [24] [25] [63].

This paper presented three components of an A-Life model (using EA) that we are developing to study the development of musical knowledge, rooted on the problem of beat synchronisation, knowledge evolution and emotional systems.

Although the A-Life approach to computational musicology is still incipient, this paper reinforced the notion that a new approach to computational musicology is emerging.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support of the Portuguese Foundation for Science and Technology (FCT, Portugal), the Brazilian Ministry of Education (CAPES, Brazil), and The Leverhulme Trust (United Kingdom).

REFERENCES

[1] M. V. Mathews, "The digital computer as a music instrument," *Science*, vol. 142, no. 11, pp. 553–557, 1963.

[2] E. R. Miranda, *Computer Sound Design: Synthesis Techniques and Programming*. Oxford, UK: Focal Press, 2002.

[3] A. Horner, J. Beauchamp, and L. Haken, "Machine tongues XVI: Genetic algorithms and their application to fm matching synthesis," *Computer Music Journal*, vol. 17, no. 4, pp. 17–29, 1993.

[4] R. Garcia, "Growing sound synthesizers using evolutionary methods," in *Proceedings of ALMMA 2002 Workshop on Artificial Models for Musical Applications*, E. Bilotta, E. R. Miranda, P. Pantano, and P. M. Todd, Eds. Cosenza, Italy: Editoriale Bios, 2002, pp. 99–107.

[5] D. Gabor, "Acoustical quanta and the theory of hearing," *Nature*, no. 1044, pp. 591–594, 1947.

[6] P. Thomson, "Atoms and errors: towards a history and aesthetics of microsound," *Organised Sound*, vol. 9, no. 2, pp. 207–218, 2004.

[7] E. R. Miranda, *Composing Music with Computers*. Oxford, UK: Focal Press, 2001.

[8] J. Mandelis and P. Husbands, "Musical interaction with artificial life forms: Sound synthesis and performance mappings," *Contemporary Music Review*, vol. 22, no. 3, pp. 69–77, 2003.

[9] C. Dodge and T. Jerse, *Computer Music*. London, UK: Schimer Books, 1985.

[10] D. Cope, *Computers and Musical Style*. Oxford, UK: Oxford University Press, 1991.

[11] D. Worral, "Studies in metamusical methods for sound image and composition," *Organised Sound*, vol. 1, no. 3, pp. 183–194, 2001.

[12] I. Xenakis, *Formalized Music: Thought and Mathematics in Composition.* Bloomington (IN), USA: Indiana University Press, 1971.

[13] D. Millen, "Cellular automata music," in *International Computer Music Conference ICMC90*, S. Arnold and D. Hair, Eds. San Francisco (CA), USA: ICMA, 1990, pp. 314–316.

[14] A. Hunt, R. Kirk, and R. Orton, "Musical applications of a cellular automata workstation," in *International Computer Music Conference ICMC91*. San Francisco (CA), USA: ICMA, 1991, pp. 165–168.

[15] E. R. Miranda, "Cellular automata music: An interdisciplinary music project," *Interface (Journal of New Music Research)*, vol. 22, no. 1, pp. 03–21, 1993.

[16] B. Degazio, "La evolucion de los organismos musicales," in *Musica y nuevas tecnologias: Perspectivas para el siglo XXI*, E. R. Miranda, Ed. Barcelona, Spain: L'Angelot, 1999, pp. 137–148.

[17] R. Waschka II, "Avoiding the fitness bottleneck: Using genetic algorithms to compose orchestral music," in *International Computer Music Conference ICMC99*. San Francisco (CA), USA: ICMA, 1999, pp. 201–203.

[18] K. McAlpine, E. R. Miranda, and S. Hogar, "Composing music with algorithms: A case study system," *Computer Music Journal*, vol. 223, no. 2, pp. 19–30, 1999.

[19] P. Dahlstedt and M. G. Nordhal, "Living melodies: Coevolution of sonic communication," *Leonardo*, vol. 34, no. 2, pp. 243–248, 2001.

[20] J. Mandelis, "Genophone: An evolutionary approach to sound synthesis and performance," in *Proceedings of ALMMA 2001 Workshop on Artificial Models for Musical Applications*, E. Bilotta, E. R. Miranda, P. Pantano, and P. M. Todd, Eds. Cosenza, Italy: Editoriale Bios, 2001, pp. 37–50.

[21] J. Manzolli, A. Moroni, F. von Zuben, and R. Gudwin, "An evolutionary approach applied to algorithmic composition," in *VI Brazilian Symposium on Computer Music*, E. R. Miranda and G. L. Ramalho, Eds. Rio de Janeiro, Brazil: EntreLugar, 1999, pp. 201–210.

[22] J. Impett, "Interaction, simulation and invention: a model for interactive music," in *Proceedings of ALMMA 2001 Workshop on Artificial Models for Musical Applications*, E. Bilotta, E. R. Miranda, P. Pantano, and P. M. Todd, Eds. Cosenza, Italy: Editoriale Bios, 2001, pp. 108–119.

[23] P. M. Todd and G. Werner, "Frankensteinian methods for evolutionary music composition," in *Musical networks: Parallel distributed perception and performance*, N. Griffith and P. M. Todd, Eds. Cambridge (MA), USA: MIT Press/Bradford Books, 1999, pp. 313–339.

[24] E. R. Miranda, "At the crossroads of evolutionary computation and music: Self-programming synthesizers, swarm orchestras and the origins of melody," *Evolutionary Computation*, vol. 12, no. 2, pp. 137–158, 2004.

[25] E. R. Miranda, S. Kirby, and P. Todd, "On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars," *Contemporary Music Review*, vol. 22, no. 3, pp. 91–111, 2003.

[26] S. Handel, *Listening: An Introduction to the Perception of Auditory Events*. Cambridge (MA), USA: The MIT Press, 1989.

[27] P. Fraisse, *The Psychology of Time*. New York, USA: Harper & Row, 1963.

[28] E. W. Large and J. F. Kolen, "Resonance and the perception of musical meter," *Connection Science*, vol. 6, pp. 177–208, 1994.

[29] E. Scheirer, "Tempo and beat analysis of acoustic musical signals," *Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 588–601, 1998. [Online]. Available: http://web.media.mit.edu/ eds/beat.pdf

[30] R. Dawkins, *The Blind Watchmaker*. London, UK: Penguin Books, 1991.

[31] A. Cox, "The mimetic hypothesis and embodied musical meaning," *MusicæScientiæ*, vol. 2, pp. 195–212, 2001.

[32] L. M. Gabora, "The origin and evolution of culture and creativity," *Journal of Memetics - Evolutionary Models of Information Transmission*, vol. 1, no. 1, pp. 1–28, 1997.

[33] S. Jan, "Replicating sonorities: towards a memetics of music," *Journal of Memetics - Evolutionary Models of Information Transmission*, vol. 4, no. 1, 2000.

[34] M. Gimenes, E. R. Miranda, and C. Johnson, "Towards an intelligent rhythmic generator based on given examples: a memetic approach," in *Digital Music Research Network Summer Conference 2005*. Glasgow, UK: The University of Glasgow, 2005, pp. 41–46.

[35] S. Strogratz and I. Stewart, "Coupled oscilators and biological synchronization," *Scientific American*, vol. 26, pp. 68–74, 1993.

[36] J. A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music*. Oxford, UK: Clarendon Press, 1985.

[37] "Paderewski," http://www.wyastone.co.uk/nrl/gpiano/8816c.html, Last visited 27 April 2005.

[38] P. White, *Basic Midi*. London, UK: Sanctuary Publishing, 2000.

[39] "An essay on chopin," http://www.geocities.com/Vienna/3495/essay1.html, Last visited 27 April 2005.

[40] C. Darwin, *The Expression of the Emotions In Man and Animals*, P. Ekman, Ed. New York, USA: Oxford University Press, 1998.

[41] W. James, "What is an emotion?" *Mind*, vol. 9, pp. 188–205, 1884.

[42] W. Wundt, *Outlines of Psychology*. London, UK: Wilhelm Engelmann, 1897.

[43] R. Lazarus, *Emotion and Adaptation*. New York, USA: Oxford University Press, 1991.

[44] W. Cannon, *Bodily Changes in Pain, Hunger, Fear and Rage*. New York, USA: Appleton, 1929.

[45] S. S. Tomkins, "The positive affects," *Affect, imagery, consciousness*, vol. 1, 1962.

[46] ——, "The role of facial response in the experience of emotion," *Journal of Personality and Social Psychology*, vol. 40, pp. 351–357, 1981.

[47] R. Plutchik, "Emotion: Theory, research, and experience," *Theories of emotion*, vol. 1, pp. 3–33, 1980.

[48] ——, *The Emotions*. University Press of America, 1991.

[49] C. E. Izard, *The face of emotion*. New York, USA: Appleton-Century-Crofts, 1971.

[50] ——, *Human Emotions*. New York, USA: Plenum Press, 1977.

[51] ——, "Differential emotions theory and the facial feedback hypothesis activation," *Journal of Personality and Social Psychology*, vol. 40, pp. 350–354, 1981.

[52] ——, "Four systems for emotion activation: Cognitive and noncognitive processes," *Psychological Review*, vol. 100, no. 1, pp. 68–90, 1993.

[53] P. Ekman, "Basic emotions," in *The Handbook of Cognition and Emotion*, T. Dalgleish and T. Power, Eds. Sussex, UK: John Wiley and Sons, Ltd., 1999, pp. 45–60.

[54] ——, *Darwin and facial expression: A century of research in review*. New York, USA: Academic, 1973.

[55] A. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, USA: Avon Books, 1994.

[56] ——, *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. London, UK: Vintage, 2000.

[57] ——, *Looking for Spinoza: Joy, Sorrow and the Feeling Brain*. New York, USA: Harcourt, 2003.

[58] P. N. Juslin and J. A. Sloboda, Eds., *Music and Emotion: Theory and Research*. Oxford University Press, 2001.

[59] E. Coutinho, H. Pereira, A. Carvalho, and A. Rosa, "Livia - life, interaction, virtuality, intelligence, art," Faculty of Engineering - University of Porto (Portugal), Tech. Rep., 2003, ecological Simulator Of Life.

[60] N. Gracias, H. Pereira, J. A. Lima, and A. Rosa, "Gaia: An artificial life environment for ecological systems simulation," in *Artificial Life V: Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*, C. Langton and T. Shimohara, Eds. MIT Press, 1996, pp. 124–134.

[61] E. Coutinho, E. R. Miranda, and A. Cangelosi, "Towards a model for embodied emotions," *Accepted to the Affective Computing Workshop at EPIA 2005*, 2005.

[62] J. S. Bach, *The Well-Tempered Clavier, Book I*, vol. BWV 846, ch. Prelude I (C major).

[63] E. R. Miranda, "Creative evolutionary systems," in *On the Origins and Evolution of Music in Virtual Worlds*, P. J. Bentley and D. W. Corne, Eds. Morgan Kaufmann, 2002, ch. 6, pp. 189–203.