Collaborating with the Behaving Machine: Simple Adaptive Dynamical Systems for Generative and Interactive Music

Alice Eldridge

Submitted for the degree of D. Phil.

University of Sussex

February, 2008

Declaration

I hereby declare that this thesis has not been submitted, either in the same or different form, to this or any other university for a degree.

Signature:

Acknowledgements

This is perhaps the smallest of many projects which was spawned by the energy and enthusiasm of Drew Gartland-Jones. I must thank him for support, encouragement and inspiration. To Phil Husbands I am indebted not only for a powerfully zen approach to supervision, but for being an inspirationally Renaissance man in modern academia. Thanks must go to Jon Bird for useful discussion and priceless support in times of need. Thanks also to Chris Thornton for lending his ear in the absence of Phil, and to Thor Magnusson and Peter Copley for discussion of less than conventional ideas. Beyond Sussex, I am indebted to the Live Algorithms for Music group which has provided a rare community for exchange and debate of matters core to this thesis and in particular to Ollie Bown whose intelligent enthusiasm encouraged me to pursue this project and continues to inspire its future moves. Finally I would like to thank members of blip, and all the people and places that have hosted performances of the work described here.

Preface

Much of the material in this thesis has been previously published and presented in various Journals, conferences, festivals and concerts.

In particular, parts of Chapter 6 have appeared in:

- Eldridge, A. (2002b). Adaptive systems music: Algorithmic process as musical form. In *Proceedings of the 2002 Generative Art Conference*, Milan.

Some of the ideas in Chapter 3 and models from Chapter 5 were presented in:

- Eldridge, A. (2005b). Extra-music(ologic)al algorithms for automated composition. In Rothlauf, F. et. al., editors, *Evomusart workshops 2005, LNCS 3449*, Heidelberg. Springer- Verlag Berlin.

The work presented in Chapter 8 has been previously published in:

- Eldridge, A. (2005c). Fond punctions: Generative processes in live improvised performance. In Edmonds, E., Brown, P., and Burraston, D., editors, *Proceedings of Generative Art Practice*, pages 41-51, Sydney, Australia. Creativity and Cognition, Creativity and Cognition Studios Press. andChapter
- Eldridge, A. (2005a). Cyborg dancing: generative systems for man-machine musical improvisation. In Innocent, T., editor, *Proceedings of third Iteration.*, pages 129-142, Mel- bourne, Australia. Centre for Electronic Media Art (CEMA).

Results from the study presented in Chapter 4 were discussed in a different context in:

- Eldridge, A. (2006). Issues in auditory display. *Artificial Life*, 12(2):259-274.

All of these papers should be considered as my own work. In addition, a summary of the ideas presented here are due for publication in a co-written book chapter:

- Husbands, P., Copley, P., Eldridge, A. and Mandelis, J. (2007). An introduction to evolutionary computing for musicians. In Miranda, E. R. and Biles, A., editors, *Evolutionary Computer Music*. Springer-Verlag.

Many of the works presented here have appeared in public exhibitions, festivals and gigs as follows:

AdSyMII (Chapter 6)

- July 2002: Blip5, Bar Sumo, Brighton, UK
- December 2002: Concert of Generative Art 2002, Politecnico di Milano University, Milan, Italy

Organised Entry (Chapter 6)

- October 2005: The Big Blip 05, Brighton Fringe Basement, Brighton, UK

Self-Karaoke Pond (Chapter 8)

- October 2005: The Big Blip 05, Brighton Fringe Basement, Brighton, UK

Ashby's Grandmother's Footsteps (Chapter 7)

- April 2006: Process Revealed at the Artpool Gallery, Budapest, Hungary (in association with the EvoMusArt workshop at EuroGP 2006.)

Performances of *Fond Punctions*, which uses the **Self-karaoke Machine** described in Chapter 8, include the following events and venues:

- July 2005: *Interactive Mind and Art(efacts)*, The Sussex Arts Club, Brighton, UK
- August 2005: The Great Escapade, Sussex, UK
- September 2005: Live Algorithms for Music meeting, The Great Hall, Goldsmiths college, London, UK
- October 2005: Wrong music, The Volks, Brighton, UK
- November 2005: University of Sussex Lunchtime Concert, Friends Meeting House, Brighton, UK
- November 2005: Third Iteration. Monash University, Melbourne, Australia
- December 2005: Generative Arts Practice. University of Technology, Sydney, Australia
- December 2005: Lan Franchis memorial discotheque, Sydney, Australia
- April 2006: Process Revealed, Artpool, Budapest, Hungary
- May 2006: Sunday Relay, The Albert, Brighton, UK

Compositions made with some of the systems here have been publicly played or distributed as follows:

- *Sines* (2002) a piece composed using the basic homeostat described in Chapter 5 was commissioned for the Lux OPEN 2002, Royal College of Art, London, UK
- *It Didn't Happen at Lan Franchis*, *Picket Fence Study 2: The Ant's ear view* and *Picket Fence Study 4: The Larvae's ear view* are due to appear on the cd accompanying *Evolutionary Computer Music*. (Miranda, E.R. and Biles, A.J. 2007). Springer-Verlag

Finally, the 'Behavioural Objects' project outlined in Chapter 9 was presented at:

- Improvising with Computers: IRCAM, Paris as part of NIME 2006

Collaborating with the Behaving Machine: Simple Adaptive Dynamical Systems for Generative and Interactive Music

Alice Eldridge

Summary

Situated at the intersection of interactive computer music and generative art, this thesis is inspired by research in Artificial Life and Autonomous Robotics and applies some of the principles and methods of these fields in a practical music context. As such the project points toward a paradigm for computer music research and performance which complements current mainstream approaches and develops upon existing creative applications of Artificial Life research.

Many artists have adopted engineering techniques from the field of Artificial Life research as they seem to support a richer interactive experience with computers than is often achieved in digital interactive art. Moreover, the low level aspects of life which the research programme aims to model are often evident in these artistic appropriations in the form of bizarre and abstract but curiously familiar digital forms that somehow, despite their silicon make-up, appear to accord with biological convention.

The initial aesthetic motivation for this project was very personal and stemmed from interests in adaptive systems and improvisation and a desire to unite the two. In simple terms, I wanted to invite these synthetic critters up on stage and play with them. There has been some similar research in the musical domain, but this has focused on a very small selection of specific models and techniques which have been predominantly applied as compositional tools rather than for use in live generative music. This thesis considers the advantages of the Alife approach for contemporary computer musicians and offers specific examples of simple adaptive systems as components for both compositional and performance tools.

These models have been implemented in a range of generative and interactive works which are described here. These include generative sound installations, interactive installations and a performance system for collaborative man-machine improvisation. Public response at exhibitions and concerts suggests that the approach taken here holds much promise.

Submitted for the degree of D. Phil.

University of Sussex

February, 2008

Contents

Track Listings

The attached DVD contains:

- Max/ MSP externals for the main algorithms described in Chapter 5. Also included are help files using mappings similar to those described in this chapter.
- A video of one performance made with the Self-karaoke Machine described in Chapter 8.
- Audio examples described in Chapters 5, 6, 7 and 8 are provided on the following tracks:

1	[5:1]	hom-perturb
2	[5:2]	hom-stabilise
3	[5:3]	hom-sines
4	[5:4]	planting-trees-excerpt
5	[5:5]	hom-sam-remix
6	[5:6]	hom-filter
7	[5:7]	hom-wrongbeats1
8	[5:8]	hom-wrongbeats2
9	[5:9]	nosc-change-weights
10	[5:10]	nosc-minima
11	[5:11]	nosc-change-inputs
12	[5:12]	nosc-perc-simple
13	[5:13]	nosc-perc-delta-tau
14	[5:14]	CA-chaotic
15	[5:15]	CA-complex
16	[5:16]	CA-ordered
17	[5:17]	GLV-inuit-pitch
18	[6:2]	AdSyMII
19	[6:3]	Organised Entry
20	[7:1]	Ashby's Grandmother's Footsteps
21	[8:1]	It Didn't Happen at Lan Franchis
22	[8:2]	Picket Fence Study One: The Blackbeetle's Ear View
23	[8:4]	Picket Fence Study Three: The Ant's Ear View
24	[8:3]	Picket Fence Study Two: The Larvae's Ear View
25	[8:5]	Self-karaoke Pond-harp

Glossary: Abbreviations and resources

1D - One Dimensional
2D - Two Dimensional
AI - Artificial Intelligence
ANN - Artificial Neural Network
CA - Cellular Automata
CTRNN - Continuous Time Recurrent Neural Network
DSP - Digital Signal Processing
EC - Evolutionary Computation
GA - Genetic Algorithm
GM - General MIDI
MIDI - Musical Instrument Digital Interface
OSC - OpenSound Control (see below)
USB - Universal Serial Bus

Arduino	is an open-source physical computing platform based on a simple i/o board, and a development environment for writing Arduino software. The Arduino programming language is an implementation of Wiring, itself built on Processing. `http://www.arduino.cc/en/`
ChucK	is a concurrent, strongly-timed audio programming language for real-time synthesis, composition, and performance, which runs on Mac OS X, Linux, and Windows. Code can be added, removed and modified on the fly, while the program is running making it an ideal language for live coding. It was originated by Perry Cook and Ge Wang of Princeton University. `http://chuck.cs.princeton.edu/`
Csound	is a text based music programming language written in the C programming language. A typical Csound program will include an *orchestra* file describing the nature of the instruments and a *score* file describing the parameters of the material (pitch, duration, amplitude etc). Csound then renders these files to produce an audio file or real-time audio stream. `http://www.csounds.com/`
Jitter	extends the Max/MSP programming environment to support realtime manipulation of video, 3D graphics and other data sets within a unified processing architecture. `http://www.cycling74.com`

Max/MSP	is a graphical development environment for music and multimedia. The program is highly modular and allows the development of third-party externals as objects which can be fully integrated with the native libraries. A typical Max programme, called a ‘patch’ is based on multiple graphical objects connected into a data flow. Control rate MIDI messages can be combined with a DSP network. Max was originally developed by Miller Puckette and is now developed and maintained by Cycling’74. `http://www.cycling74.com`
Processing	is an open source programming language and integrated development environment (IDE) built for the electronic arts and visual design communities. It builds on the graphical side of Java, simplifying some features and adding new ones. It is developed by Casey Reas and Ben Fry `http://www.processing.org`
Pure Data (Pd)	is a graphical programming language developed by Miller Puckette in the 1990s for the creation of interactive computer music and multimedia works. Though Puckette is the primary author of the software, Pd is an open source project and has a large developer base working on new extensions to the program. It is released under a license similar to the BSD license. `http://puredata.info/`
OpenSound Control	is a protocol for communication among computers, sound synthesisers and other multi-media devices. It is optimised for networking technology allowing very fast data sharing between machines. It can transport over many protocols but is commonly used with UDP or TCP/IP. It can be compared to MIDI, but does not suffer the same time lags and allows an open-ended url-style symbolic naming scheme. `http://www.cnmat.berkeley.edu/OpenSoundControl/`
SuperCollider	is a real time audio synthesis programming language. The Language combines the object oriented structure of Smalltalk and features from functional programming languages with a C programming language family syntax. Originating as proprietary software, it was released in 2002 by its author James McCartney under the free sofware GPL license. `http://www.audiosynth.com/`
Wiring	is a programming environment and electronics i/o board for exploring the electronic arts, tangible media, teaching and learning computer programming and prototyping with electronics. It is an open project initiated by Hernando Barragàn and builds on Processing. `http://wiring.org.co/`

List of Figures

Chapter 1

Introduction

> "In spite of being scientists, three gentleman consented to an experiment which must have seemed very strange at first sight, namely the marriage between music and the world's most potent machine ... With the help of an electronic brain the composer turns into an astronaut pressing buttons of his musical spaceship to introduce co-ordinates and keep the course of his vessel on its journey through constellations and galaxies of sound, controlling from his easy-chair what the imagination of yesteryear could have envisaged only remotely in its wildest dreams." - Xenakis (1971a) p.124–133

It is an exciting time to be a musician. The potency and portability of Xenakis' electronic brain has increased to the point where we no longer even need to press buttons in the cockpit to keep it on its journey: we can steer it through unimagined realms with remote gestures or sonic provocations; we can programme it to learn from and replicate the works of composers past and present; we can even programme it to to *lead* us around spaces beyond our *own* imaginations. And these new universes are not just of the sonic variety, but inhabited, explored and invented by practitioners from every corner of the arts allowing cross-fertilisation of ideas and techniques and opening possibilities for new practices. What is more, we are not alone in developing navigational strategies: our contemporary scientific colleagues offer a multitude of conceptual and technical know-how to be put to use on our expeditions.

Art has always been driven by an urge to explore, to create, to mimic and to come to terms with the world around us. In this respect the marriage of computers and music is not very strange, but entirely expected. Art, technology and science have always been locked in an intimate coevolution, the products of technological development continually fuel our creative endeavours, that in turn drive new technical innovation. At the same time the processes that support our creative outpourings have become hot topics in scientific research.

Throughout history the development of musical culture has been intimately influence by technical developments with new technologies constantly changing the ways in which music is composed, performed, preserved and distributed. Digital techniques have refined and expedited practice across the arts, simulating existing tools to make them more efficient, more flexible and easier to use: we can record, produce and master an album on a laptop, pitch shifting out some bad intonation; we can clean up photos, removing the blemish on the bride's nose and brightening the sky whilst undoing any errors; we can tween between keyframed poses of an animation, saving hours of time drawing transition frames by hand. But computers are not only good at replicating existing tools: they offer many new ways of working.

The development of creative software is inevitably tied closely to trends and advances in the computer sciences. In the last few decades there has been extensive research into applications of information processing, drawn in particular from Artificial Intelligence (AI) research, to the development of musical composition and interactive performance software. In fact it has been suggested that the development of systems that can create music in established styles is one of the major achievements of AI to date. In this area techniques of mutual interest to musicologists, cognitive scientists and composers have been developed through which computers can not only be imbued with musical know-how in the form of explcit instructions, but can be programmed to derive representations from existing musical works. Learning mechanisms and search strategies from AI have been deployed to develop programmes that can harmonise chorales, improvise bee-bop or extract patterns from existing works and generate convincing pastiches of the supplied material.

Similar techniques have also been developed to assist the contemporary composer's search for new musical idioms: new sounds, new textures, new means of expression. The picture of the composer or artist, set apart from worldly influence, fed by divine inspiration is common but mythical. We might like to uphold the Mozartian image of Composer as Visionary, but in reality most artists are fuelled in part by serendipity, gathering ideas that spring up in their active interaction with the tools of their trade. Composers for instance might sit and improvise open mindedly, saving ideas that arise for future development. Computers can not only help in this process, but offer new ways of systematically exploring new ideas.

Whilst the computer can only do what we tell it to, it can do so very quickly, and the results are not always necessarily something we could have predicted. For composers wishing to systematically expose themselves to novel ideas, or explore whimsical complexities, computers are not only a time saving device, but open up whole new realms of possibility. In wanting to step outside the confines of their own imagination, some composers have employed random processes - more or less sophisticated digital dice rolling techniques - to shake up the material of their practice. Other have implemented probabilistic frameworks, not dissimilar to Xenakis', defining broad fields of possibilities and employing the computer to generate the detail. Approaching composition in this way, the implicit process underlying the act of composition becomes reified in formal language: crudely put, arranging notes becomes designing processes.

The use of processes to induce unknown outcomes is nothing new. Most infamously perhaps advanced by the Experimentalists who dismissed fixed forms and prespecified structural relationships in favour of exploring ways of "outlining situations in which sounds may occur" (Nyman (1999) p.10). John Cage in particular is renowned for his use of various sources of chance to specify processes that bring about "acts the outcomes of which are unknown" (Nyman (1999) p.10). More recently Brian Eno has enthused about his fascination with inventing systems and machines that " . . . make music with material and processes I specified but in combinations and interactions I did not . . . (Eno (1996)). This interest in relinquishing control to a definite process with an *in*definite outcome characterises much current activity in the broad church of 'generative art'. Whilst the scope and rate of evolution of generative art practice defies any strict definition, Philip Galanter's description of the key elements of the approach is widely accepted:

> "Generative art refers to any art practice where the artist uses a system, such as a set of natural language rules, a computer program, a machine, or other procedural invention, which is set into motion with some degree of autonomy contributing to or resulting in a completed work of art." - Galanter (2003), p.4

On these terms, we can include the use of the computer to facilitate our specific artistic aims, as in Xenakis' use of the electronic brain to expound the sonic virtues of Gaussian Galaxies. We can include the use of the computational procedures to mass produce media, or to demolish the hierarchies of authorship. We could also include the adoption of AI techniques to mimic existing styles mentioned above. But for many generative artists the appeal is not just the ease of production, absence of repetition and never-ending something-else: its the promise of something *more*, something new and surprising.

The promise of excess and emergence of new ideas and new forms is one reason why many generative artists have adopted techniques from Artificial Life (Alife) research. In contrast to traditional AI, which concerns itself with developing representations of a highlevel central processor capable of human-level cognitive tasks, Alife concerns itself with studying the low level interactions of distributed processes from which coherent behaviour emerges as a global product of the system. One of the simplest and most famous examples is the set of cellular automata (CA) rules called the *Game of Life*, devised by John Conway (first described in print in Gardner (1970)).

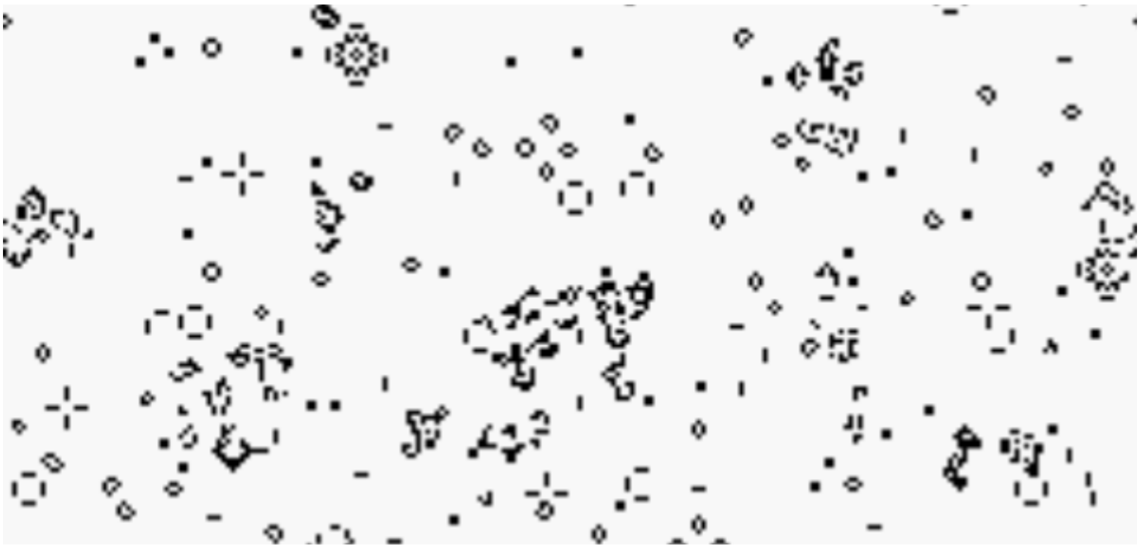


Figure 1.1: Screen shot from the Game of Life, a CA rule set devised by John Conway.

The Game of Life is a set of rules that specify the state of cells on an infinite two-dimensional grid i.e. whether they are 'dead' or 'alive'. In Figure 1.1, black dots are alive, dead cells are white. The state of each cell is determined by its eight immediate neighbours – those horizontally, vertically or diagonally adjacent. At each time step, the following set of rules gets applied to all the cells, setting their state for the next time step:

- Any live cell with fewer than two neighbours dies, as if by loneliness.
- Any live cell with more than three neighbours dies, as if by overcrowding.
- Any live cell with two or three neighbours lives, unchanged, to the next generation.
- Any dead cell with exactly three neighbours comes to life.

From a computer scientific point of view the significance of this rule set was that it had the power of a Universal Turing machine (i.e. anything that can be computed algorithmically can be computed with the Game of Life), but for physicists, biologist, economists, mathematicians, philosophers and Alife researchers, it represents a prime example of emergence and self-organisation.

According to the state of the cells at the start, different configurations and patterns emerge. Imaginatively named after things in the real world ('boats' and 'toads') or worlds of science fiction ('lightweight space ships', 'gliders 'and 'glider guns'), these are not only static forms, but coherent patterns of organisation that move across the grid. The rules describe what happens in each static cell, and yet what we see can only be sensibly described in terms of movement *across* cells. A verbal description of what is going on demands recourse to a formal level of description that is absent in the initial specification.

These systems are not alive in any sense of the word, but emulate operational aspects of biological convention without simply mimicking nature. There is something about their movement – their behaviour – that invites attributions of intentionality. Similar models have been used by visual artists to create forms that similarly display an uncanny agency. We have seen flocks of strange silicon bird-like forms or schools of abstract digital creatures, breeding and chasing one another. Robotic forms that are not explicit simulations, not merely automaton but substantially themselves.

Their sense of presence is due in part to their reactive nature. Many are defined in terms of local rules that are sensitive to their environment, creating digital creatures can be convincingly engineered to recoil at our touch, to follow us through space or hide in dark corners of their silicon worlds when we enter the room. Many visual artists have created virtual critters of varying degrees of abstraction whose movements and behaviours exhibit a degree of agency that belies their formally specified origins. The dynamic and flexible ways in which they respond to us (informed by sensors, video cameras and the like), has led researchers such as Ken Rinaldo (1998) to suggest that Alife techniques present "opportunities for both artists and viewer/participants to develop true relationships with the computer that go beyond the hackneyed replicable paths of "interactivity" which have thus far been presented by the arts community". For Rinaldo the time has come to indulge in a "cybernetic ballet of experience."

An impression of intentionality also arises from their often unpredictable nature. Although every aspect of a system is formally defined, the outcomes can at times be surprising, exceeding the expectation of the programmer-artist. In discussing their work many artists give the impression that they are pursuing a general interest in creating something that goes beyond what they specified: to achieve that "something more" (Whitelaw (2004)). Simon Penny writes, "I am charmed and fascinated by the possibilities of complexity theory and emergent order" (Whitelaw (2004) (p.216)). Robb Lovell expresses his interest in "going after creating something that gives me more than I expected" (ibid).

Similar techniques have been explored to an extent in musical applications. Many researchers have employed biologically inspired search mechanisms such as Genetic Algorithms (GAs), which harness the power of Darwinian evolution to systematically explore a defined, yet vast, space of musical possibilities. Others have used agent-based models of evolving ecosystems to create fluctuating populations of sound. CAs like the Game of Life and models of biological growth such as L-systems have also been extensively explored to generate abstract melodic paths or deployed in sound design tasks. But investigation has been limited to a small handful of models which have almost exclusively been applied in compositional tasks.

Rinaldo used the phrase 'Cybernetic Ballet of Experience' in the context of installation art. The vision behind the current project is to bring this sense of a digital 'other', that has made an appearance in the visual installation arts, into the sonic domain and onto the stage. The techniques that inject life into these visual creatures scurrying about on screens are crying out to be let loose on stage and offer an intriguing alternative to traditional approaches to interactive music. The origins of life and mind are far from solved, but the conceptual and technical tools used to tackle these epistemological issues offer a tool box which inspires a 21st Century self-steering upgrade of Xenakis' space ship. To borrow a turn of phrase from George Lewis, perhaps the time has come when we can programme computers that we not only can play our music on, but which will play their music with us (Lewis (2006)).

1.1 Summary of contributions

This thesis aims to expand upon the dominant design strategies for composing and performing with computers, introducing simple adaptive systems as mechanisms for both generating musical material and affording a novel approach to interaction with digital systems. Whilst many arts practitioners have appropriated Alife models, these invariably focus on agent-based simulations or EC methods, hankering after glamourous 'out of control' properties of emergence and self-organisation. Within music these have tended to remain in the domain of composition rather than performance.

A particular class of formal system is ear-marked as most suitable for this task. These are described as *simple adaptive systems* that are rich in both generative and interactive possibilities. This set of models adds to the compendium of tools available for computer musicians to play with. These tools are of both a practical variety and conceptual utensils which it is hoped will inspire the extension of the current collection.

The thesis is split into two halves. The first half provides a context for the current work and surveys and discusses work in the fields of Interactive Computer Music and Computer Automated Algorithmic Composition. The second half presents a practical exploration of the use of simple adaptive systems in different musical contexts from standalone generative music systems, to physical interactive installations, culminating in performance systems for man-machine improvisation. These practical applications represent initial, and minimal implementations of a fresh perspective on digital generative and interactive arts which is developed in the first half.

Chapter 2 examines the notion of interaction and considers how this is affected by the move from acoustic to digital instruments. It is suggested that the *active* nature of the digital medium makes new forms of interaction possible and inspires a model of interaction that is different to that associated with acoustic instruments. A conversational metaphor is introduced highlighting the mutual influence between performer and instrument in contrast to the one-way control we have over standard acoustic instruments.

In order to provide a conceptual framework for understanding these differences, heuristic concepts drawn from behavioural robotics and the philosophy are introduced. Specifically the notion of autonomy is explored and an operational definition raised as a useful conceptual and practical tool for the development of interactive systems. Conceptually this is useful as it provides a perspective for appreciating how a software system can be at once independant from, yet sensitive to its environment. Practically an operational definition outlines the type of system architecture that can realise this, albeit minimally.

With this in mind, Chapter 3 focuses on existing approaches to the generation of musical material and proposes that simple adaptive systems offer attractive features as generative composition tools. Included in this chapter is a gentle reminder to practitioners of algorithmic composition that music is not a natural kind, and the relationship between music as experienced and music theory is not the same as that of theory and phenomenon in the physical world. In this respect a mild warning is raised that music theoretic tenants may not necessarily be a suitable basis for designing mechanisms for generating music.

An experimental offering is made in Chapter 4, with the results of an experimental psychology style study that was designed to investigate whether formal properties of complex systems, which are evident in graphical representations, could be similarly appreciated from an auditory display. The results showed that at least in the current conditions, the highlevel properties of complex systems can be appreciated in audio. This is an assumption which is made by practitioners who employ extra-musical algorithms in music, but has not been experimentally verified in this context in the past.

The practice based investigations of Chapters 5, 6, 7 and 8 contribute finished works to the broad field of digital generative arts in offering a balance between the 'out of control' Alife systems explored most extensively in visual and installation arts, and the use of mathematical models in algorithmic composition. Specific contributions are made by each of the practical investigations to the distinguishable practices of generative art, algorithmic composition, installation arts and interactive computer music performance within which they sit.

The studies presented in Chapter 5 contribute to extra-musical algorithmic composition by bringing a new set of workable tools. Chapter 6 presents two generative music installations which use these components in combinations, and explore the use of multiple mappings from model to sound. Chapter 7 opens these systems up to the real world and investigates both the use of simple adaptive systems in responsive environments and a cybernetic take on providing feedback.

The Self-karaoke Machine presented as both an installation and performance system in Chapter 8 speaks to both the Alife generative art world and interactive computer music players. To the interactive music world, it offers an alternative approach to man-machine collaboration that allows full expression of both generative and traditional forms of improvisation. To the Alife artists preoccupied with emergence and user's creative freedom, it demonstrates the power of bringing the user into the generative loop. In coupling the formal system with the open-ended dynamical system that is the performer's sonic improvisations, possibilities escape the programmer's intentions and open-ended exploration can occur.

The body of work has acted as a spring board for the creation of *behavioural objects*, a project funded by the EPSRC network Live Algorithms for Music, which aims at to develop tools and an understanding of adaptive systems in improvised computer music. This project promises to gather enough momentum to fulfil my main wished-for contribution: that this work may in some way inspire future explorations of the use of adaptive digital processes in conjunction with acoustic instruments in live performance, bringing together the boundless possibilities of computer simulation with the wonderfully productive constraints of performing live music with acoustic instruments.

Chapter 2

All Computers are Interactive: But Some Are More Interactive Than Others.

> "One must take computers into account, and take them to task, because it is a response to the technology of our time, to the situation of our time. To make music with the technology of our time, and specifically the computer, poses a tremendous challenge for the artist. To address this challenge, in itself, will help keep music alive and significant. To address this challenge in a way that acknowledges, directly and deeply, the human production of that music brings together into a new art form the diverse elements of performance, with its millennia of history, and the age of the computer, with its bare decades of history. . . . interactive computer music takes the fullest advantage of the ideas and technologies of today and unites them with a vision of what they could be." - Garnett (2001), p.31

As outlined in the introductory chapter, the ultimate practical aim of this body of work is to develop upon existing digital music systems which bring us the sense that we are "playing with", rather than "playing on" our computers: to be able to invite the laptop onto the stage and pursue a man-machine collaborative improvisation. Acoustic instrumentalists might talk about their relationship with their instrument as a form of collaboration, but digital instruments have the potential to play a more literally *active* role. As a starting point then, this chapter provides an overview of current approaches to interactive computer music and questions how digital instruments of various kinds have affected our understanding of interactivity itself.

Interactivity has been a buzz word in New Media Arts for decades. Nearly 30 years ago Kay and Goldberg (1977) recognised the significance of the active nature of the computer as an artistic medium. Since then researchers in the field of Human Computer Interaction (HCI) have been exploring the ways in which we interface with digital technology, artists have been developing interactive artworks and installations and more recently critics of New Media Art have begun to construct a framework for appreciating the aesthetic implications of interactive art. Indeed there has been so much activity under the rubric 'interactivity' that Manovich (2002), suggested that the concept has become "too broad to be truly useful" (p.55).

In some respects, the impact of a new dynamic, active medium was less alien to musicians than to other artists working in static visual or plastic arts: music is quintessentially temporal, and instruments inherently interactive. In fact some critics have offered the musical instrument as a model of interaction for the rest of New Media Arts. Although unqualified in this case the term 'instrument' is not a simple concept. Understanding the impact of technological advances on the nature and role of the instrument is central to appreciating the evolution of interaction.

Long before real-time interaction became a possibility, the musical potential of digital technology had been exploited, creating sound worlds which would have made the futurists weep. (e.g. Chowning (1973), Xenakis (1971b), Stockhausen (1964)). For these pioneers, their instrument was the studio. Meanwhile the use of live electronics in performance by Cage and other experimentalists, notably Gordon Mumma began to alter the traditional roles of instrument, performer and composer in performance. Even before digital technology entered the musician's world the term 'instrument' no longer referred solely to a passive device requiring a manual action to create each sound. Instruments now encompassed explicit temporal structure which had previously been the sole reserve of composition.

Composers such as David Tudor and Gordon Mumma have used the term 'composed instrument'. The term has also discussed by Schnell and Battier (2002), to highlight the fact that computer systems used in musical performance "carry as much the notion of an instrument as that of a score" (p.1). Computers can be used to predetermine aspects of a musical work as much as they can be used to realise it in performance. The concept is equally applicable to electronic instruments by virtue of the fact thay they are dematerialised. There is no longer a fixed or direct correspondence between the interface and sound production mechanism. Electronic and digital technologies make it possible to conceive of a sound producing device which is independant of its gestural control.

To give a simple example, a piano responds with a single note when a single key is pressed: there is a one-to-one correspondence between the player's action and the sonic output. We could say that there is an isomporphism between the gesture and the rhythmic and pitch content of the emitted sound. With wind or string instruments, a similar isomorphism exists, and we are in addition aware of a more obvious match in timbral terms. For example the speed and depth of hand or diaphragm movement corresponds directly to the rate and depth of vibrato in a string or wind instrument. On an analogue or digital synth, we are most likely to be presented with a physical or graphical knob, slider, or number box which we adjust in a single movement to create an ongoing change in depth and/or speed of modulation. The differences become more obvious when we consider a simple harmoniser or arpeggiator, or a Max/MSP patch loaded with a rhythmic sample - one simple gesture (a button press) unleashes a potentially endless stream of structured musical material.

The decoupling of interface and sound producing mechanism has great impact from the performer's perspective and, as will be presented in Section 2.1.2, the development of controllers that facilitate expressive control of the sonic output has become a major topic of research. The independence of gesture and sound production also means that physical gestures are not tied to particular sonic gestures but can be re-mapped at will. This is achieved in modular analogue synthesisers for example, by re-patching modules with physical patch cords. This was a trick explored by early electronic composers. Morton Subotnick, for example, applied an envelope follower to his vocal utterances, producing voltages changes which were then used to control a Buchla analog synthesizer (Winkler (2001)). This process was developed into one of the early multimedia operas *Ascent into Air* (1983).

If electronic instruments began to blur the boundaries between instrument and composition, and patch-able electronics demonstrated the effects of de-coupling the interface and sound engine, then the unique aspect of the *programmable* digital instrument goes further and facilitates not only 'composed' instruments with structure, but active, responsive properties that we would traditionally associate with the performer. The aim of this chapter is to examine how these properties have affected the model of interaction that underlies the design of digital instruments, and how this could be developed and realised.

The first section of the chapter outlines and illustrates the principle approaches to interactive computer music systems, bringing attention to the underlying metaphors on which models are based. The model of one-way interaction with an acoustic instrument is contrasted with an alternative two-way, conversation model that has been proposed by those who feel that one-way reactivity does not exploit the full potential of the digital medium. In search of some examples of this conversational form, a brief survey of some examples from interactive Alife-inspired installation art is given in Section 2.2. Section 2.3 begins by contextualising the one-way and conversation models of interaction in terms of the design protocols that subserve them. This is developed further by introducing the philosophical concepts of autonomy and heteronomy which illuminate the implications of adopting either model. In turn, these perspectives are allied with two contrasting computer science paradigms which have quite different methodologies. Consideration of these philosophical perspectives and methodological issues helps in developing a more solid understanding of these metaphors of interaction, but more importantly offers suggesting a set of conceptual and practical tools to aid in the implementation of interactive performance software that begins to realise this model. These are considered in Section 2.4.

2.1 Interaction in Live Computer Music

Almost all newly observed phenomenon are initially described using combinations of concepts drawn from phenomenon that we already understand. This is as true of new technologies as it is of natural world phenomenon: shortly after the invention of the telephone, people discussed the possibilities of broadcasting concerts directly into homes; the first motion pictures were shown in theatres as a backdrop to the actors on stage. Rapidly of course, the telephone and film came to be understood on their own terms and even provided metaphors for the next generation of technology or scientific understanding. Arguably the modern personal computer has not yet reached a similar level of maturity. This is well illustrated within the interactive computer music community, where digital systems are commonly conceptualised in terms of the elements of classical musical culture: virtual instruments, virtual performers, virtual composers, virtual improvisers, virtual listeners, virtual critics (Winkler (2001)).

Variations on these positions can be seen implicitly in various researcher's definitions of interactive music. Robert Rowe (1992) seems to have some kind of dynamic instrument in mind when he writes that interactive music systems are "those whose behaviour changes in response to musical input" (p.1). Other discussions reflect the changing musical roles associated with interactive music: "Interaction has two aspects: either the performer's actions affect the computer's output, or the computer's actions affect the performer's output." (Garnett (2001), p.23). Increasingly, however, there is a sense that this 'one-way reactivity' is not enough, that interactive systems should strive to amalgamate the characteristics of all of music's tools and personnel – the sound of an instrument, ears of an instrumentalist and mind of a composer – such that the musical flow is mutually influential. Winkler hints at this in the opening of his book with a conversational analogy:

> "*Interaction is a two way street*. Nothing is more interactive than a good conversation: two people sharing words and thoughts, both parties engaged. Ideas seem to fly. One thought spontaneously affects the next." - Winkler (2001), p.4

More recently Bert Bongers (2006) has extended this metaphor, stressing the point that interaction should involve a mutual influence which causes both partners in the discourse (whether machine or human) to have changed state, frame of mind or views after the interaction. The reciprocal nature of conversation captures the sense of interaction which we observe in fluid musical situations between two or more human performers, or indeed between any entities in the living world.

2.1.1 Metaphors from Traditional Performance Practice

In one of the earliest texts on interactive music systems, Robert Rowe (1992) proposed a classification of interactive computer music systems along three principle axes. Taking the central concepts of classical western performance practice, he made the distinction between composed and improvised (score-driven versus performance-driven), methods of composition (*transformative*, *generative* or *sequenced*) and between what he calls *instrument* or *player* paradigms.

Todd Winkler (2001) later expressed the need to establish new models of interaction in order to inspire the development of digital performance systems. Attempting to answer key questions such as *What role does the computer play?* and *What is the relationship between computers and humans?*, he suggested that consideration of the interactive relationships which occur in traditional performance ensembles may be a useful starting point to evolve “new modes of thought based on the computer’s unique capabilities” (p.21). Focusing on the issue of control (who is in charge, who follows, who leads?) Winkler offers four models based on different types of ensemble and their associated musical idioms.

The Conductor Model (a la Symphony orchestra) describes the situation where the performer acts as conductor, influencing the computer’s delivery of a pre-scored part. The earliest interactive music system GROOVE[1], developed at the Bell Labs, operates within this framework. The performer adopts the role of conductor, controlling the tempo, dynamic level and balance of a computer programmed with a pre-written score. This model epitomises a common approach to interactive computer music research known as *score-following* which is described in more detail in Section 2.1.3.

The Chamber Music Model (a la String quartet) proffers a richer model of interaction, alluding to the mutual influences between several players. “In a string quartet . . . the interplay between musicians demonstrates shared control. Intonation, phrasing and tempo are constantly in flux” (p.25). As an example of this model in interactive composition, Winkler offers the first movement of his *Snake Charmer* for clarinet and computer. The interplay here is achieved by explicitly switching control. The piece opens with a computer introduction set at a fixed dynamic level. When the clarinet enters, it is able to influence the dynamic level of the computer line for a short time, after which, the computer stops listening to the performer and continues on its own. Control switches several times during the performance, with the computer exerting most obvious influence over the clarinettist in its occasional outputs, which demand the player increases his dynamic level to match the computer.

The Improvisation model (a la Jazz Combo) outlines a more complex model of interaction, reflecting the fact that in traditional Jazz combos, not only do musicians mutually influence the interpretation of the head (i.e. the scored motifs), but commonly whip each other up into frenzied improvised solos. Winkler comments that “what makes these relationships function to produce music that does not sound like a random babbling is that there are a huge number of shared assumptions and implied rules based on years of collective experience” (p.26). This musical intelligence, as he calls it, is typically simulated with software, which contains both sets of analysis tools to recognise patterns in rhythm, melody and harmony, and generative components consisting of coded sets of rules and assumptions which respond according to the outputs of the analysis module.

---

[1] Generating Realtime Operations On Voltage-controlled Equipment

Free improvisation (a la AACM, Down town NYC, European free improvisers) is deemed to be the most challenging and complex model for interactive music. The spontaneous, expressive and unpredictable nature of free improvisation, evokes a much richer notion of interaction and demands software which allows more freedom. Such systems typically combine analysis components with a processing engine which has considerable generative power. These engines draw from a variety of AI methods such as GAs, markov modelling, directed graphs, rule based systems as well as 'extra musical' models, all of which will be discussed in Chapter 3. The aim is not only to turn the computer into an instrument you can play, but also to achieve a sense that the instrument it playing with you. This model begins to capture the sense of mutual influence central to the conversation model.

By drawing on our understanding of interactions between players in traditional ensembles, these models help to focus attention on the direction of influence and control. This is an issue that has become an intense area of discussion in other electronic arts with the introduction of dynamic and interactive media (Bongers (2006)). However, over-emphasis on traditional models and musical frameworks brings with it the danger of getting stuck in past paradigms, using new technology to parody old practices, rather than exploring new possibilities. Winkler himself recognises the limitations of using metaphors from traditional performance practice as models for interactive computer music, suggesting that "simulation of real-world models is only a stepping stone for original designs idiomatic to the digital medium" (Winkler (2001) p.23).

Since this time, there has been considerable research effort focusing on real-time analysis tools such as pitch detection, beat induction, phrase segmentation etc. which allow us to design software capable of tracking the sonic gestures of a separate performer. There is also a huge community of researchers exploring novel possibilities for gestural control, both in terms of hardware interfaces and video analysis of bodily movements, in order to explore the potentials of the computer itself as an instrument. Both of these research efforts concentrate on creating a front-end, an means of interfacing with a digital system rather than the model of interaction itself. Whilst the tradition of interaction in music may seem to provide it with a head start in comparison with other art forms, the apparent similarities between the interactive nature of traditional instruments and performance networks, and those afforded by the computer, mean that comparatively less time has been spent considering how the interactive possibilities of the computer differ from traditional performance practices.

This section presents a survey of some of the current approaches and issues in interactive computer music practice. Rowe's original conception of instrument verus player paradigm may be too polar in the current climate, but approaches can be usefully considered along a continuum between these two extremes. This section explores the nature of interaction in a variety of projects from the design of *New Musical Instruments*, through *responsive accompanists* and *virtual improvisers* to *audible ecosystems*. Comprehensive reviews and histories are available in Roads (1996), Chapters 14 and 15, Dean (2003), Impett (2001), Jorda (2002), Rowe (1992) and Rowe (2001)

2.1.2 New Musical Instruments

One of the major research efforts in interactive computer music focuses explicitly on the computer as an instrument in its own right, or as an extension to existing acoustic instruments. The term *Hyperinstrument* refers to a software augmentation of an existing acoustic instrument, whereas *Virtual Instrument*, in the broadest sense, refers to any software-based sound-making system that takes user input. Developing powerful interfaces has become a major field of investigation and research in itself, with institutes like

STEIM[2] and conferences such as New Interfaces for Musical Expression (NIME [3]) dedicated to the development and deployment of new musical instruments and interfaces. **Hyperinstruments**, as the name suggests, aim to augment the musical possibilities of acoustic instruments whilst preserving their expressive potential. In some respects the approach represents a digital exploration akin to the development of extended playing techniques, or the practice of 'preparing' acoustic instruments: one aim is to broaden the possible sound world. Beside the obvious extension of sonic possibilities, digitally extended instruments leave open the possibility to retain the original acoustic output of the instrument, creating space to explore the interplay between the acoustic and processed sounds. The interface between the acoustic and digital instruments can be purely sonic, either employing real-time analysis tools to capture key aspects of performance gesture, or directly treating the acoustic signal. Alternatively, physical interfaces can be implemented by adding sensors and switches to the physical instrument as in the *Hyperbow* developed by Young (2001), Jonathon Impett's *Hypertrumpet* or Ernest Rombout's *Electronic Piccolo Heckelphone*. The best known work in this area has been carried out by the Music and Cognition group of the MIT Media Lab.

> "Our approach emphasises the concept of 'instrument', and pays close attention to the learnability, perfectibility, and repeatability of refined playing technique, as well as the conceptual simplicity of performing models in an attempt to optimise the learning curve for professional musicians." - Machover and Chung (1994) p.186

Many of the early Hyperinstruments developed at MIT were typically used in notated compositions where composers could pay close attention to developing the interaction between acoustic instrumental lines and live electronics. Increasingly however, players are commissioning their own controllers. Performers such as Jonathon Impett and Ernest Rombout demonstrate the new possibilities for improvisation that emerge when a virtuosic player has control over an acoustic instrument *and* its manipulations. Where acoustic instruments are used as the principle controllers, the traditional sense of interaction with an instrument remains essentially unchanged. However, this approach develops upon the sort of interaction we talk about between musical parts themselves. No longer is this the off-line domain of the composer, but the performer now has control via one, albeit augmented, instrument over an indefinite number of musical parts.

Perhaps the most familiar **virtual instruments** are those found in commercial software such as the Virtual Instrument plugins available for sequencers such as Cubase and Logic. These are typically software simulations of familiar analogue instruments such as synthesisers and samplers. But digital instruments are by no means restricted to simulation of existing analogue systems, and experimental research in this area extends far out into the art world: musical (and audio-visual) interfaces also appearing in exhibition contexts at festivals such as Ars Electronica.

That the design of technological interfaces has become worthy of critical appraisal by artists is perhaps testimony to the fact that this represents one of the major facets of virtual instrument design. If expanding the sound world was top of the agenda for early computer music, current research arguably focuses most heavily on approaches to constraining, controlling and expressively exploring the expanded horizons. One of the major problems facing designers and users of virtual instruments is their virtual nature itself, i.e. their lack of physical interface.

---

[2] `http://www.steim.org/`

[3] `http://www.nime.org/`

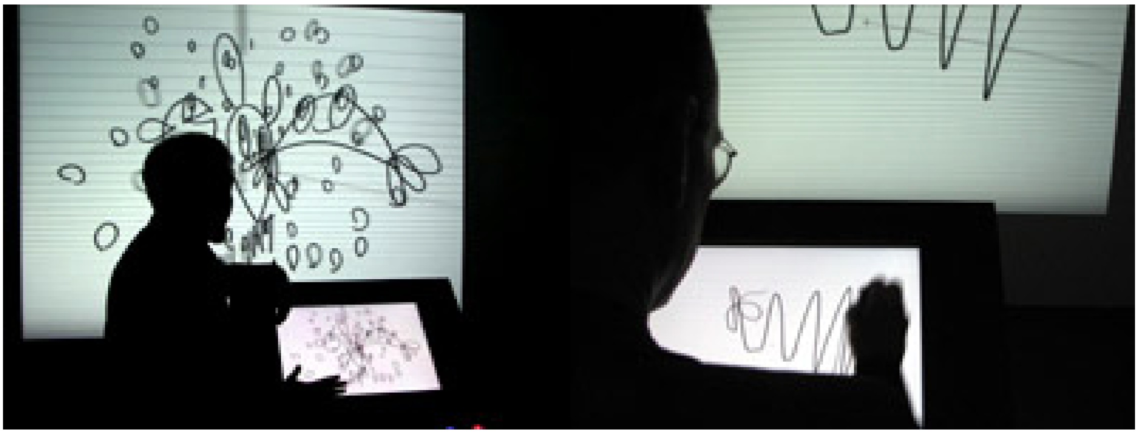


Figure 2.1: Sonic Wire sculpture by Amit Pitaru, an AV interface exhibited at Ars Electronica 2005. (photo: Thomas Petersen)

The physical form of acoustic instruments governs both the nature of the sound they can produce and the physical skills which must be developed in order to shape the sound. The minimal nature of most virtual instrument interfaces can evoke frustrations in the audience as well as the performer. The performance gestures associated with the physical existence of acoustic instruments provide not only sonic character and idiosyncrasies, but a visual correlate that is an integral part of performance from the audience's perspective. For many audiences (the eye-closed classical concert-goer being the exception), *seeing* the trombonist deftly alter his slide position, observing the finegrained coordination between the fingers of a violinist's left hand and right arm or watching the singer's bosom rise, head kick back and jaw drop are almost as important as hearing their sonic results. A look of studied concentration illuminated by the glow of a laptop screen just isn't the same.

For the performer the situation is perhaps even more severe. The physical form of many traditional instruments has evolved over hundreds of years to provide a very high number of degrees of freedom, each with fine-grained continuous control. The instrument is often the locus of integration of physical movements throughout the entire motor-system. Such interfaces take time to develop. I am not suggesting that the emulation of this physical congruence is a good model for digital instruments. As noted above, the decoupling of the interface and sound engine is a characteristic unique to digital and electronic instruments, and the dynamic reconfiguration of these connections is one of their powerful features which deserves exploration. Under such flexible conditions even if a comparably expressive physical interface was developed, the hours of mechanical practice endured by acoustic instrumentalists would be irrelevant: learned sensory-motor contingencies are powerless in such a dynamical setting. This represents an interesting line for future research.

There are a range of approaches that focus on the issues of control and expression which are relevant to both performer and audience. These are *hardware interfaces*, *gestural sensors*, *software interfaces* and *live coding*.

Much work is being done in the development of new *hardware Interfaces*, physical controllers which are used to control underlying Digital Signal Processing (DSP) engines. Many reappropriate existing instrument controllers, such as the MIDI guitar. Others explore idiosynchrasies of the digital or electronic medium such as Michel Waisivic's *Crackle Box*. Some also aim to capture expressive movements made by parts of the body not normally deployed in instrumental performance. Todd Machover's group at MIT for example developed the Sensor Laden Dancing Shoes (Paradiso and Hu (1999)). These shoes are fitted with sensors that send sixteen streams of control data relating to elevation, acceleration, orientation and pressure etc. Data gloves, made popular in the early days of Virtual Reality experiments have been extensively explored, most famously perhaps by Laetitia Sonami with her 'Lady's Glove' [4] (shown in Figure 2.2), the more adventurous or athletic deploying similar technology in full body suits. In recent years, more sophisticated sensors are being developed which provide haptic feedback such as Bert Bonger's force feedback and vibrotactile tools (Bongers (2006)).

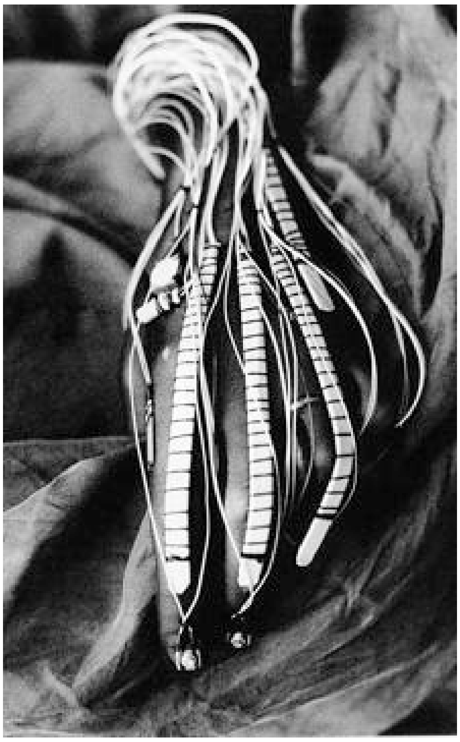


Figure 2.2: Laetitia Sonami's Lady's Glove

Gestural controllers are not only used in performance but also in the studio, one of the earliest examples is Paul de Marinis' use of a data glove to control a voice synthesiser in *Power of Suggestion*[5]

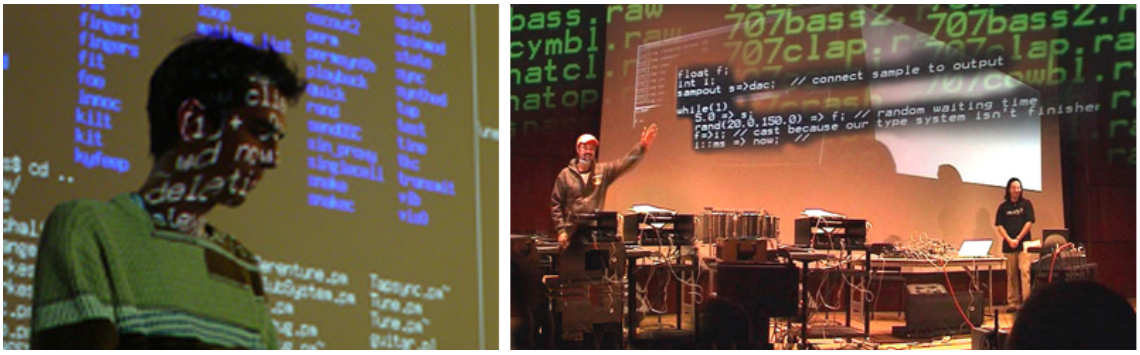
Gestural sensors, such as video analysis, infrared, or ultrasonics bypass the hardware interface and track the physical movements of performers, transforming aspects of their gestures directly into control signals. These might be continuous and high resolution as in the heterodyning oscillators of the theremin, which demands a refined technique comparable to that of a playing a string instrument. In other situations discrete sets of switches or longer distance sensors are used, triggered by larger movements such as that of dancers. In both areas physical movement is transduced with the aim of maximising the performer's expressive control and invariably creating observable correspondences between what the audience can see and hear.

Software interfaces, which add an extra layer of control between mouse-keyboard-screen and DSP engine, aim to put back some of the structural restraints of physical instruments. Interesting examples of work in this area comes from the ixi-software group who have developed a series of graphical front-ends to sound engines such as Pure Data or SuperCollider.

---

[4] `http://www.sonami.net/lady_glove2.htm`

[5] This is a track on de Marinis album *Music as a Second Language*. 1993. (Lovely Music).



```
float f;
int i;
sampout s=>dac;  // connect sample to output

while(1)
    5.0 => s;
    rand(20.0,150.0) => f; // random waiting time
    f=>i; // cast because our type system isn't finishe
    i::ms => now;  //
```

Figure 2.3: Alex McClean (left) and Perry Cook and Ge Wang (right) live coding

Finally *live coding* renounces all attempts to regain the physical, but instead embraces the power of dynamic interpreted languages, such as SuperCollider or ChucK, to write code 'on-the-fly' in performance situations. Rather than attempting to engage the audience with physical gestures, deftly typed lines of code are projected on a screen behind the laptopist, giving the audience an insight into the processes that create what they hear.

A major focus in this area then is in creating flexible and expressive interfaces between man and machine. The computer is employed as a powerful effects unit or DSP engine under the guidance of a human pilot: the main focus for interactive research parallels the concerns of the HCI community, i.e. the efficiency and effectiveness of the man-machine interface.

2.1.3 Responsive Accompanists

Another major aim of early research into interactive techniques was the exploration of methods that allowed a musican real-time control over the tempo of a predetermined computer music score (Winkler (2001) and see Dannenberg (1989) for review of early approaches). In this model, the musician acts as conductor (as in Winkler's conductor model presented in Section 2.1.1) defining the tempo of a programmed score. The interface may be made physically via a hardware controller or sonically using beat-tracking algorithms to follow the musician's performance. Classic examples of this score-following approach include Max Mathews' *Radio Baton*, which enables a musician to control tempo and other aspects of a score stored in computer memory, and Todd Machover's *Bug Mudra*, in which the conductor wears a specially designed glove to control the reverb, panning and mix of the digital score whilst simultaneously conducting three performers (Machover (1992)).

In more recent years there has been an at once more pragmatic and more sophisticated application of these techniques. Christopher Raphael's *Music Plus One* (MPO) (Raphael (1999)) is an 'intelligent' version of the more familiar *Music Minus One* system which provides recorded accompaniments for musicians to rehearse concertos and sonatas. In line with the standard score following approach, representations of both the solo part and the score are input into the system. A Listening process (based on a hidden markov model (HMM)) and Play process (which utilises a Baysian belief network) run concurrently to track and respond appropriately to nuances in the perfomer's interpretation.

> "Specifically, my goals are that the program must respond in real time to the soloist's tempo changes and expressive gestures ... In this way MPO *adds* to the soloist's experience by providing a responsive and nuanced accompaniment rather than *subtracting* from it by imposing a rigid framework that stifles musical expression." - Raphael (2004)

Some of these features are available in the *Suivi* Max/MSP patches developed at IRCAM[6] (Orio et al. (2003)) which perform score following from audio input using two levels of HMM. One model tracks low level spectral features such as attack, sustain and decay of individual notes; the higher level implements transitions according to notes in the score.

This approach combines the active qualities of digital media with an automated instrument. Essentially we have a pianola which can adjust to the performer it is accompanying. The real development in this approach lies in the analyses devices necessary for the machine to track the performer's progress and respond accordingly. Interaction here then, also focuses on the interface between man and machine, where the interface itself is active.

2.1.4 Virtual Improvisers

At the far end of Rowe's instrument-performer spectrum, sit the 'virtual improvisers'. The aim in this area is to construct an 'artificial player', with a musical presence and personality of its own. It is in this area of interactive computer music then, that the type of 'mutual influence' discussed by Bongers and others is most explicit.

In an attempt to preserve a global musical context, some researchers take what Rowe describes as a transformative approach, modifying either the human player's improvisations directly, or transforming a predetermined database of musical fragments according to analyses of player's input. This approach underlines the current incarnation of Al Biles' *GenJam* (Biles (2002)) which improvises within a traditional Jazz framework by selecting and transforming phrases from a ready made database of Jazz licks or mimicing or modifying the human performer's phrases.

Presented as a practice or didactic tool, Francois Pachet's *Continuator* (Pachet (2002)) uses Markov techniques to build databases of sub-sequences that enable the creation of responses derived from the performer's improvisations. Pachet's system operates in real time to capture key structures of a performer's musical statements. The system operates in both 'Autarcy' mode where it progressively catalogues the input of the current performer, or 'Virtual Duo' mode, where a database built from another musician is used as the transition matrices for the system. The system's performance is impressive, capturing the idiosynchratic harmonic and gestural moves of professional Jazz pianists, and can also act as a structured learning environment for musical novices.

Others aim to increase the independence of the computer system by creating a generative engine that may be influenced by the performer whilst maintaining a certain independence. The best known, longest running (and least disclosed) project in this area is George Lewis' *Voyager* system which is designed to perform improvisations with a human instrumentalist. He regards the computer as "just another musician in the band" (Lewis (1999)), or more specifically "a multi-instrumental player with its own instrument" (p.103). The 'players in the orchestra' are controlled by global behaviour specifications, which are influenced by analysis of pitch and velocity data taken from the player's improvisations. The generative behaviour of the system is developed from white noise, which is shaped and filtered with a series of stochastic rule sets. The generative engine produces musical output regardless of whether or not a human performer is playing. When sonic input *is* detected, feature analysis of many aspects of pitch and velocity data is used to represent the state of the input at a given moment. These in turn influence the behavioural specifications, altering the musical behaviour of the system – or not.

Lewis' move away from the fixed idioms of many systems reflects his view of improvisation per say: "Musical improvisation is one domain among the various possible do-

---

[6] L'Institut de Recherche et Coordination Acoustique/Musique, a computer music research centre set up by Boulez in Paris.

mains of improvisation – an interaction within a multi-dimensional environment, where structure and meaning arise from the analysis, generation, manipulation and transformation of sonic symbols" (p.101).

This dedication to the independence of a digital generative process and a 'bottom up' conception of improvisation is explicit in Tim Blackwell's approach as embodied in his various *Swarm* systems (e.g. Blackwell (2003), Blackwell (2004)). On the premise that structure in free improvisation arises spontaneously from the interactions between players, Blackwell employs particle models of swarming phenomenon. In its simplest incarnation, the swarm is based on Craig Reynold's (1987) *boids* algorithm.

The boids algorithm is often used to illustrate the process of self-organisation, demonstrating how coordinated global behaviour can arise out of simple local interactions without the need for any supervisory control. The basic flocking model consists of three simple rules. These determine the movements of individual particles, or boids, according to the positions and velocities of their neighbours. The rules consider three factors: *separation*, which ensures that individuals do not bump into each other; *alignment*, which causes them to adopt the average heading of their nearby flockmates and *cohesion*, which causes them to move toward the average position of their local flockmates. Rules like this are sufficient to cause co-ordinated flocking in an initially randomly distributed cloud of particles.

Blackwell implements a similar system in an N-dimensional space which is then mapped into musical dimensions. In early incarnations the axes of 'music space' were loudness, pulse and pitch (Blackwell (2003)), later versions used the swarms to parameterise a granular synthesis engine (Blackwell (2004)). Interaction takes place by analysing the performer's input in the same dimensions of musical space in which swarms exist. Musical events in the outside world then become 'targets' in the swarm space to which the swarms are attracted.

Blackwell insists that interactivity, or 'strong interactivity' as he calls it depends upon 'instigation and surprise as well as response' (Blackwell (2006)). Using this swarming model, he creates a system which generates behaviours internally, giving a sense of independence from the performer and moving away from the one-way model of interaction. Lewis similarly eschews the instrument metaphor with its attendant notions of control, taking pains to stress that "the computer system is not an instrument, and therefore cannot be controlled by a performer . . . The computer's own musical behavior is the product of its own initiatives, and its response to outside input when the program has determined that such input is present." (Lewis (1999)).

2.1.5 Audible Ecosystems

Suggesting that the vast majority of computer music systems which are described as interactive should more properly be thought of as 'non-reciprocating reactions', Italian composer and sound artist Agostino Di Scipio submits a more systemic approach in what he describes as *ecosystemic signal processing* (Di Scipio (2003)). In projects such as *Audible Eco-systems Interfaces* (AESI) he broadens the traditional network of performance components from just humans and computers to explicitly include the performance space itself. The traditional roles of each are often inverted as a circular relationship between human performer(s), machine(s) and the surrounding environment is implemented (Anderson (2005) p.16).

Di Scipio's determination to express an alternative to the one-way model, which he sees as an expression of a common conception of interaction as a "determinate machine reaction to a planned human action", is clear. The design of the network of influences between performers, environment and DSP units reveals serious consideration of the ways in which mutual and continuous influence can be achieved: the computer's output may affect the instrumentalists, the instrumental sounds might affect the computer processes, the sound might elicit a resonant response from the room, and this response might feedback to drive the computer processes or suggest changes in the instrumentalist's performance. In works such as *Texture-Multiple* (1993), and *5 difference sensitive circular interactions* (1998), he sets up relationships between the performers, machines and environment which cannot be understood in terms of simple input-output relationships[7].

As well as using instruments as passive sound-producing devices, Di Scipio develops methods by which they influence the digital transformations of the sounds they are delivering. For example a signal processing module might be set up to automatically alter its internal configuration according to changes in the input sound. As he puts it: ”sound sets the conditions and boundaries for its own transformations”. In this setting the production of sound is the by-product of a set of interactions, rather than their purpose.

2.1.6 Summary

In many areas of the field of computer music, the dominant model of interaction is a one-way street. Extended instruments focus on techniques for expressive control of DSP engines. Research in new controllers focuses on the interface between humans and computers, a consideration which is similarly central, if more complex, in score following approaches. Virtual performers that take a transformative approach employ some 'intelligent' methods of altering material but are driven cunningly by the performer.

The nature and abilities of digital instruments may have departed from those of traditional acoustic instruments, but the basic operational metaphor remains the same. At times aspects are automated, and at times the digital instrument unfolds ready made or cunningly transformed material, but in many instances of published research, there is little evidence of 'mutual influence': traffic down Winkler's street is essentially one-way. This is not to say that much exciting new music is not being made, and not meant to undermine the usefulness of the one-way model. However, its seems that some of the new interactive idioms which Winkler suggested lay ahead may not yet be being fully explored.

In systems such as Lewis' *Voyager* and Blackwell's *Swarm music* we see explicit exploraration of conversation style models. Di Scipio is not only closing the loop, but setting up circular, self-controlling processes in components throughout the performance system. The software takes on a decidedly active role in the performance. Notions of control have disappeared and are replaced by cooperation creating the possibility for more flexibility in improvised situations and arguably supporting a more spontaneous form of conversation. This system goes beyond any of the traditional performance practice models offered by Winkler. Others seeking novel approaches suggest looking to the broader field of interactive and generative arts. In the next section then, consideration will be given to interaction in the Alife installation arts which adopt a similarly inclusive, cooperative approach.

2.2 Interaction in Alife Installation Arts

Since the late 1990s, many artists interested in exploring this shift from control to cooperation and the creation of emergent behaviours in systems with a degree of autonomy have been exploring techniques drawn from the field of Alife. Alife is an interdisciplinary scientific field concerned with the creation and study of artificial systems that manifest low-level properties of living systems. In contrast to traditional AI which concerns itself primarily with high level cognitive competencies specific to human, Alife focuses

---

[7] A similar set up was explored in using analogue electronics by Gordon Mumma in *HornPipe* (1967)

on the basic adaptive abilities common to all living creatures: rather than the top-down approach of AI which specifies representational models of cognitive capacities directly, Alife is characterised by a bottom up approach in which adaptive behaviours emerge out of the organised interactions of low-level dynamical processes. The differences between these paradigms are considered in more detail in Section 2.3.

Many artists have looked to Alife techniques specifically because of the richer possibilities they afford for interaction. Ken Rinaldo, for example writes:

> "Perhaps the greatest potential for the arts and Artificial Life techniques is that they have presented opportunities for both artists and viewer/participants to develop true relationships with the computer that go beyond the hackneyed replicable paths of "interactivity" which have thus far been presented by the arts community." - Rinaldo (1998), p.374

For the visual arts, Alife techniques brought about a significant change in the nature of the artwork. As Sommerer and Mignonneau (1998) describe " the art work . . . is no longer a static object or a predefined multiple choice interaction but has become a process-like living system." (p.158). This change has brought many of the concerns and ambitions of visual installation arts much nearer to those in the performing arts. For example Rinaldo's (1998) vision of "a cybernetic ballet of experience, with the computer/machine and viewer/participant involved in a grand dance of one sensing and responding to the other." allies closely with the aims of those developing music performance systems.

This section looks at some of the ways interactivity has been developed within Alife installation arts. Consideration is given to the balance of influence between user and the system and how much freedom there is for 'spontaneous' conversation.

2.2.1 Breeding Artificial Forms: Interacting with Evolution

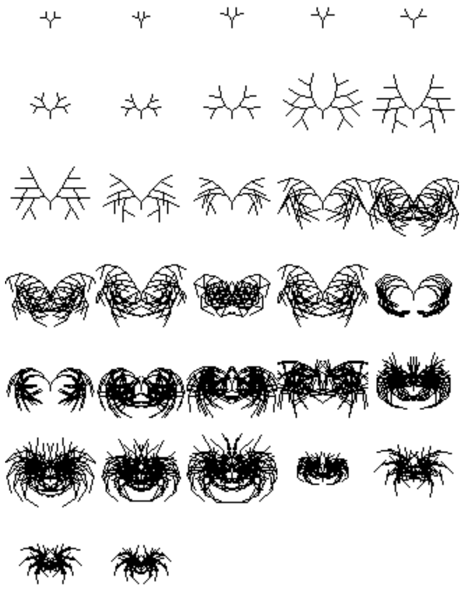


Figure 2.4: Screen shots of the 'stages of evolution' of one form in a biomorph environment. Each image (left to right, top to bottom) represents a mutation of the previous.

Some of the earliest explicit applications of Alife techniques in the interactive arts is seen in the work of William Latham and Karl Sims. This work has spawned a large and growing community of what has been dubbed 'breeder art' (Whitelaw (2004)). These original works were heavily inspired by *Biomporph*, a program which evolutionary theorist Richard Dawkins developed and discussed in his book *The Blind Watch Maker* (Dawkins (1986)). Dawkins used Biomorph to illustrate his arguments concerning the creative capacity of Darwinian evolution, namely that random variations created by chance mutations can be shaped into complex forms by natural selection.

The program presents the user with a selection of stick-built forms. The user selects an individual to become parent. The genes of the selected individual are copied and each copy is altered slightly creating a set of similar, but not identical, children. The 'genes' of each individual determine its visual appearance, specifying aspects such as colour, number of segmentations, depth of recursion, seperation of segments etc. From a starting point as simple as a five pixel cross, complex insect-like forms can be evolved. Some examples are given in Figure 2.4. Dawkins' premises and perhaps implementation are debated in the biological and evolutionary communities, but to the artistic community *biomporph* demonstrates that artificial genetics and artificial evolution, guided by human aesthetic preference, can give rise to complex visual objects (Whitelaw (2004)).

Karl Sims' (1991) *Genetic Images* was shown at Ars Electronica and the Pompidou Centre in 1993. The installation itself consisted of an arc of sixteen video screens, each displaying colourful abstract images. Users interact with the piece by pressing on the touch sensitive pads placed at the foot of each screen. Their selections form the basis of the next generation; over succesive generations they exert a steady influence over the nature of the generated images as they guide the process of graphic variation. The genome in this case is a mathematical equation. If one image is selected this single equation will be altered randomly to reproduce another sixteen forms. When two images are selected they are spliced together using a process analogous to cross-over in biological sexual reproduction to produce another sixteen images, each bearing hallmarks of the parents.

Figure 2.5: Karl Sims' Genetic images on display at the Pompidou Centre, Paris (1993).

William Latham's *Mutator* employs a similar process of aesthetic selection but operates on geometrical procedures rather than mathematical equations. The basic geometric building blocks and methods of transformation and accumulation are evident in the final forms which are reminiscent of mutant space age crustacaens (shown in Figure 2.6). Similar techniques have been widely adopted throughout the arts community. Although essentially a constraint satisfaction algorithm, or search tool, emulation of Darwinian processes raises discussion of the Promethean status of the artist. Just as electronic instruments blur the boundaries of instrument maker, composer, and performer, many working in this area discuss the dual roles that the artist takes on. Todd and Latham (1991) in particular discuss the way that 'evolutionism' changes the role of the artist. Rather than creating a work directly, the artist's task now is "the creation of generative systems and structures" on one level and "the selection of specific forms and animations" on the other. Introducing an analogy, partly perhaps inspired by the organic forms, Latham and Todd analogise these roles: "The artist first creates the virtual world ... then becomes a gardener within this world he has created" (p.12).

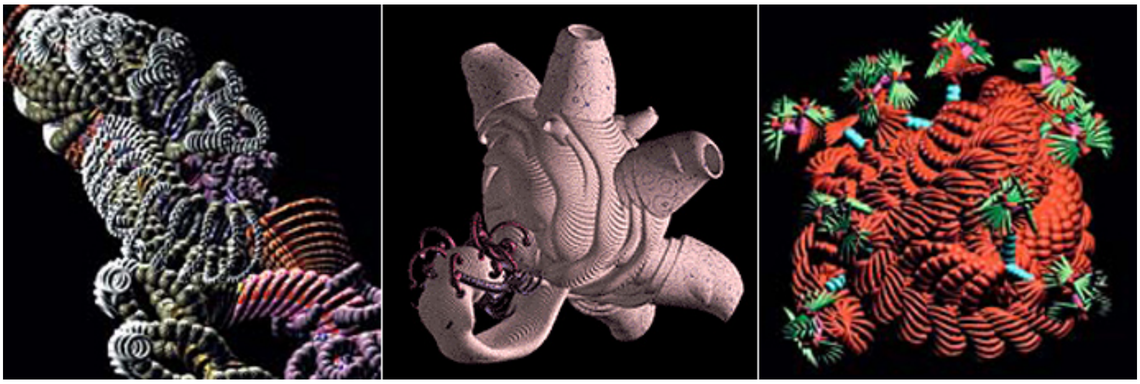


Figure 2.6: Examples of forms evolved in William Latham's Mutator.

But how far does this really take us from the one-way model of interaction? The role of the artist-creator allies with that of the instrument maker in defining a space of possibilities; the artist-gardener nurtures and explores these just as the player explores the bounds of their instrument. Part of the process is automated of course, giving some degree of agency which can surprise or upset the gardener player, but the gardener essentially remains in control of a collection of computational specifications. Existing musical applications of these techniques which will be examined more closely in Chapter 3 support this.

Others have developed more sophisticated interfaces. In Christa Sommerer and Laurent Mignonneau's *Interactive plant growing* (1993)[8] for example, users can influence the development of plants. As in both Sims' and Latham's program, these are virtual silicon graphic plants, but rather than pressing buttons to generate a series of static forms, a small garden of real plants acts as interface to a dynamic jungle of virtual plants. A user touching one of the potted plants may cause a three-dimensional fern like plant to start growing on the screen; another user brushing past a different potted plants triggers the growth of a vine, tree or moss. Each potted plant is fitted with electrodes that pick up the electrical potential of the plant. The signals differ according to the way the plants are touched. Voltage changes are then mapped to parameters which influence the growth patterns of the synthetic plants.

Although the technical processes do not depart wildly from the interactive GA, the juxtaposition of organic and synthetic plant life illustrates an interplay of natural and synthetic forms which resonates with considerations of integrating acoustic and digital musical practices.

---

[8] `http://www.mic.atr.co.jp/~christa/WORKS/CONCEPTS/PlantsConcept.html`, `http://www.medienkunstnetz.de/works/the-interactive-plant-growing/`

Figure 2.7: Christa Sommerer and Laurent Mignonneau's A-Volve.

2.2.2 Artificial Ecosystems: Interacting with an Environment

As the complexity of the simulated world grows, so the influence of the user's action diminishes. Many artists have developed systems that are based on the ecological concept of an ecosystem. These systems tend to model not just the genetic code of individuals, but the behaviours between individuals and the complex interactions between multiple species and their environment.

Many artists have explored these system in non-interactive situations where visitors are invited to observe the evolution of new forms and behaviours as the virtual entities themselves interact – mating and being born, eating, dying and competing for resources. Some forms of agent based modelling have been explored by composers, which will again be considered in Chapter 3, but they are used as closed compositional systems rather than in any interactive performance situations.

Those that have opened up these environments to human influence in an exhibition setting provide an interesting model of interaction that is subtle yet reminiscent of our interactions with aspects of the real world. In *A-volve*[9] Sommerer and Mignonneau (1997), develop the intuitive physical interface, or natural interface as they call them, of *Interactive Plant growing*. The piece consists of a virtual pond projected on the under side of a shallow water-filled glass tank. The pond is stocked with strange aquatic creatures: digital coelenterates slowly pulse along the edge, attracted to the visitors' hands that are dipped in the water.

A-volve not only simulates the basics of genetic evolution but incorporates rudimentary vision and a range of simple behaviours. These extend beyond predation and mating to deal with aspects of the physical and social worlds such as collision avoidance and parental protectiveness. The system also includes a model of basic jelly-fish-like

---

[9] `http://www.mic.atr.co.jp/~christa/WORKS/CONCEPTS/A-VolveConcept.html`, `http://www.medienkunstnetz.de/works/a-volve/`

propulsive locomotion which takes the radial shape of the creature into account, relating physical form to swimming ability and thus fitness.

Visitors can interact with the installation in two ways. By drawing on a pressure sensitive tablet on the side of the pond, they can introduce a new species into the pool. Their sketched outline is then rotated in three dimensions to create the body shape of their new species which is then released into the pool. Because radial form determines swimming ability, users can introduce a streamlined predator capable of wiping out the more sluggish members of the population, or a cursed defenceless blob which may provided refreshments for the faster fish to feed on. Motion tracking and shape-detection techniques bind the physical and virtual tanks, allowing visitors to interact with the creatures in an open and intuitive way: by dipping their hands in the tank, visitors can protect a creature from being eaten, or herd two together to encourage mating (or predation if they are that way inclined). The artists report that many become protective over their own creations, selecting its mates, or even gathering food for it.

Figure 2.8: Jon McCormack's *Eden* in installation.

Jon McCormack's *Eden* similarly allows visitors to influence the development of a virtual ecosystem. Derived from Holland's ECHO system (e.g. Forrest and Jones (1994)), Eden is a virtual projected space inhabited by rule-based, evolvable agents. These are represented visually as abstract forms that similarly predate, mate and evolve over time. Rather than the simulated sight of A-Volve populus, McCormack encodes a complex mechanism for the evolution of sonic communication. Unlike many models (some of which will be discussed in the next chapter), the ability to 'sing' and locate sound sources, does not come with a hardwired survival value. Never-the-less, in some runs McCormack (2001) reports that agents can evolve to utilise their sonic capabilities. Some creatures evolve altruistic behaviour, calling to invite others to share an abundance of food, others learn to exploit this altruism by developing siren-like tricks to lure their neighbours to their death.

Sadly the creatures of Eden (at least in current incarnation) do not respond to heckles or wolf whistles from exhibition visitors, however video analysis fuses the virtual and physical space, allowing visitors to influence the evolution of the ecosystem. The presence of people in the real environment increases the rate of biomass growth at the corresponding point in the virtual world. As this is a spatial world, the idea is that agents with the most interesting behaviour will attract more people to their local vicinity, increasing their potential food resource and so chances of survival. Conversely the level of movement in the physical world increases mutation rates of agents in the localised virtual world: increasing mutation rates means that new (although not necessarily better) behaviours or forms are more likely to evolve.

By McCormack's own admonition, the time scales involved in Eden's evolution, and the subtlety of influence means that many people are completely unaware that they exert any evolutionary pressure on the system. Indeed a major motivation for this twist to the straightforward interactive evolution of the gardening variety is to try and create an situation in which a more open-ended evolution may occur: interaction is a means to an end rather than the end in itself.

Both McCormack and Vorn and Demers, who work with robotic sculptures, describe their systems as reactive rather than interactive, emphasising the fact that users do not gain control of the self-steering system, but rather influence the unfolding of higher level events. Interaction in these systems is a far cry from determined control of the one-way model. The dynamic, adaptive nature opens possibility for spontaneous conversation, but the complexity of the system and indirect affect of the human visitor is such that the power of influence has swung the other way: the system is doing more talking than listening.

2.2.3 Single Synthetic Agents: Interacting with an Autonomous Other

Figure 2.9: Simon Penny's Petit Mal in action.

The sense of agency which emerges from these digital ecosystems is at once more minimal and more powerful when the visitor interacts with a digital creature on a one-to-one basis. One of the simplest, but perhaps most universally appealing examples of 'cybernature' is Richard Brown's *mimetic starfish*[10]. This is an almost photo-realistic, although over-sized, starfish which is projected onto a low circular white table. A transparent interface tracks the movements of visitors' hands and the image responds incredibly evocatively: tentacles stretch out languidly in response to kind, soft stroking and recoil sharply at fast, aggressive movements. Even inanimate objects placed in the vicinity of a tentacle cause it to reach out as if the starfish is inquisitively trying to examine the object.

Although perhaps not the most sophisticated example of Alife art, the starfish's basic reactive responses induce visitors to enter into a form of gestural turn taking, forming a prototypical communication of gesture. The physical reality of the system is nothing more than an array of projected light, but the starfish is attributed with a degree of agency and even personality.

This sense of autonomous agency is increased considerably when artworks escape the confines of the virtual and are palpably present in the form of physical robots. Artists such as Rinaldo, Yves Klein, and Vorn and Demers have all taken this leap into the real world which enables an interface-less open form of interaction. *Petit Mal* for instance is Simon Penny's 'autonomous robotic artwork'[11]. Petit Mal stands just over a metre tall and physically comprises a scrawny counterweighted column encircled with ultrasonic and pyroelectic sensors balanced precariously on two bicycle wheels.

> "The robot presents someone with the impression of a non-human, non-animal sentience, which then has to be dealt with in some way. If they run away, it will chase them. If they want to play, it will play. If they are aggressive and advance, it will back off. At some point, if you're boring, it gets bored and goes away." - Davis (1996), p.32

In contrast to the lavish graphical worlds of the virtual ecosystems described above, Penny's work presents the minimal requirements for a basic illusion of sentience. Visually he makes an effort to present the public with something that is neither a biological simulation, nor an automaton, but something that is 'substantially itself'. He describes his approach as under-engineering, and capitalises on mechanical or electronic quirks as the generators of emergent behaviour such as the dynamics of the double pendulum structure which is the central control system, or deficiencies in sensor readings. It is these quirks, he suggests, that give rise to its "personality". Petit Mal is an attempt to explore interactive machine behavior in a real world setting. The reflexive nature of interactivity is a focal issue: interactive behavior is defined by the cultural experience of the human visitor.

Pieces like these support a simple but powerful conversational interaction. Petit Mal in particular engages those it encounters, seemingly judging their moves and altering its behaviour accordingly. These basic behaviours support a strong sense of interaction. Much of this perhaps is associated with its embodiment, as Penny himself says, 'evaluation of interactivity is subjective', this raises the importance of considering not only the formal system, but how it is dressed and presented to the world. The alluring personalities exuded by some of these artificial forms inspire a possible style of conversation for our digital performance partners. The sense of agency that we attribute to some of these systems would be a very attractive property in a musical performance partner, offering a novel alternative to transformational approaches for generating musical material.

2.2.4 Summary

In many respects, the nature of interaction in these systems is much closer to the fluid models of communication sought by members of the interactive music community: no-

---

[10] `http://www.mimetics.com/starfish.html`

[11] `http://www.ace.uci.edu/penny/works/petitmal/petitcode.html`

tions of control and masterability central to the one-way model are replaced by indirect influence and cooperation. This represents an attractive shift for those aiming to create a sense of distinct musical personality in man-machine improvisation. But in many cases, the systems are perhaps *too* independant. McCormack and Vorn and Demmers describe their systems as reactive, stressing that visitors can only indirectly and inconsistently influence their path. This is an issue which Rinaldo sees as an obstacle:

> "One difficulty with some Artificial Life artworks is that the systems may not seem to be responsive to the changing environment, as the work demonstrates its own internal desires. This can make the work seem unresponsive or uncaring." Rinaldo (1998), p.373

Control of the dialogue has switched: the machine has gained control and is not listening. How then could we gain the balance desirable for man-machine improvisation? Simply reducing the complexity of the system so that the human visitor engages with a single virtual entity rather than an entire ecosystem as in Richard Brown's starfish and Penny's Petit Mal seems to begin to readdress this balance and evoke an attractive form of conversational interaction. Not only is there a proto-conversation, but the aesthetic achievement of a sense of artificial agency. In contrast to the transformative approaches used in artificial improvisation software, the system demands to be encountered on its own terms. How can we understand the nature of these differing forms of interaction so that we can preserve the desirable qualities of the fluid conversation but temper the balance such that we can employ these systems on stage for musical performance? How does the framework adopted by Alife artists differ from that of traditional interactive music?

2.3 Frameworks for Understanding and Implementing Interactive Systems

This first part of this section considers the dominant design protocol in current interactive computer music research, and contrasts this with the scheme of those promoting the conversation model. These two approaches ally closely with two different philosophical perspectives on interaction which are introduced in Section 2.3.2. It is suggested that existing approaches tend to operate within a computationalist paradigm and that the adoption of a dynamical perspective may support the development of software that is capable of engendering a more conversational style of interaction.

2.3.1 Design Schemes: Pipeline vs Circular Causality

Traditionally, interactive software design is split into three principle parts: sensing, processing and responding (Rowe (1992)). Shown schematically in Figure 2.10 the first two steps can be divided into sub-tasks as in Winkler's description:

- Listening (input). Human activity is translated into digital information and sent to the computer.
- Listening (analysis). The computer receives human input and analyses the performance information for timing, pitch, dynamics etc.
- Processing (interpretation). The software interprets the computer listening information, generating data that will influence the composition.
- Processing (composition). Interpreted information is used to guide the generation of composition data.

- Responding. The computer plays the music, using sounds generated internally or by sending musical information to devices.

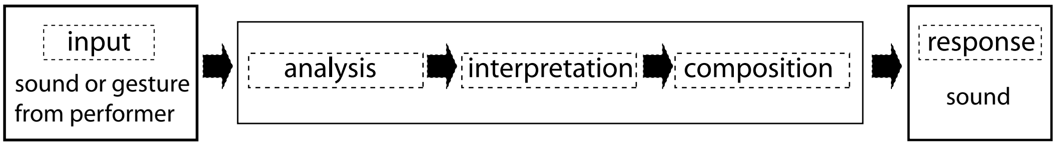


Figure 2.10: Schematic of system design protocol in interactive computer music.

This basic protocol has remained unchanged, appearing in recent reviews (Schnell and Battier (2002)), and is virtually ubiquitous throughout the community. Although we may assume that most performers think in terms of a conversational model when playing music, a notion that presumably carries over to performance in digital systems, this schema is used to describe, and thus presumably influence, the design of performance systems. As is evident, the basic ontology is linear, what we might call a pipe-line model of information processing. The performer exerts control over the system by providing sonic input. The system analyses it and computes a response. We may assume that the future actions of the performer in most circumstances will be affected by the sound, so affecting the state of the computer system and closing the loop, but this is only implicit. Although this scheme is presented again and again in the context of performance systems, it sidelines the fundamental feedback from the system output to the performer.

The other noteable feature of this design scheme is the decomposition of the task into functional components: input, analysis, interpretation, composition, response. Naturally, many of the steps in this process are based on cognitive models of the comparable processes in humans. For example improvisation systems designed by Wessel in the late 1980's, for performance with Roscoe Mitchell and later with Ushio Torikai, consisted of a collection of objects divided into three modules: *listening assistants*, which extracted musically meaningfull information, i.e. DSP objects for pitch extraction, parsing, and tempo extractors and objects for musical analysis; *composing assistants*, which helped the improvisers manage the construction and set-up of their performance materials on the fly; and *performance assistants*, which supported gesture mapping, phrasing and articulation. The listening assistants were based on a model of memory influenced by cognitive psychology, comprising both direct short and more abstract long term memory. The design of the phrase boundary was based on the elementary grouping mechanisms taken from Lerdahl and Jackendoff's Generative Theory of Tonal music (Lerdahl and Jackendoff (1983)), whilst the tonal field estimation scheme was based on a method developed by Krumhansl (1990). This approach has proved very powerful in certain areas, and the development of sophisticated and powerful audio analysis tools are undoubtedly central to the successful development of digital performance systems.

A potential problem though, is that this pipeline model reinforces the metaphor of one-way interaction. Coupled with the power of these intelligent analysis modules, the motivation for considering other forms of interaction is obviated. As noted above, electronic and digital instruments are distinct from acoustic instruments in that the interface is decoupled from the sound producing mechanism. It is very natural then to approach the design of new instruments by focusing on the input device and response mechanism as distinct design problems. Layers of complexity are then added as interactive systems become more sophisticated and this model is applied not only to instrument design, but becomes the de facto approach to 'artificial players'. In a recent review and project proposal, Robertson and Plumbley (2006) describe their aims of creating an 'autonomous player' and break down the task in exactly this way. "The first stage is the development of a system that is capable of predictive score following and sequencing. We consider that this is a sub-problem of our eventual aim of an interactive system capable of autonomous generation of output." (p.3).

Many researchers wanting to push the boundaries and explore what they feel are richer interactive relationships invariably invoke discussion of feedback and circular causality within a systemic view. Garth Paine for example takes inspiration from Cybernetician Norbert Wiener in presenting a more inclusive consideration of the elements of an interactive system:

> "In a sense, the ... interactive musical environments creates an ecosystem formed by the human presence and nature of behaviour, the response of the technology (the aural or visual response as experienced by the inhabitant of the installation) and the space itself. The process of understanding this dynamic relationship between the human condition and the physical space is supported by the study of cybernetics, and in particular the closed causal loop." - Paine (2002), p.301

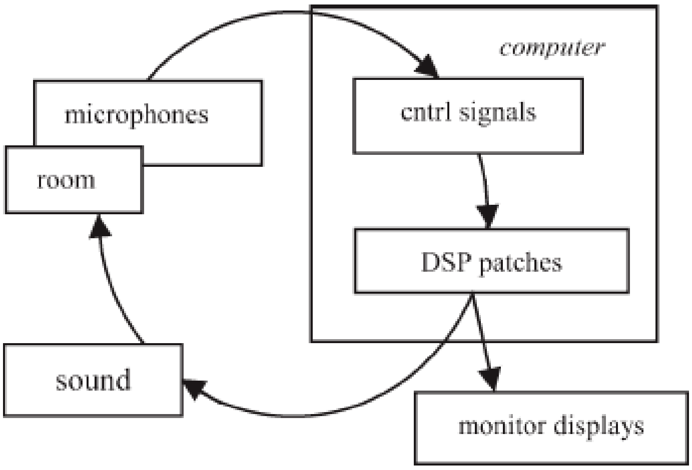


Figure 2.11: Basic design of Di Scipio's Audible Eco-Systemic Interface.

This is almost precisely the approach taken by Di Scipio in AESI, the basic design scheme for which is shown in Figure 2.11. This closed causal loops captures much more accurately the ongoing negotiations that can occur simultaneously between individuals in a group of performing musicians. In aiming to build systems with enough freedom to improvise with, the basic design scheme should ideally facilitate the possibility for this continuous and simultaneous influence throughout all parts of the system. The switch from a linear to a circular scheme intuitively supports this more open, dynamic relationship between the system and performer. Closing the loop infers a continuous flow of influence. Establishing a continuous flow, rather than set of determinate feed-forward commands in turn increases the flexibility and so the possibility for more spontaneous communication.

2.3.2 Heteronomy vs Autonomy

The distinction between the determinate control of the one-way model and the mutual influence of the conversation model are well recognised within certain philosophical circles where they represent the paradigms for interacting with *heteronomous* and *autonomous* systems respectively. Autonomy and heteronomy literally mean self-governed and other-governed: an autonomous system is described as a self-determining and self-maintaining system; a heteronemous systems is determined and controlled externally.

Biological entities across many scales – a living cell, a jellyfish, an ant colony, or a human being – behave as a coherent, self-determining unity. On the other hand, most human-designed technology such as an automatic bank machine is determined and controlled from the outside (Thompson (2007)). In terms of organisational structure, a heteronomous system is defined by an input-output information processing flow which is controlled externally. By contrast an autonomous system can be characterised by highly recursive network of dynamic processes that generate and maintain internal variants in the face of external (and internal) disruptions (Varela (1979)).

Traditional computationalist systems – both cognitivist and connectionist – typify the heteronomous perspective. For example a typical connectionist network has an input layer and an output layer. The inputs are assigned by the observer of the system, and the output performance is evaluated in relation to an externally imposed task. Rule-based or statistical models such as the Markov chains used in the improvisation systems of Pachet and others work in a similar way, albeit with one layer of abstraction. The observer (composer or performer) selects or plays a set of notes (inputs) into the system, the system creates an abstracted representation of this input data. This is then used to define the output according to subsequent given inputs. These systems are typically constructed from the 'top down', comprising independent processing modules which pass information to each other.

In contrast, an autonomous perspective assumes a dynamical approach in which there are no inputs or outputs in the usual sense, but rather a closed loop of circular causality. The dynamical processes of the performance networks composed by Di Scipio, or the evolving Alife ecosystem models lie closer to this notion of autonomy by virtue of the recursive loops which bolster endogenous, self-organising and self controlling dynamics. Within this framework, design is attacked from the bottom up. As Di Scipio notes, in a musical context the final sonic output is the 'by product' of sets of composed interactions.

In these next sections, a brief outline of the computationalist approach is given, followed by a closer consideration of autonomy and an outline of how the attendant dynamical approach differs.

2.3.3 Cognitive Science and the Computationalist Approach

Mainstream cognitive science can be characterised by two central pillars. Firstly the computationalist theory of mind is upheld, which supports the 'physical symbol system hypothesis'. This states essentially that a physical symbol system like a computer is both necessary and sufficient for general intelligent action (eg Newell and Simon (1976)). Secondly, internal activity is divided into a sense-think-act (or sense-model-plan-act) sequence as in David Marr's description of how a three-dimensional world model could be constructed from a two-dimensional image in order to generate appropriate action.

These two principles were brought together in a classic paper by Lachman et al. (1979) in the late seventies where they define cognitive science as "how people take in information, how they recode and remember it, how they make decisions, how they transform their internal knowledge state, and how they translate these states into behavioural outputs" (p.99). If we compare this to the dominant scheme in interactive computer music we see just how strongly the two are allied: 1) translate human activity into digital information, 2) analyze the performance information and extract pitch, timings, dynamics etc., 3) interpret information for use in influencing the composition, 4) process this information to generate a composition and 5) send information to sound generating device and output. (Winkler (2001) p.6.) The approach, both in terms of functional decomposition and the linear information processing flow represents the classical cognitivist approach to cognition as practised in classical AI.

2.3.4 Autonomy

Definitions of autonomy are couched in many different ways, from practical sets of engineering constraints in robotics research, to extreme theoretical accounts of the fundamental organisational principles of biological life (Varela (1979)). Since the 1990s, the term autonomous robotics (Maes (1991)) has been used to refer to a set of engineering constraints on the design and evaluation of robots in both cognitive science and engineering. These include conditions like mobility, and real-time response in real-world environments, no remote control, no external energy supply and no human intervention in task solving. Brooks' situated robotics (e.g. Brooks (1991)) develops upon these ideas. Many Alife artists working in hardware acknowledge the work of these researchers as a major influence, and it is perhaps the achievement of these engineering constraints which Rinaldo and Penny refer to when they talk of autonomy.

Practical research in autonomous robotics has driven a deeper consideration of the notion of autonomy. Increasingly consideration is given to the types of interactive processes that are established between the robot and its physical environment, as well as the properties and dynamic structure of its control mechanisms. Engineers such as Tim Smithers (1997) and Randall Beer (1995) strongly criticise the classical computational information processing approaches which provide little room for considering these aspects, putting forward dynamicism, embodiment and situatedness in place of the virtual, formal approach of traditional AI.

Many working in this field have been greatly influenced by the biologist and philosopher Francisco Varela. He approaches autonomy in terms of the organisational structure of an entity. For Varela, autonomy comes about by virtue of a systemic organisation which defines its own identity: a kind of self maintaining, self-reinforcing and self-regulating system subserved by a highly recursive network of dynamic processes, capable of generating and maintaining internal variants in the face of disruptions both internal and external (Varela (1979) p.55). Varela defines an autonomous system as one which has organisational and operational closure. Closure doesnt mean that the system is cut off materially and energetically from the outside world (that would be impossible) but refers to a system whose organisation is constituted by a network of internal processes. The operation of the network is sufficient for those constituting processes to be generated and sustained without any of them being driven from outside the system.

Ruiz-Mirazo and Moreno (2004) give an account of *basic autonomy* in terms of the energetic and thermodynamic requirements in the physical world. This brings with it specific and demanding physical-implementation requirements: "the system must be made up of certain types of components, specifically a semipermeable active boundary (a membrane), an energy transduction/conversion apparatus (an energy currency like ATP in living cells, which transfers energy from chemical bonds to energy-absorbing reactions within the cell), and at least one type of component that controls and facilitates the self-construction processes (catalysts)". (p.252).

In the biological domain, this form of autonomy is exemplified by a living cell. The recursive constituent processes in this case are chemical. Their recursive interdependence takes the form of a self-producing metabolic network which also produces its own membrane. This network constitutes the system as a unity in the biochemical domain and determines a domain of possible interactions with the environment. This kind of autonomy in the biochemical domain is referred to as *autopoiesis* (Maturana and Varela 1980).

A system does not have to be autopoetic in the strict sense of maintaining a self-producing, bounded molecular system, but it remains unclear exactly how the autonomy of an artificial agent can measured or implemented. Aiming to develop a more operational definition for use in Alife simulation, Barandiaran (2004) has recently developed a definition of autonomy in the behavioural domain. Working within a dynamical systems framework, Barandiaran models the metabolic constructive processes of basic autonomy as a set of essential variables which tend to stay away from equilibrium. He describes a *behavioural adaptive autonomy* which is defined as "homeostatic maintenance of essential variables under viability constraints through self-modulating behavioral coupling with the environment, hierarchically decoupled from metabolic (constructive) processes." (see Barandiaran (2004) for explication).

Now none of the art works or music systems we have looked at are autonomous on these terms. Indeed we may not really want a truly autonomous system as a partner in improvisation. However, the framework adopted by researchers taking this autonomous perspective on cognition provides much inspiration for the current project. Unpacking some of the tenants central to those adopting autonomous approaches to cognition and Alife helps contextualise the instrument and conversation models of interaction, clarifying the differences between them and providing conceptual and practical tools for the current project.

2.3.5 Dynamical Approaches

Proponents of autonomous robotics and Alife are a sub-population of a large community of research adopting a dynamical approach to mind science (e.g. Kelso (1995), Lewis and Granic (2000), Port and Van Gelder (1995), Thelen and Smith (1994). In contrast to the cognitivist hypothesis mentioned above – that cognitive agents (natural and artificial) are digital computer or physical symbol system and that cognition should accordingly be explained in symbol processing terms – the dynamical hypothesis postulates that the cognitive systems instantiated in natural agents are dynamic systems and therefore that action, perception and cognition should be explained in dynamic terms.

The cornerstone of the dynamical approach is the emphasis it places on time. Traditional computational models are static, in that they specify a sequence of discrete states which a system passes through. Dynamic-systems models specify how a process unfolds in real time. As Tim Van Gelder (999a) states, "Although all cognitive scientists understand cognition as something that happens *over* time, dynamicists see cognition as being *in* time, that is, as an essentially temporal phenomenon." (p.244). Elsewhere he compares the approaches with a series of oppositions: change versus state; geometry versus structure; structure in time versus static structure; time versus order; parallel versus serial; and ongoing versus input/output (Van Gelder (1998)). At the lowest level then it should become apparent that the adoption of a one-way instrument metaphor implemented within a pipe-line model of information processing allies with a computationalist approach to cognition, whereas the conversation model based on circular feedbacks sits within a dynamicist understanding.

Dynamicists conceive of state changes in terms of their position and trajectory in phase space, i.e. geometrically: computationalists focus on the internal formal or syntactic structure of combinatorial entities. Computationalists think of cognition as the rule governed transformation of one formal static structure into another, whereas for dynamicists cognitive structures are laid out as temporally extended patterns of activity. Cognition is seen as the flow of complex temporal structures mutually and simultaneously influencing each other. Dynamicists are therefore interested in the timing (rates, periods, durations, synchronies) of processes, whereas computationalists have traditionally not been interested in these temporal details, but only in the order in which cognitive states occur (Thompson (2007)).

The serial progression of listen-process-respond is the musical equivalent of the computationalists' sequential ordering of cognitive subtasks into sense, plan, think, act. This contrasts with the dynamicists' conception of cognition as the unfolding of a continuous coevolution of acting, perceiving, imagining, feeling, and thinking. These basic comparisons promote two very different understandings of the relationship which an entity has with its environment, of interaction. Let's assume that the entity is a human musician and the environment is the piece of performance software. Under the one-way model, the human sends a trigger which sets off a series of events that flow down the processing chain. The response is fully determinate so it can be controlled and mastered. The conversation model assumes that both entities are autonomous systems each maintaining their identity and thus independence but influencing each other through a process of structural coupling. Coupling refers to the fact that the conduct (dynamics or behaviour) of one system is a function of the conduct of the other. In dynamic systems language, the state variables of one system are parameters of the other system and vice-versa (where a variable determines the state of the system along a trajectory within a particular field, and parameter determines the field in which it currently exists). In a truly autonomous system, the domain of interaction is determined internally, thus we cannot absolutely predict what its response will be. We cannot control it. A completely autonomous system in this sense is neither desireable nor achieveable as a performance partner. But *some* degree of internally generated state is the first step toward creating a sense of distinct musical personality.

The dynamical perspective provides a framework within which we can understand how our action can have differing levels of influence on a system: why some Alife systems are 'uncaring', whilst others such as starfish are so attendant to our moves.

Agent-based evolutionary ecosystem models such as that used in McCormack's Eden are described by a large set of interdependent processes operating across different time frames. The global system is constituted by a complex set of evolving processes operating on multiple time scales. The behaviour of any one agent is determined by the state of its local environment (which could include other agents) and the state of its internal system. This internal system mediates the sensors and actuators, so determining its interaction with the environment and is subject to lifetime learning as well as evolution. Population dynamics are influenced by the inbuilt seasonal variation in biomass density, which will presumably cause population levels to fluctuate even in the absence of any visitors. The affect we have on the system then, in influencing either biomass growth or mutation rate is only one of many factors that determine any observable changes at the level of agent behaviour, sub-population or global population. Additionally any affects will firstly be extended in time, and possibly space. This makes our influence perceptually inseparable on a micro level from the effects of default seasonal variation or the effects induced by others in the exhibition.

In Brown's starfish however, there is a much stronger and more direct coupling between our behaviour and the behaviour of the agent. Firstly there is only one agent, secondly the interaction occurs in the same time scale and is local, and finally although there may be a small amount of noise or minor oscillation injected into the process which controls movement, the principle variable is the motion detection system which directly initiates reaching or recoiling behaviour.

2.3.6 Bottom Up Design

These examples also illustrate the impact of adopting a dynamical approach in terms of the conceptualisation and design of a system. As noted above, system design under a computationalist approach typically proceeds in a modular fashion, where for example the interactive interface is designed and implemented separately from the composition module. As becomes apparent in considering Eden, what the audience interacts with isn't some separate interface module, but the resource levels in the environment which are intrinsic to the fundamental processes by which the whole system is constructed, maintained and evolves.

These opposing perspectives ally closely with the top down versus bottom up approaches taken by cognitivist and dynamical approaches to understanding cognition: whereas cognitivists focus on some kind of central processor or homunculi that controls behaviour, a dynamicist considers the distributed and functionally integrated network of recursive processes from which a coherent behaviour emerges as a global product of the system. The switch in perspective is neatly summarised by Di Scipio's suggestion that his ecosystemic approach represents " a shift from creating wanted sounds via interactive means, towards creating wanted interactions having audible traces." (Di Scipio (2003) p.271).

The notion of emergence is central to dynamical and Alife approaches to life and mind, and is an equally close to the hearts of Alife artists. Within Alife research and amongst those concerned with autonomous systems the concept of emergence as a process underpins the methodological approach to understanding, and attempting to simulate, life. As Langton put it in his inaugural speech:

> "The "key" concept in Alife is emergent behaviour. Natural life emerges out of an organised interaction of a great number of nonliving molecules, with no global control responsible for the behaviour of every part. Rather, every part is a behaviour itself, and life is the behaviour that emerges from out of all of the local interactions among individual behaviours. It is this bottom-up, distributed, local determination of behaviour that Alife employs as its primary methodological approach to the generation of life like behaviours" - Langton (1989), pp. 2-3

This bottom up approach deviates substantially from the standard design process of interactive music. As noted above, the channels of interaction are no longer constructed front-end interfaces, but slip streams into the internal dynamics which constitute the larger model. Recognition of this is apparent in Di Scipio's description of his construction of a performance network. "System interactions, then, would be only indirectly implemented, the by-product of carefully planned-out interdependencies among system components, and would allow in their turn to establish the overall system dynamics, upon contact with the external conditions."

2.4 Summary and Implications for Design

Consideration of these different frameworks provides a set of conceptual tools for thinking about and realising a more conversational style of interaction. Strong autonomy may not be achievable or desirable for the current project but provides a very useful set of metaphors.

On a conceptual level, the autonomy/heteronomy distinction provides a useful framework for understanding how a musician engaged in a musical dialogue with others can retain their individual identity, whilst being a part of a larger musical unit. We can conceptualise a cell as either an autonomous entity, structurally coupled to the biological environment of the body, or as a component functionally defined in relation to the larger organism. By extension we can see the individual musician as both a distinct musical personality and as an instrumentalist with a defined role relative to the ensemble of which they are a part. This points to the possibility of developing software processes that have their own musical agenda yet are able to collaborate with a human musician in an improvised performance. Rather than being a unit in the pipe-line model, the software algorithm and human musician are elements in a closed causal loop.

In very simple terms, changing the system boundary – separating human performer from computer system, or encircling both as a unified system – assists in matters of assessment. For some, the important thing is that the audience get a sense that the machine has its own musical agenda. For example Mari Kimura stated at a recent NIME workshop ″my job as a performer is to give the audience the impression that we are equal partners″. Talking specifically about Eric Singer's robotic guitar GuitarBot, with which she performed at NIME 06, she described some of the rules that defined the relationship between what she did on the violin and what the robot did. These consisted of things like: 'if the note is higher than E2, then play, else don't play'. The simplicity of such rule sets, she suggested allowed her as a performer to learn to play the system, and to create a sense of intelligence and intent on the behalf of the robot.

Figure 2.12: Mari Kimura performing with Eric Singer's Guitar Bot at NIME 2005

Other researchers feel that the real litmus test, and therefore the aim of the enterprise, must be that the performing musicians themselves gain the sense that the system has a musical voice and initiative of its own, that it instigates as well as responds to musical suggestions. Achieving this will often, although perhaps not inevitably, equate with the audience's engagement with the spectacle on stage.

In either case it seems crucial to consider the effects of the overall performance network as in the systemic perspective which is well illustrated by Di Scipio's AESI project. If we consider system design solely in terms of what happens between the input and the output of the digital system, we may fail to take into account the potentially rich effects of the real-world environment. These considerations are key in the embodied situated approaches to robotics mentioned in Section 2.2.3. In these areas, *behaviour* is defined as the observed agent-environment interactivity (in line with our every-day understanding), and *mechanism* is defined as the structure inside the agent which subserves this interactivity.

In biological systems, the important message is that behaviour is a product of the joint activity of agent, environment, and observer, so the (agent-side) mechanisms underlying the generation of any behaviour should not be assumed to be identical to the behaviour itself. An important consequence is that if a behaviour appears complex to an external observer, this does not imply that the underlying mechanisms are also complex. The classic illustration of this is the description Herbert Simon (1969) gave of an ant walking across a beach. The internal mechanisms of the (hypothetical) insect comprise simple obstacle-avoidance rules such as, if there is a rock or clump of sand to the left, go right and vice versa. The ant responds to every miniscule lump of sand, flotsum, jetsum and pebble, turning left and right and right and left as it negotiates the rough terrain. Simons points out that from the perspective of an external observer the trajectory traced by the ant is strikingly, and perhaps irretrievably, complex. This classic example serves to illustrate the possibility of achieving complex behaviours from simple mechanisms, something which is an attractive possibility for any designer of creative software. In a musical context, where the environment might be sonic, we can potentially generate behaviours that are not only apparently complex, but that are also contingent on current events in the sound world.

The situation where both the agent and the ant are capable of dynamically adapting is captured by the concept of structural coupling and puts forward an understanding of interaction that describes our experience of musical interaction with other human musicians in a much richer way than any models currently used within interactive computer music. The notion of coupling itself (where-by each system is a function of the conduct of the other) captures the sensation of togetherness experienced when playing. In free improvisation, and more subtly in an ensemble of scored parts, there is often no 'leader'. Whether subtle changes occur in expression, or dramatic changes in pace or texture, it is often impossible to pin down their origin.

The linear notion of cause and effect implicit in the predominant sense-plan-act software design fails to capture these dynamics which are central to the coherence of the group, and would be valuable characteristics of musical performance to move toward in artificial systems. This sense of ensemble can be achieved by attentive musicians who have never met, but we have all witnessed the phenomenon of a group who have played together over a long period of time: whether a string quartet, a rock band or a free improv group, there is something about the co-ordination of an established ensemble which belies their existence as individuals, it is as if their musical selves have somehow aligned, dragging their arms, fingers and minds with them. This type of structural congruity between autonomous entities is understood from an autonomous systems perspective as precisely the result of a history of interactions between autonomous systems, i.e. structural coupling.

Forms of these dynamic interdependencies can already be seen in Di Scipio's AESI , and perhaps in Blackwell's Swarms system. Di Scipio's performance network can itself be conceived as a dynamical system instantiated across digital and acoustic media. The closed causal loop defines an interdependency between each element and each can dynamically adjust according to reciprocal influences. Similarly, the low level rules of the boids algorithm underlying Blackwell's Swarm system defines every element in relation to each other, such that each adjusts to varying local conditions. In practical terms then, a dynamical bottom up approach can be used to place interdependent adaptive processes at the very core of the system. This facilitates the achievement of flexible and spontaneous form of interaction in a continuous, adaptive circuit that goes beyond the explicit switching of control implemented in systems such as Winkler's *Snake Charmer* and offers an alternative form of man-machine interaction.

The final important implication in adopting a dynamical perspective is the suggestion of an alternative to functional modularity. Recall that for dynamicists, cognition isn't the transformation of formal structures by distinct modules, but the temporally extended pattern of activity across the brain. Applied to the design of performance systems, this suggests that rather than designing separate interaction and composition modules, one process may be used to both subserve a response *and* generate musical material. The next chapter therefore presents an overview of the many different approaches to algorithmic composition, and considers the potential for adaptive dynamical techniques for generating musical material.

Chapter 3

Computer Automated Algorithmic Composition in Research and Practice

In the 1980s, Berg (1987) suggested that the fundamental contribution of the computer to music was that it "empowers the composer to hear that which could not be heard without the computer, to think that which could not be thought without the computer and to learn that which could not be learned without the computer. The computer can allow a composer to write music that goes beyond that which she is already capable of." (p.161). One of the earliest experiments in computer automated composition put its navigational powers to test in exactly this way. The Illiac suite (Hiller and Isaacson (1959)) consisted of four 'experiments' carried out on the ILLIAC I mainframe of Illinois University. The first two experiments were designed to demonstrate that standard musical techniques could be handled by computer programming, the third to show that computers might be used by contemporary composers to extend present compositional techniques, while the fourth experiment was intended to "show ... that computers might be used in highly unusual ways to produce radically different species of music." (ibid p.4). The article Hiller wrote at the time was pounced upon about by the popular press who whipped up a storm of controversy which lasted throughout much of the 1980s. The musical establishment of the time were so antagonised by this scientist operating his machines in the name of art that apparently neither Baker's Biographical Dictionary of Musicians nor the New Grove Dictionary of Music and Musicians recognised his existence until just before his death. Fortunately this hysteria has now subsided, and the computer, if not the algorithm, is now largely accepted as an integral part of compositional practice.

It was not only players and composers who took an interest in computational approaches: music theorists also looked to the quasi-scientific rigour of computational modelling as a means of raising their profile. In post-war universities, the hard sciences occupied pride of place, particularly in American academies. Disciplines like music theory tried to gain credit, making themselves look as hard as possible by adopting scientific language and symbol systems (Cook (1998)). By the late 1970s, attempts were made to implement the more objective of existing analysis techniques such as Schenkerian analysis in computer models (Cook (1998)). Others such as Sundberg and Lindblom (1976) attempted to objectify musicology by applying cognitive science methods. They developed new grammars specific to musical styles aiming "to describe facts in music theory by means of generative rule systems" (p.100).

Several researchers interested in the cognitive processes underlying musical composition also built algorithmic composition systems as cognitive models. Steedman (1984) declared that "generative rules are only really interesting when they can be used to drive

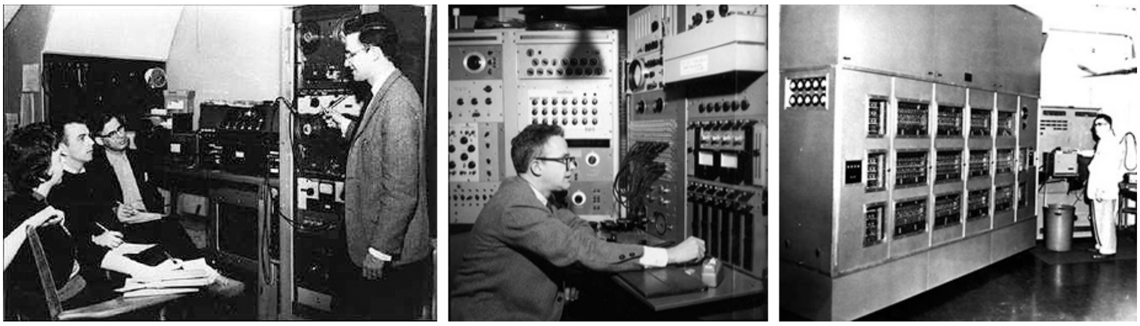


Figure 3.1: Hiller with colleagues in the Electronic Music Studios, University of Illinois (left), at the controls of the console (middle), and in front of the mainframe ILLIAC I (right).

a model of human performance," (p.75), and implemented a model of blues progressions. Johnsson-Laird (1991) set out to develop a theory of "what the mind has to compute in order to produce an acceptable improvisation" (p.291). Related research activities have gathered momentum in recent years with the burgeoning field of computational creativity which is beginning to tackle creativity in general as well as issues such as the evolution of musical behaviours (e.g. Miranda (2000b), Bown and Wiggins (2005)).

As this rapid historical jaunt suggests, the field is wide, and has a rather large number of different teams playing on it: those developing digital performance systems are only one of a large number of different groups exploring the use of algorithmic methods in music. Under the most inclusive use of term, the field draws ideas and motivations from a range of disciplines including music practice, musicology and music theory, AI, computer science, cognitive science, and more recently, evolutionary theory. Research, practice and investigation take place not only within AI departments, music departments, dedicated computer-music research centres, on new media courses, and in arts colleges, but increasingly outside institutes on the laptops of the world's more experimental electronic music producers and performers.

Testimony to the breadth of the field are the number of labels which have been used to describe overlapping, sometimes identical projects: algorithmic composition, automatic composition, composition pre-processing, generative modelling of music, generative music, active style synthesis, empirical style modelling, computer aided composition, computer composing, computer music, procedural composition, score synthesis, computer aided automated composition, machine musicianship, computational creativity and computational musicology. Attempts to distinguish these terms have been made in Spiegel (1989), Cope (1991) (p.220), Burns (1994) (p.195), Miranda (2000a) (pp.9–10), Taube (2004), Gerhard and Hepting (2004) (p.505), Ariza (2005) Roads (1985) (p.175), Ames and Domino (1992) (p.55) and Pearce et al. (2002).

This chapter focuses on compositional applications. Section 3.1 provides a rapid survey of the principle techniques that have been applied across the whole field. There are a great many of reviews and surveys of the field, so this overview is meant merely as a means of marking out the corners of the pitch and demonstrating the range of equipment available. Methods used to extract or imbue forms of musical 'knowledge' are considered separately from 'extramusical' approaches that are developed to explore new musics. Section 3.2 introduces some distinctions suggested by Pearce et al. (2002) which differentiate between scientific and creative applications. This section also raises some issues, perhaps specific to the academic community, which arise due to the cross-disciplinary nature of the field. Section 3.3 takes a more practical perspective and examines some of the motivations for using generative processes in composition or performance. The desirable characteristics of algorithms are discussed from this more practical perspective. Section 3.4 aims to gather together these different perspectives to mark out the framework for the current project.

3.1 Techniques

This section gives a brief overview of the some of the main techniques that have been applied to computer automated algorithmic composition. This whirlwind tour is meant only to provide a taste of the principle methods that have been applied to date. There exist many reviews positioned in different corners of the expansive field, adopting a number of different classification and evaluation systems. More detailed surveys of algorithm types are given in Loy (1989), Papadopolous and Wiggins (1998), Dodge and Jerse (1997) p.341, Miranda (2000a), and Todd and Werner (1999). Chronological overviews are given in Hiller (1981) and Burns (1994), whilst Ariza (2005) gives a taxonomy according to various descriptors[1]

At the start of the new millenium Miranda (2000b), asserted that the discussion as to whether or not computers could compose music was no longer relevant. Suggesting that one of the greatest achievements of AI to date was the demonstration that machines can compose music "of incredibly good quality", he cites Cope's EMI system (e.g. Cope (1991)), which is famously reported to have fooled the critics with its compositions in-the-style-of X, to demonstrate that computers *can* compose if programmed accordingly. What these AI systems are good at is mimicking established musical styles, either by being hard wired, deriving statistical patterns from exemplar of the extant cannon, or searching spaces of musical possibility using Evolutionary Computation (EC) techniques guided by a human ear or imbued more implicitly with musical knowledge. He goes on to suggest that the study of 'new' music is trickier, partly because "it is hard to judge what the computer creates in such circumstances because the results normally sound very strange to us. We are often unable to judge these computer-generated pieces because they tend to lack those cultural references that we normally hold on to when appreciating music." (ibid, p.1). Issues of 'judgement' will be discussed in Section 3.2 and 3.4, but this distinction between the replication of existing styles versus exploration of new ideas forms a natural division which is adopted here.

In the first half of this section, an outline is given of the various techniques that have been applied in an AI framework to imbue a system with some degree of musical knowledge. These are divided into *rule based*, *probabilistic*, *learning* and *evolutionary* systems. The boundaries are a little fuzzy as will become evident. These are often considered in terms of the extent to which a good balance of novelty and structure can be achieved. This may be a little over simplistic, but provides a good starting point from which to compare the characteristics of algorithms. In the second half a range of more experimental approaches is given. In current practice, there are a vast number of musicians using all sorts of personalised methods, many of which are never discussed, so the examples are primarily representative of those presented in academic research.

3.1.1 AI Methods of Embodying Musical Knowledge

Algorithmic composition systems are often compared according to the potential for achieving a suitable balance of structure and novelty (e.g. Todd and Werner (1999)). A random

---

[1] These descriptors are scale, process-time, idiomaffinity, extensibility, event production, sound source, and user environment.

number generator is an algorithm, but by itself has little structure[2]. The procedure:

```
IF note = A then play B, ELSE IF note = B THEN play A; note = A
```

is an algorithm but such an incessant trill, although structured, would get a little dull after a second or so. The attraction of many AI techniques then, is that they can not only be used to embody, or derive musical 'knowledge', but this 'knowledge' can then be used to generative a variety of different examples within the bounds defined by it.

Explicit Rules

The simplest form of algorithmic composition comes in the form of explicit sets of rules. These may be simple sets of transformations used to organise pre-existing material, or more complex sets of rules seen to embody aspects of a specific style (for a review see Loy (1991), Todd and Loy (1991)). The *Musikalisches Würfelspiel*, a popular parlour game of the 18^{th} century is often used to illustrate the former. A set of Minuet and Trios are composed, lets say twelve. Each bar is interchangeable across sets. The players then throws a pair of dice, once for each bar of the piece. The original material is composed such that any combination of individuals bars works harmonically, giving the players the joy of composing their own variations. Such an approach obviously guarantees a well formed result but leaves little room for novelty.

William Shottstaedt's automatic species counterpoint program writes music based on rules from Johann Joseph Fux' *Gradus ad Parnassum*, a counterpoint instruction book from the early 18th-century aimed at guiding young composers to recreate the strictly controlled polyphonic style of Palestrina (1525-1594) (Grout and Palisca (1996)). These 'guidelines' are then implemented in a formal language which allows some probabilistic deviation.

> "The program is built around almost 75 rules, such as 'Parallel fifths are not allowed' and 'Avoid tritones near the cadence in lydian mode.' Schottstaedt assigned a series of 'penalties' for breaking the rules. These penalties are weighted based on the fact that Fux indicated that there were some rules that could never be broken, but others did not have to be adhered to as vehemently. As penalties accumulate, the program abandons its current branch of rules and backtracks to find a new solution" - Burns (1997)

Some see such programs as helpful in advancing a more generalised understanding of compositional practice, but in practical terms such programs leave little room for flexibility, and the adherence to such strict rules in actual practice is questionable (see Section 3.2.2).

Grammars

Various forms of grammars have been used in musical analysis as a means of representing hierarchical structure (e.g. Schenkerian analysis or Lerdahl and Jackendoff's (1983) Generative Theory of Tonal Music). When implemented as a computer program, grammars can be used to generate musical material within the specified structural form. The generation of traditional Jazz progressions has been tackled using a grammatical approach. Steedman (1984) devised a generative grammar for chord progressions in twelve-bar blues and refined it using categorial grammars (Steedman (1989)). Johnsson-Laird (1991) also applied grammars to the generation of Jazz chord progressions and bass line improvisations.

---

[2] Random number generators are probably the most widely used devices in Generative Music practice, but generally in the context of some other algorithm.

The grammars themselves can be derived by hand, or created automatically. For example David Cope's *Experiments in Musical Intelligence* project, which focuses on the understanding of musical style and stylistic replication of various composers (Cope (1991), Cope (1992)), uses pattern matching to extract the 'signatures' from two or more existing works. Augmented transition networks are then used to rearrange these patterns into meaningful structures.

Probabilities and Stochastics

Combining a rule based approach with some form of probabilistic reasoning gives greater scope for variation in output. Arguably some of the most successful examples of algorithmic composition are based on probabilistic processes which have designed intuitively by seriously musical minds. Although the inner workings of George Lewis' *Voyager* system are somewhat shrouded in mystery, in talks and discursive papers (e.g. Lewis (1999)) he outlines that the generative behaviour of the system is developed essentially from white noise, which is shaped and filtered with a series of stochastic rule sets. These stochastic processes have been designed and cunningly ordered by his intuitions as a Jazz improviser:

> "For the Voyager program, in addition to the idea of state, it is the sheer number of decisions, as well as their character and order of preference, that leads to a sense of directedness in the music that belies its humble origins in white noise." - Lewis (1999), p. 108

Similarly Nick Collins' Squarepusher patches, released as part of his BBcut SuperCollider library (Collins (2003a)) do an incredibly convincing job of emulating the intricately hand carved break-beats of Tom Jenkinsen, to the point where many cognoscenti of the breakcore universe adopt these procedures rather than sweat over sequencers with the GUI magnification turned up high.

	C3	D3	E3	F3	G3	A3	B3	C4	
C3	0.2	0.2	0.2	0.0	0.2	0.0	0.0	0.2	If C3, then C3, D3, E3, G3 or C4
D3	0.33	0.0	0.33	0.0	0.33	0.0	0.0	0.0	If D3, then either C3, E3, G3
E3	0.0	0.5	0.0	0.5	0.0	0.0	0.0	0.0	If E3, then either D3 or F3
F3	0.33	0.0	0.33	0.0	0.33	0.0	0.0	0.0	If F3, then either C3, E3 or G3
G3	0.25	0.0	0.0	0.25	0.25	0.25	0.0	0.0	If G3, then either C3, F3, G3 or A3
A3	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	If A3, then B3
B3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	If B3, then C4
C4	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0	If C4, then either A3 or B3

Table 3.1: Example of a first-order Markov transition matrix (right) derived from probabilities specified in rules (right) for a hypothetical melody in C major.

Stochastic processes have also been widely employed. Stochastic processes evolve in time according to probabilistic law. One of the most widely explored are Markov chains. Markov chains have a discrete number of steps, with probabilities governing the transition from one to the other. These can be designed by hand. For example if we wanted to make a first-order Markov transition matrix to generate a melody in the key of C (Table 3.1, left), we could define which notes we could move to given the current note (Table 3.1, right) and then convert these into probability chains for each note. A second-order process would take the last two notes into account to define the subsequent note and so on. Simple applications include Brooks et al. (1993) who use matrices derived from an analysis of a series of traditional hymns to generate hymn-like melodies. A more sophisticated implementation was carried out by Ames and Domino (1992) who generated music in specific styles using Markov chains with transformation probabilities created from analyses of existing musical arrangements.

Rather than deriving probabilities by hand, learning systems have been employed to extract statistical information about the structure of a given set of inputs. Loy (1991) and Jones (1981) give early examples where analysis of overall pitch-transition probabilities is performed on a collection of set musical examples. Based on how often particular pitches follow each other in the examples, new compositions can be constructed with similar statistical structure. This can also be done in real time. *Jam Factory* (Zicarelli (1987)) for example, allows the user to train four 'artificial musicians' using real time MIDI input. Analyses are performed on notes and durations to create transition tables which are then used to generate music with the same transition probabilities. A similar technique is developed in Francois Pachet's *Continuator* (Pachet (2002)). This uses harmonic analysis of performance data and applies harmonic constraints on the material generated by the Markov engine to create a computer response which is impressively close to the performer's style with respect to phrasing as well as harmonic and rhythmic motifs.

Neural Networks

In theory, feed-forward and recurrent artificial neural networks (ANNs) are capable of learning more abstract, or longer term patterns than Markov-processes. This potentially allows higher level of musical structure to be preserved in the output (see Todd (1900), Todd (1989) and Todd and Loy (1991) for early examples of musical applications).

Very briefly, ANNs are connectionist models of information processing which are loosely inspired by the neural network structure of the brain. They consist of interconnected processing units which send signals to one another, changing levels of activation according to the sum of their incoming signals. They are commonly used as non-linear statistical data modelling tools, or for finding patterns in data as different learning techniques enable them to model complex relationships between inputs and outputs. (For a good introduction see e.g. Gurney (1997), and for a more detailed excursion, Haykin (1999)).

In musical applications, ANNs hold promise of being able to extract structural information which can then be used to re-synthesise music bearing a similar structure. Attempts have been made to model particular styles directly, as in Tovainen (1995) who worked on Jazz improvisation, and Hornel and Dagenhardt (1997) where a similar approach is adopted for the emulation of baroque melodic improvisation. Others have taken a more indirect approach, aiming to capture specific features of music. Melo (1988) for example used two cooperative ANNs operating on different levels in an attempt to capture harmonic tension in music. 10 listeners were asked to listen to the last movement of Prokofiev's 1st Symphony and to indicate their estimation of dynamic musical tension by pushing a sprung wheel. This data was split between use as training sets for the ANN and as means of testing their results. After training, the ANN's could predict quite well the tension of an unseen part of the piece. However music generated using the trained system was reported to be 'not as successful'.

Despite initial excitement, this application of ANNs does not seem to be living up to its promise: Mozer famously wrote of his own system that it produced "Music only its mother could love" Mozer (1994). One of the problems is that whilst formally ANNs are capable of extracting pertinent structure and generalising from it, it is often hard to predict exactly *what* structure will be learned: what is pertinent to the musical ear may not be statistically most pertinent.

Having said this, increasingly researchers are thinking more carefully about how

ANNs can be applied to musical situations. For example Ollie Glass' *Breakage*[3] project uses a feedforward ANN to generate breakbeats in-the-style-of the user. The nets are used to extract common patterns from a set of user-supplied drum patterns. The system then generalises, creating a variety of drum patterns bearing similar structure. This approach represents an exciting new take on automation in computer music composition which is further discussed in Section 3.3.

Evolutionary Approaches
Evolutionary computation (EC) techniques, such as genetic programming (GP) (Koza (1993)) and genetic algorithms (GAs) (Holland (1975)) model the Darwinian conception of evolution by random variation and natural selection and act as powerful search mechanisms. Various forms of EC have been extensively explored (see Burton and Vladimirova (1999) for a good overview), spawning a sub field of Evolutionary Art and Music (e.g. Bentley and Corne (2001)).

With respect to the structure/ novelty tradeoff, the approach is promising. The process of artificial evolution itself relies upon the accumulation of 'good solutions' which gives a degree of continuity, variation being provided by the recombination and mutation operators. The problem comes in ascertaining what a 'good' solution is. In natural evolution the criterion is ultimately the ability to survive and reproduce. In engineering applications, or optimisation problems, although non-trivial, it is possible to define a fitness function that can potentially be fulfilled by members of a population. In many music situation there is rarely a unique optimal solution, rather an ongoing process of innovation and refinement. This is arguably something that evolution, and artificial evolution systems excel at, however the problem of defining some measure of fitness criterion or selective pressure remains.

Several approaches have been developed to tackle this problem: some employ theory-derived rule sets to define operators, or desired patterns in output (e.g. Horner and Goldberg (1991)), or introduce a learning module to train an artificial critic (e.g. Spector and Alpern (1995)). One of the technical problems with simple GAs is premature convergence: the population can get stuck in a local minima. In addition, even if a solution is achieved that satisfies the fitness function no further evolution will occur. With no other pressure to do anything else, no further innovations will result. This can be a problem especially when the the population finds a loop hole in the fitness function, producing solutions that meet the criterion, and dominate the population but do not fulfil the goal (i.e. interesting music) that the function was intended to describe.

One solution that circumvents the difficulty of having to come up with a formally defined fitness function which has been popular in art and music applications is to use an Interactive Genetic Algorithm (IGA). In this case the fitness function is simply the subjective preference of a human listener. This is the technique used in the 'breeder art' described in the last chapter. However this approach is not without problems. GAs necessarily require a large population of solutions, each of which must be assessed. This is commonly referred to as the 'fitness bottleneck'. Some solutions to this include the development of hybrid systems, with multiple critics that include rule-based or trained critics operating on a lower level. This has proved successful in terms of reducing the number of examples which must be screened by a human listener (Biles (1994)).

Coevolutionary Approaches
An alternative technical means of circumventing premature population convergence, which does not require human assessment, is to coevolve critics and solutions. It has been shown that coevolution reduces the chance of solutions finding easy ways to 'trick' critics (Hillis (1992)). Coevolution can also produce diversity within a population. This

---

[3] `http://www.blackholeprojector.com/`

synchronic diversity can be generated through sexual selection leading to speciation - splitting the population into subpopulations of individuals with distinct traits and preferences (see e.g. Todd and Miller (1991)). Coevolution can also lead to diachronic diversity, producing traits in a population which change over time as in predator prey 'arms race' models such as described in Futayama and Slatkin (1983). This has been exploited by Werner and Todd (1997) in their model of the evolution of bird song.

These interactions between critic and composer in the coevolutionary approach is seen as a proto-social behaviour, and has been described as a 'cultural approach'. Miranda has produced a simulation in which a society of distributed, autonomous, cooperative agents evolve sound repertoires from scratch by copying each other (mimesis) (Miranda (2002)). The agents produce musical signals to which other agents react, influencing the signals they produce or their proclivity to mate. He argues that as the music created in the system influences the behaviour of the other agents living in the system, it gives it a social role and that the corresponding increase in the richness of the dynamics promises a musically more interesting output. The problem with this is that as a simulation, the system is cut-off from the real-world. Whilst the system may be argued to have cultural significance, culture references will accord with an agent, rather than human aesthetic. In this sense the approach is more amenable to scientific modelling rather than composition.

3.1.2 Techniques Employed in Exploration of New Musics

The application of AI methods described above frame composition as a 'problem solving' exercise aimed at capturing the essentials of existing musical idioms, (Cope (2005), Jacob (1996)). Other practitioners have adopted a more exploratory approach, using mathematical models or data drawn from other sources to explore new musical possibilities. The idiosynchratic nature of this approach makes it impossible to give a fully comprehensive review of all the techniques which are used, but this section aims to provide a representative overview. There may be many more innovative and effective algorithmic procedures in use which are never published or even described in words - primarily because discussion of compositional techniques is not considered relevant for many composers and producers (or maybe they are just so darn good they don't want anyone else to know).

Many of the techniques described above can be adapted to create new music. For example it has been proposed that you could train an ANN or derive Markov tables from two quite different musical styles (e.g. Tuvan throat singing and Black Sabbath) to create strange hybrids. In practice it is likely that the results will be a statistical rather than musical hybrid. In many instances composers have taken an established technique and applied it in unconventional ways. Many different algorithms have been explored, but by far the most common method of employing them is to directly sonifying the numerical outputs of the selected model. Where models are appropriated from existing mathematical or scientific theory (which is common), these tend to be models of growth, pattern formation or movement.

Evolutionary Approaches

Many of the techniques discussed above can be implemented in other ways which take advantage of the *structure* in the process of the algorithm, rather than its problem solving capabilities. In some instances the adaptations made to algorithms are such that its original function is no longer even preserved. For example Waschka II and Cristyn Magnus both used GAs, but rather than employing them as a search mechanism and 'listening' to the winning individual at the final generation, the process itself is sonified, conveying the changes that occur in the population across generations.

Waschka II used this technique to create one of the arias in his opera *Sappho's breath*.

The initial population was small and consisted of Greek and Medieval songs, and melodies that he had composed himself. There was no fitness function (and so no evolution in the traditional sense), individuals being randomly selected to form the next generation. The programme was run for only five generations. Even within this small population Waschka reports that there is a clear diminution of diversity in the population and a tendency towards replication of one or a few individuals over very few generations. These characteristics, i.e. the reappearance of individuals that survive unchanged across generations, and the population convergence, are used to create psychologically meaningful musical effects.

> "As the piece proceeds, the repetition of certain elements or whole individuals allows the listener to make connections with other parts of the work. Finally, the increased similarity of the musical material and the repetitions of motives or measures can provide for both the composer and the audience a sense of closure, and help to bring the piece or section to a natural and satisfying end."
> - Waschka II (2001), p.5

More traditional implementations of GAs have been applied as a means of setting synthesis parameters in complex DSP chains. For example *AudioServe* from Yee-King (2000) is an online interactive evolutionary synthesis tool. This uses an IGA to evolve frequency and amplitude modulation (FM/AM) circuits. A similar technique is employed by Palle Dahlestedt's *MutaSynth* (Dahlstedt (2001)) which can be applied to arbitrary synth engines.

Dahlestedt also experimented with simple coevolutionary programs to generate musical material. *Living Melodies* (Dahlstedt and Nordahl (2001)) uses a spatial multi-agent evolutionary model where populations of singing creatures wander about on a discrete virtual plain. Each agent's genome specifies their actions (WALK, TURN, REST, SING) and also includes 'IF' and 'LOOP' statements, allowing these actions to be combined (loops and ifs cannot be nested). Each agent moves around the discrete 2D space according to its specified actions. Reproduction occurs between agents which are spatially adjacent, have enough 'energy points', are of sufficient age, have walked about recently and have 'heard' some singing recently. A child is produced by cross-over, and genome lengths are dynamic. In the simplest mapping, each agent sings its specified note (voiced on a MIDI piano), which is determined by its species. This produces pulsing patterns which vary as species wax and wane. In order to thin out the cacophony of large populations, they also experimented with applying filters such as constraining note amplitude with 'listening pleasure' or energy points. What we hear in effect is a representation of various aspects of the population dynamics.

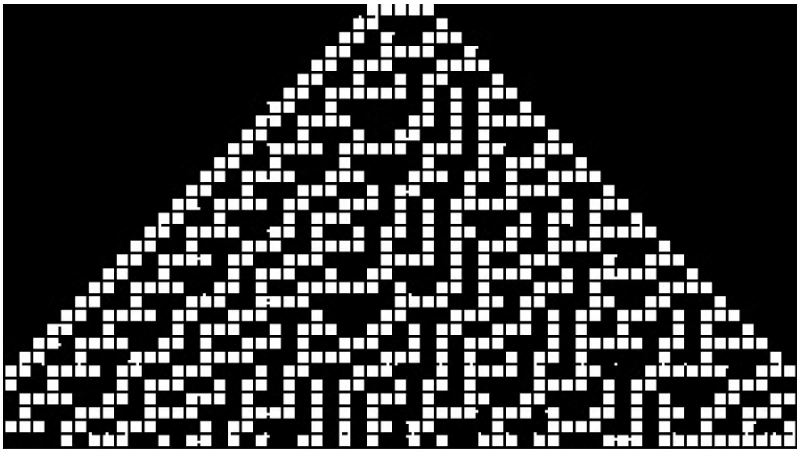
Cellular Automata

Evolutionary techniques are generally presented under the heading of Alife music alongside Cellular Automata[4] (CA). CAs are discrete dynamical systems operating on a regular lattice which change state according to rules based on their local interactions, creating self-organised patterns. In this respect they are comparable to the Swarm model used by Blackwell. The dimensions of the lattice and the number of possible states can vary. A one dimensional (1D) CA with binary states is depicted in Figure 3.2 (top) where each successive horizontal row represents the states of each cell at consecutive time steps. The state of each cell at time *t* is determined by the state of the cells in its defined neighbourhood

---

[4] In fact many prominent authors seem to conflate the two. Miranda for example repeatedly in papers and books refers to CAs as evolutionary, and Cope even describes CAs as "A form of GA". Here the term evolutionary and GA are reserved for models which deploy some form of fitness based selection and reproduction operators to specify the evolution of a population.

according to a sets of rules (Figure 3.2 (bottom)). In this example, the neighbourhood of each cell are the three cells immediately above it. Different rule sets produce different types of patterns; some very repetitive, some chaotic and some complex.



neighbourhood states at time (t-1):	111	110	101	100	011	010	001	000
cell state at time (t):	0	1	1	0	1	1	1	0

Figure 3.2: Graphical representation of a one-dimensional CA and its rule set (bottom).

Perhaps due to their inherent pattern propagation properties, CAs have been extensively explored by many researchers and composers as a basis for generating musical material (for a recent review of applications in MIDI based music see Burraston et al. (2004)). Bilotta and Pantano (2002), draw parallels between the development of variations on a theme, and the fact that CA rules can be run from different initial conditions, producing different, though correlated numerical sequences which can be transformed into musical passages. Others such as Miranda (1993) have applied CAs in the synthesis domain and suggest that they are more suited to specifying timbral rather than structural aspects of music (details of this are discussed in Chapter 4).

L-systems

Another model of structural growth that has been frequently implemented are L-systems. L-systems are generative grammars developed by biologist Prusinkiewicz and Lindenmayer (1990) to model the growth of plants. These have been used by many composers (e.g. Prusinkiewicz (1986)) as a method of generating or organising fragments of musical material. The attraction of such systems lies in their ability to model complex growth processes that, certainly graphically, develop in "aesthetically challenging opposition" (Supper (2001) p.50). In addition, non-predictable, often self-similar patterns are produced which would be hard to specify by hand. Examples of the sorts of branching structures which can be generated are given in Figure 3.3.

Statistical Models

Beside statistical models of musical structure, many other probability-based models are commonly used – most famously explored by Iannis Xenakis. Xenakis' background as an engineer and architect gave him a radically different perspective on music from his classically trained contemporaries. He used probabilistic models drawn from areas such as thermodynamics and game-theory to realise his fascination with the 'out of time' structures of the ancients. One of his best known works, *Pithoprakta* (1955-1956) is essentially a sonic incarnation of the Maxwell-Boltzmann kinetic theory of gases. The Boltzmann distribution can be used to simulate the speeds of gas particles. Xenakis mapped these onto

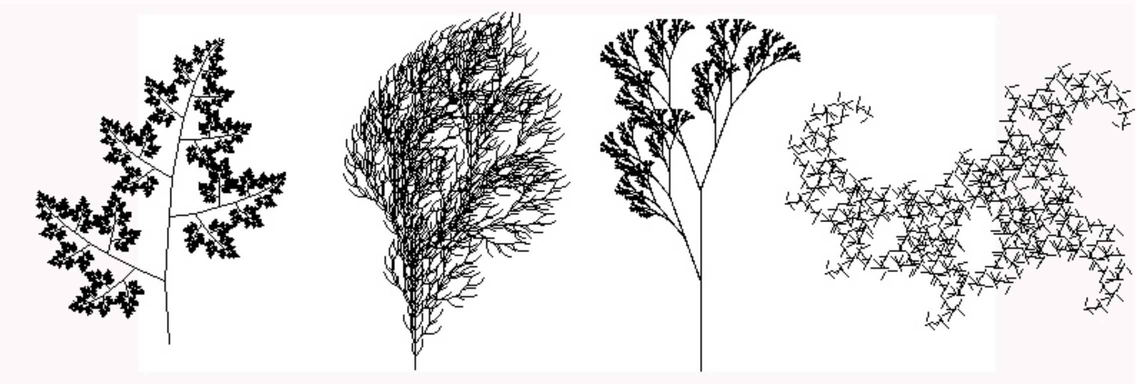


Figure 3.3: L-system generated plant forms.

tuning pitches for glissandi across different orchestral instruments. By slowly changing the temperature parameter in the gas simulation, he created a broad, slowly changing sonic cloud.

Chaos

Following the mathematics communities' enthusiasm for chaos theory, many researchers in the late 1980s and early 1990s explored the musical potential of fractal geometry and chaotic non-linear dynamical systems (e.g. Pressing (1988), Di Scipio (1990), Gogins (1991), Bidlack (1992), Leach and Fitch (1995)).

These are sometimes referred to as iterative functions due to the fact that they take the form $y_{(t+1)} = F(y_{(t)})$. Motivation for adopting these systems is often given in the form of a description of the existence of chaotic patterns in natural phenomenon such as the weather, population cycles of animals, the structure of coastlines, trees or leaves, bubble-fields or the dripping of water. Others point out similarities between the dynamics they exhibit and common musical structures:

> "Musical development or variation can be viewed as the transformation or distortion of a simple entity (a motif), often followed by some sort of return to the original motif. When certain values are chosen for the input parameters to these equations, very similar behaviour can be obtained from them. Thus a series of solutions can act like a repeated group of n notes for a number of steps in the iteration process, and then break away to more unpredictable (quasi-chaotic) behavior before eventually returning to the original n-note group, perhaps somewhat altered." - Pressing (1988)

Specific nonlinear dynamic equations that composers have experimented with include: the logistic map which is traditionally used to model a species' change in population; the Henon map which was originally introduced as a simple and efficient model of chaotic systems in general; and the three-dimensional Lorenz system which was developed from a simplified model of atmospheric turbulence.

Four pioneers of these methods are Jeff Pressing, Michael Gogins, Rick Bidlack, and Jeremy Leach. Gogins (1991) worked with what he called Iterated Functions System, a collection of functions which were selected from at each iteration giving an incredibly (perhaps excessively) broad selection of complex dyanmics. Bidlack (1992) explored higher dimensional dissipative chaotic systems as well as 'conservative' chaotic systems in which energy is conserved, producing a constant orbit rather than transience toward an attractor.

Pressing (1988) focused on the edge of the chaotic regions, where dynamics alternate between quasi-periodic and chaotic attractors: "The output shows unpredictability, but also traces of the nearby cyclic behavior ... In musical terms, the overall effect is like a variation technique that inserts and removes material from a motive undergoing mildly erratic pitch transformations, in the style of an adventurous but development-oriented free jazz player, perhaps" (ibid. p.4). Maurer (1999) compares this to techniques of free improviser Ornette Coleman, who is stylistically notorious for improvising "motivic-chain-associations" (Jost (1994)).

In all these cases, the outputs of equations were generally used to define (MIDI) pitch information in a generally monophonic stream of notes. Pressing (1988) also experimented with two dimensional maps, mapping each dimension to two different characteristics of a musical event (e.g. pitch and length of a note), but reports that these results sound no better than the single dimension.

This approach was not without critics. Truax in particular was scathing about the use of non-linear functions at this level of musical form:

> "From a more philosophical or aesthetic point of view, it is not clear than an arbitrary mapping of a non-linear function [onto the pitch of successive notes] is inherently more musical than, for instance, a random or stochastic function. The musicality may reside in the musical knowledge of the mapper more than in the source function. The audience, if suitably primed with program notes, may be convinced there is more value or interest in the result because of the technique used, but the half-life of such interest seems to be short."
> - Truax (1990)

Despite his scepticism for using models at the level of pitch, Truax (1990), and later Di Scipio (1999) applied similar models to sound design tasks. They both experimented with models such as the logistic map to parameterise granular synthesis engines. The models were used to control a granulation process operating on a source sound and used to create chaotic textures. The degree of chaos could tuned, creating varying degrees of deviation from the original source sound: "Depending on the 'degree' of granulation on the equation system and on the region of the relative logistic map to be explored, the output sounds are clearly derivative from the original or completely extraneous to it" (Di Scipio (1990)). These equations were used to reproduce the dynamic progression of natural sounds such as increasingly heavy rain fall, or spreading fires.

There is a strong belief amongst both critics and practitioners that much of the creativity in using algorithms, the real compositional work, comes in defining the mapping that is used. This is true to an extent but the issue is more complex. In this case Truax expressed disdain at the use of 'arbitrary' mappings of non-linear systems to define musical pitches, but enthusiastically deployed very similar models in the synthesis domain. Whilst the details of the mappings used are of course vital, ensuring a good marriage between the algorithm and the level of musical organisation is an equally important consideration. This issue is discussed in more detail in Chapter 4, Section 4.4.

Summary

As this overview suggests a range of techniques, drawn largely from AI and other scientific endeavours, have been explored which could potentially assist in many different aspects of the compositional process. As Miranda noted, some are very good at generating music in their target style. Having said this surprisingly few techniques, either AI or extra-musical, seem to actually get used. In the case of AI techniques this may well be because until recently, the implementation of much more than a random number generator or probability model demanded sophisticated programming know-how which was perhaps beyond the average composer. Even when implemented, some require significant technical knowledge to set up and use: ANNs and GAs for example require a good understanding of their operation in order even to be able to pre-process and encode the input data in a suitable way. Such an understanding is also crucial to being able to tune them if their initial performance is less than perfect.

Information processing techniques have revolutionised many areas of engineering and data management, but musical composition is a very different task to the pattern recognition or optimisation problems in which these techniques excel. Within AI music literature many examples suggest that the compositional 'problem' has not been framed in quite the right way. Testimony to this are a number of reports of a mismatch between formal and aesthetic results. For example in their exploration of GAs in a harmonisation task, Papadopolous and Wiggins (1999) report that "The solo generator GA converged very quickly to high fitness" (p.11). However they also add that: "from an aesthetic viewpoint, the results are still far from ideal." (p.12).

Perhaps another reason that these techniques are less than ubiquitous in the musical community is that research tends to focus on the machine learning capabilities in the abstract, rather than their application in a specific musical domain. This is to be expected where algorithmic research is carried out as a computer scientific project, but as discussed below in Section 3.2, these divisions are not always made clear. Increasingly, those focused on music making are considering how these tools can be altered for more creative use, rather than applying them straight off the shelf. Al Biles for example focuses on EC techniques and contrasts the musician's creative perspective with the engineering perspective, which is more prevalent in EC circles and suggests how the musician's perspective alters how EC is applied in musical domains (Miranda and Biles (2007) , Chapter 2).

The dominant computer science/ AI research perspective may be one reason why the use of so called 'extra-musical' algorithms gets somewhat sidelined in academic writing. Beside Miranda's comment on the lack of 'cultural references', Papadopolous and Wiggins (1999) write that "It is difficult to judge the quality of their output, because, unlike all the other approaches, their knowledge about music is not derived from humans or human works" (p.12). In their context, judgement may refer to some form of empirical measure, which is indeed hard to define. But as Pearce et al. (2002) later suggests, the value of any algorithmic approach is best 'judged' in the same way that other music is, i.e. by a listening public.

It is difficult to know to what extent specific or similar techniques are adopted by performing musicians[5] (many of whom may not have access to the academic journals in which they are predominantly published). However, it is generally agreed that few formal models explored within the research community have been applied to anything more than generating small fragments of musical material which if used at all, are incorporated by hand into larger compositions. As a compositional aid, this is fine, indeed many composers view such processes as useful ways of throwing up new material to inspire new compositions. If we aim to develop generative systems for live performance however, we need something a little more powerful.

3.2 Aims and Inspirations: Ambiguity in Cross-disciplinary Enterprises

If the range of techniques, implementations and applications makes the field seem a little schizophrenic, it is because it is. Interdisciplinarity is trendy and increasingly encouraged in academic communities, and an operational interdisciplinary research pro-

---

[5] Allegedly Aphex Twin has employed some examples of generative music packages that are distributed with Miranda's 'Composing Music with Computers' (Miranda (2000a)) (Miranda, personal communication).

gramme would be welcome in algorithmic composition. But arguably, the field has not yet reached this level of maturity, existing more as a collection of cross-discipline projects situated in different fields. This section aims to highlight a few of the issues that arise in such situations.

In many academic papers describing algorithmic composition projects, authors, particularly those aiming to replicate existing musical idioms in some way, often present ambiguous and sometimes opposing representations of what they are aiming to achieve. Biles for example introduces *GenJam* – a GA which improvises jazz melodies over a supplied chord progression – as "a genetic algorithm based model of a novice jazz musician learning to improvise" (Biles (1994) p. 131). This makes it sound as if he has a cognitive scientific interest in understanding a process of musical development. In the final paragraph however, he writes that "GenJam shows that GAs can be a useful tool for searching a constrained melodic space"(ibid p.137), suggesting that GenJam is a compositional or performance tool.

Further confusion is added when authors write up the same systems in different contexts and submit them to journals in different fields. Todd and Werner for example, in one of the most cited reviews of algorithmic composition, present their coevolutionary model as a method of composition:

> "Thus, we propose a new algorithmic composition system based on a combination of all the approaches we have described here: rule-following, learning, and evolution." - Todd and Werner (1999) p.185

Exactly the same model is presented as a coevolutionary model of communications in an Alife context (Werner and Todd (1997)).

Similarly, Miranda's agent-based mimetic model appears in numerous computer music publications in the context of other sound generation/composition techniques, but it also appears as a model of the cultural evolution of music. For instance in his book *Composing Music with Computers* (Miranda (2000a)) he writes: "The adaptive musical games approach to generating musical materials for pieces of music seems a sensible way forward in order to break new ground in algorithmic composition practice." (ibid p.157)

The same system is presented in a paper entitled *Towards an Artificial Life Approach to the Origins of Music* (Miranda (1999)), in which he describes the contribution of Alife simulation to theoretical biology and proposes that "the Alife paradigm also has great potential for Musicology." (p.1), and presents his model as an investigation of "how musical forms may originate and evolve in artificially created worlds inhabited by virtual communities of musicians and listeners. " (ibid p.2). This is not perhaps a problem, if each is suitably assessed in each respective domain, but as Pearce et al. (2002) have pointed out this is not always the case.

3.2.1 Implications for Research Methodologies

Pearce et al pick up on the motivational ambiguity within the field and suggests that it is symptomatic of a methodological malaise. To remedy what they fear may hamper research progress, they propose that the field should be subdivided and propose distinct labels and methodological and evaluation procedures for each subdivision.

A principal distinction is proposed between creative and scientific projects. These are further bisected. The term 'algorithmic composition' is used to describe projects aimed at expanding the compositional repertoire, as distinct from the 'design of compositional tools'. Scientific projects they suggest should also be divided according to whether systems pursue the musicological aim of evaluating stylistic theories or attempt to make contributions to cognitive science by modelling the cognitive processes that underlie composition.

Not unreasonably, they suggest that each requires quite distinct approaches to evaluation. Creative applications, where the ultimate output is intended for public consumption they suggest can only be evaluated "in the same way that composers and compositions are usually appraised: through audience reactions at performances, record sales, critical reviews and so on" (ibid p.5). For many years this was an all-too-rare event. Any form of objective appraisal of musical effectiveness in papers tend to offer rather vague subjective judgements such as "[the program] seems to be capable of producing musical results" (Ebcioglu (1988) p. 49).

They also note that there is a lack of evaluation from those offering their systems as cognitive science models. A cognitive theory is evaluated by implementing it in a model and then making empirical comparisons between the behaviour of the model and the human behaviour that the theory is intended to explain. Granted little is know about the cognitive mechanisms underlying composition, but those who have attempted to model aspects of this process in cognitive terms pay little attention to the basic scientific modus operandii. For example, the only evaluation that Johnsson-Laird (1991) offers for his model of Jazz improvisation is that "The program performs on the level of a competent beginner." (p. 317).

Pearce's suggestions are relevant to individual disciplines using algorithmic composition methods for specific goals and they promote more rigourous application of these established methodologies. It would undoubtedly be easier for us as readers and researchers to judge the value of individual contributions if authors were a little clearer in stating their aims, intentions and evaluations. But there is also an argument that the field needs to look outside of its constituent disciplines to develop new methodologies. Computer assisted algorithmic composition is a relatively young endeavour which straddles speculative and empirical disciplines as well as arts practice. This is what makes it exciting, as well as slightly confused. Too close an allegiance to its parent methodologies could also be restrictive. It seems reasonable to encourage purposeful cross-fertilisation of ideas which may be a crucial ingredient for success. For example, in yet another paper describing his mimetic model of music evolution, Miranda proffers:

> "Should such experiments with adaptive games corroborate our hypothesis that we can improve algorithmic composition systems considerably by including mechanisms that take into account the dynamics of cultural evolution and social interaction, then we believe that a new generation of much improved intelligent composing systems will soon begin to appear over the next few years." - Miranda (2002)

Clearly defined speculative enquiries are central to the development of new methodologies and activities are arguably vital for the maturation of a nascent interdisciplinary project such as algorithmic composition.

3.2.2 Implications for Software Design

I would like to suggest that there is a more worrying side effect of cross-disciplinary projects which impacts on the use of algorithmic composition as both a practical creative pursuit and as a method of scientific modelling. These problems stem from the adoption of ideas and theories from one discipline into another practice without careful consideration of their true nature. The adoption of AI techniques, invariably, although not inevitably, is accompanied by the framing of composition as a 'problem' that needs to be solved. This is implicit in much work, and utterly explicit in others (Cope (2005), Jacob (1996)). This is fine in itself: we could agree to confer that some aspects of composition could be usefully framed as a problem solving task. In the context of AI, music becomes an engineering problem, to be solved by a search mechanism, constraint satisfaction procedure or pattern matching algorithm. In both engineering and scientific research applications, such techniques are developed into computational models. Models implement theories which are generally derived from experimental observations of the phenomenon under study. In a scientific context, the theory is then evaluated by empirical comparison of the output of the model and the natural phenomenon. In an engineering context the hope is that this theory is sufficiently powerful that the model simulates the natural world phenomenon.

As noted in the last chapter, this approach can be very effective in related areas of machine listening, where theories of human auditory perception and cognition provide the starting point for beat-induction or pitch discrimination algorithms (e.g. Collins and Cross (2003)). In algorithmic composition, for either scientific or compositional ends, Western Musicology offers an attractive body of theory which can be embodied as a rule-base or set of constraints in a similar fashion. As well as explicit encoding in Knowledge Based systems, theoretic principles can be embodied more implicitly in evolutionary or learning algorithms.

At first sight, this offers a means of imbuing the necessary domain knowledge to ensure musical structure in the output. The rules of harmony, for example, offer an apparently neat solution to the headache of designing a fitness function for a GA; spatial representations of tonal distance (e.g. Schoenberg (1954), Shepard (1982)) could provide a basis of error measurement for an ANN. And indeed this is common practice: the "Fitness function judges the fitness of each chromosome according to criterion taken directly from music theory" (Phon-Amnuaisuk and Wiggins (1999)). The problem, which does not seem to be discussed, is that music is not a natural phenomenon, it does not conform to natural law, and music theories do not serve the same function as theories of biology or physics or even cognition.

Music theory and analysis are post hoc endeavours, which aim neither to describe what the composer did, nor necessarily what the listener hears. For researchers aiming to model the compositional process or develop systems with which to compose music aimed for public consumption, it seems important to keep this in mind. This section serves as a gentle reminder that care should be taken when inspiration is drawn from music theoretic tenants either for creating models of composition or in designing systems for generating music for people to listen to.

The Analyst's Music

> "Writing about music is like dancing about architecture: it's a really stupid thing to do" - Elvis Costello, in an interview by Timothy White (1983) p. 52

Music theories in their many guises, represent attempts to *understand* music in an analytical sense; the analyses from which rules are derived aim to achieve possible coherent sets of principles and ideas with which to rationalise, analyse and investigate the structurally functional aspects of music. This is neither exhaustive, nor aimed primarily at describing music in terms of the listener's perception.

> "Each musical culture rationalises only a few selected aspects of its musical production ... [so] any cultural representation of music (i.e. music theory) must constitute a thoroughly incomplete specification of the intended musical experience... A formal analysis is a kind of mechanism whose input is the score, and whose output is a determination of coherence...In other words, it purports to establish or explain what is significant in music while circumventing the human experience through which such significance is constituted; to borrow a phrase from Coulter, it aims at 'deleting the subject'." - Cook (1990), p.241

Cook's basic argument is that there is an important and inevitable discrepancy between the experience of music aurally, and the ways in which it is imagined or thought about. He draws a useful distinction between 'musical listening' which is concerned with the aesthetic gratification in being absorbed in a non-dualistic sense, and 'musicological listening' for the purpose of establishing facts or the formulation of theories about music as a 'perceptual object'. This is not to say that the thoughts and ideas of theorists and analysts are inadequate or misplaced, but simply that they are not aimed primarily at giving an account of the listener's phenomenological experience of music.

An extreme example of this discrepancy between analytic and experiential reality is provided by experiments in which two versions of short classical piano pieces were played to music students: their original form, which began and ended in the same, or closely related key, and an altered form which had been modified so as to modulate to, and end in, a different and unrelated key (Cook (1987a)). In standard music theory, tonal closure - or more generally the influence, or organising function of the overall tonic - is the very core of the traditional forms of eighteenth and nineteenth century music. However in these trials, there were no statistically significant differences in preference for the original over the altered forms. In fact, there was a general preference for a Brahms Intermezzo in which the final section had been transposed up a minor second.

In another set of tests, Cook (1987b)) played the first movement of Beethoven's G major sonata (Op 49 No 2) to music students, stopping the recording just before the final two chords. He then asked them how much longer they thought the piece continued for and reports that they frequently predicted that the music would carry on for another minute or more. Theoretically, the recapitulation and coda are key informative aspects of form, signifying the close of a piece. These studies suggest that in aural presentations they seem to be slightly less than effective in fulfilling this function. In both these studies, there seems to be a discrepancy between what is seen as functional theoretically, and what actually affects us perceptually.

Perhaps more elegantly designed studies would be needed to make any strong claims, but it is common for musicologists to differentiate between the aural and analytic aspects of a piece of music. Bailey (1983) writes of Webern's symphony that it consists of "two quite different pieces - a visual, intellectual piece and an aural, immediate piece, one for the analyst and another for the listener." (p.195) Clifton (1970) expresses this more incisively: "For the listener, musical grammar and syntax amount to no more than wax in his ears". (p.71)

Within the music community, the differences between analytic treatments and compositional practices are well recognised. The fugue, for example is often qualified with a prefix 'composer's' or 'classroom' to differentiate between the actual fugues written by composers, and the idealised abstractions which appear in textbooks.

> "[The classroom fugue] insists on certain exigencies which place the texture upon the highest level of skill, the actual musical result being, seemingly, of secondary importance. Many noted theorists have set down all the things which should happen in every fugure, but composers have always followed their own dictates. It is impossible to align No.1 Book One of Bach's Forty-Eight, for example, and many others, with the dictates of the theorists. At one time it was not regarded as musical expression but was used solely as a mental exercise." - Demuth (2003)

In some discussion of algorithmic composition, these distinctions seem to get papered over. In his list of historical examplars of algorithms, Cope (2000) writes: "Fugues and other contrapuntal formalisms represent constraints and often require severely limiting algorithms" (p.3). He uses these and many other examples of pre-digital algorithms to support his claim that 'all composition is algorithmic'. But as Demuth notes - in a textbook of musical forms - although a textbook description of a fugue might look like an algorithm, and so be suitable for implementation and realisation by a procedural device like a computer, this is not what composers did, or do. The clean, algorithmic-looking version came *after* the composer's intuitive one.

> "Schemes of musical arrangement, even if they exist a priori, should only be discovered after they have been used." - Schoenberg (1954)

The dangers of too close a marriage between academic theory and the development of compositional algorithmic programs has been expounded by Gartland-Jones and Copley (2005). They focus on the level of form, clearly illustrating their arguments with insightful and deep consideration of sonata form which shows just how artificial the text book abstractions are that researchers such as Cope use as a point of departure to justify algorithmic research. This is perhaps especially important for those aiming to model compositional processes, as these theoretical imprints depart considerably from what composers actually do. For those interested in algorithmic composition for making music, there may be another problem with the blind adoption of music theoretic principles.

The Listener's Music

Music as we listen to it is indefatigably temporal. Music as it is studied on the score lies out of time. The contradiction between the ever-present 'inner' time in which music is experienced, and the retrospective 'outer' time which is imposed in the act of reflection and measured by musical notation is a fundamental dilemma for many theorists (Cook (1990)). Schutz (1976) goes as far as to suggest that attempts to describe musical experience in 'outer' time poses a variant of the Eleatic paradox – i.e. that the flight of Zeno's arrow cannot be described because it is impossible to represent the ongoing quality of its' motion. As he puts it "you may designate the spot occupied by the arrow at any chosen constant during the flight. But then you have dropped entirely the idea of an ongoing motion." (ibid p.30). At the heart of this lies the discrepancy between the static, symbolic nature of music in notated form - which is the principle object of music theoretic concern – and the dynamic immediate nature of music in sonic form - which is the subject of concern for the listener. Many more progressive critics and theorists both within and beyond traditional 'note-based' music discuss the distortive nature of the musical notation, and how it gets in the way of an understanding of our actual listening experience.

> "The principle point I am going to develop is that the priorities of notation do not merely reflect musical priorities - they actually create them. It is fundamentally important to grasp the point if we are going to understand an approach to music based on our listening experience .... A preoccupation with conventional notation can lead us into formalism, a situation where there is no longer any experiential verification of our theories about how to compose music." - Wishart (1985), p.11

Wishart of course works primarily in electronic music and is keen to ensure that new possibilities and ideas aren't constrained by conceptions carried over from archaic idioms. But this is not just an issue relevant to the acoustic-electronic divide. Very similar ideas were expressed years before, by Ernst Toch.

> "I never expected so much fascination to come from investigations of the nature of musical theory and composition. Aspects unfolding to me show why the rules of established musical theories could not be applied to 'modern' music, why there seemed to be a break all along the line, either discrediting our

> contemporary work or everything that has been derived from the past. To my amazement I find that those theories are only false with reference to contemporary music because they are false with reference to the old music, from which they have been deduced; and that in correcting them to precision you get the whole immense structure of music in your focus." - Toch (1948), p.xii

Ernst Toch was a masterful and original classical, and later modernist, composer who was also renowned for his Paramount Studio film scores. Later in life he became preoccupied with the reconciliation of theories of classical music with contemporary modernist trends. *The Shaping Forces in Music* (Toch (1948)) is his account of how all musical writing must respond to the psychological wants of the listener, and how similar goals may be achieved in different styles. If harmonic structure is the cornerstone of traditional music theory, Toch sees the movement of melodic 'impulses' (not dissimilar to Wishart's dynamic morphologies) as the central force of music from the listener's perspective. He describes harmony as 'arrested motion' by which he means to stress the fundamental Heraclitean flux in music.

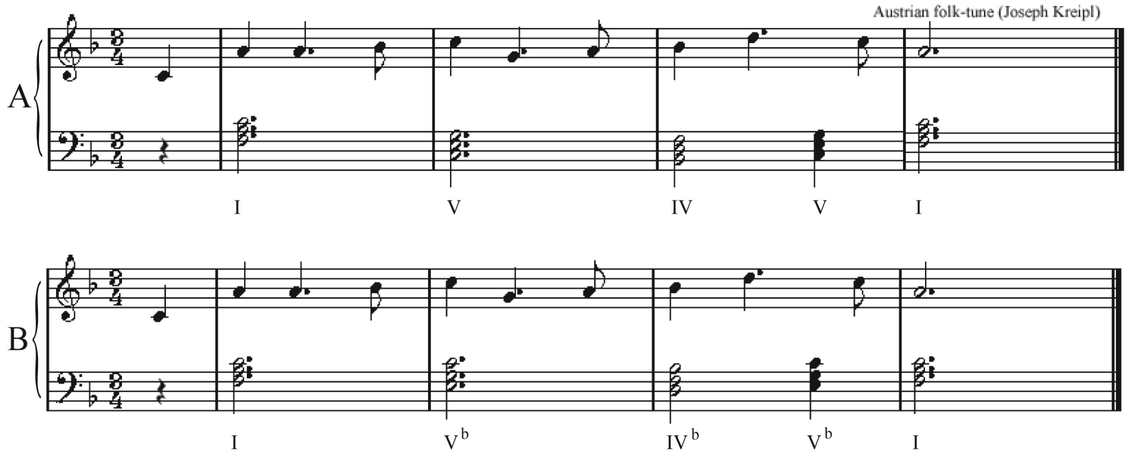


Figure 3.4: A Natural harmonisation of a phrase from a simple folk tune using I (tonic), V (dominant) and IV (sub-dominant) (top) and appropriate chordal inversions (IV^b etc) (bottom).

Harmony as Arrested Motion

In an example that is typical of the tasks used in algorithmic approaches to harmonisation, Toch presents a phrase from a folk tune, that invites a simple I, IV V harmonisation (Figure 3.4 A). This is something that a GA could perhaps achieve. We could even potentially incorporate theoretic axioms for finding appropriate chord inversions into the fitness function: by minimising the number of steps that each note must take into membership of adjacent harmonies we could feasibly find the first inversions needed to create a smoother chordal structure in bars two and three as shown in Figure 3.4 B. The apparent simplicity and efficacy of this kind of 'rule' is precisely what is attractive to the algorithmic composer, but as Toch warns: "While this axiom seems a simple expedient for the beginner, it implants in him a dangerous misconception, namely the view point of rigidly preconceived harmony as a fixed unit, within the frame of which each voice seeks to take up its appropriate place." (p.5).

This point is illustrated by considering a common or garden Chorale harmonisation shown in Figure 3.5 A., which concedes to all the traditional rules of harmony. Toch then

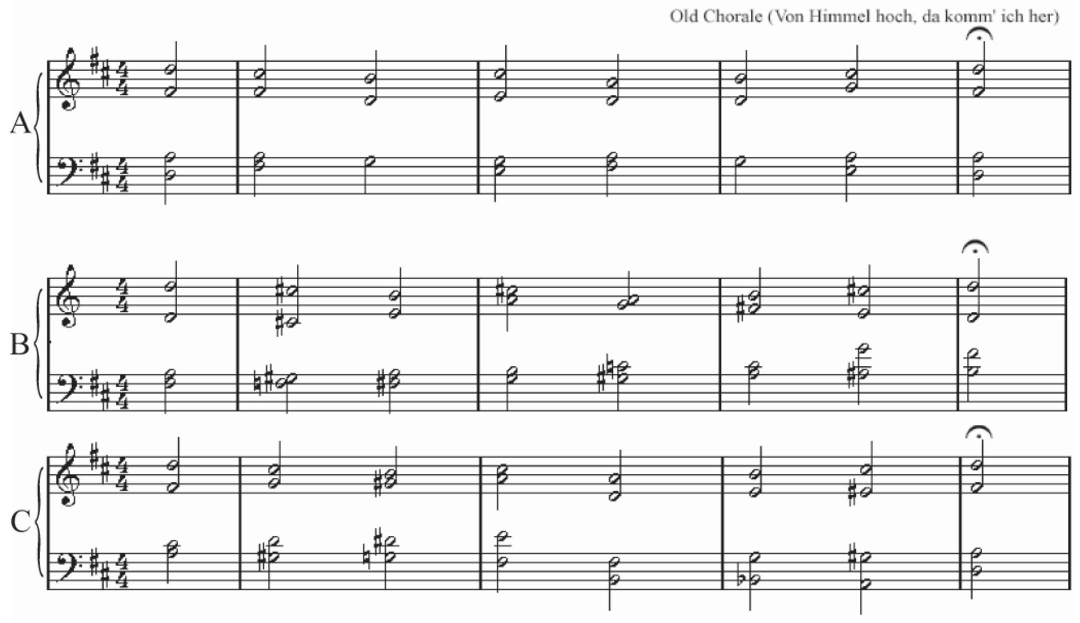


Figure 3.5: Standard harmonisation of a phrase from a Chorale (A) and examples from Toch's alternative harmonisations (B) and (C).

offers twelve other possibilities, examples of which are given in Figure 3.5 B and C. These are written using a more general principle which he calls 'linear voice leading' - a term he uses to describe the dynamic impulse of each voice. In contrast to the 'appropriate' harmonisations of Figure 3.5 A, some harmonies in Figure Figure 3.5 B and C go against every rule in the book: consecutive fifths, cross relations, arbitrary dissonances etc.

> "And yet we hope that the reader, even though these harmonisations may appear unusual and strange, will feel their logic and organic life.[6] That they are arrived at by the movement of melodically independent voices is obvious. The truth is that the melodic impulse is primary, and always preponderates over the harmonic; that the melodic, or linear impulse is the force out of which germinates not only harmony but also counterpoint and form. For the linear impulse is activated by *motion* and motion means life, creation, propagation and formation." - Toch (1948), p.10

Toch's point here, is that harmonies are not dictatorial pillars which define the pitch of the constituent notes, but snapshots of coincidences that emerge from the interplay of separate melodic lines as they develop in time. This stands in stark contrast to the way in which the 'harmonisation problem' is sometimes conceived and approached in algorithmic composition: "We apply the following criteria: we avoid parallel fifths, we avoid hidden unison, we forbid progression from diminished 5th to perfect 5th; we forbid crossing voices ... " Toch himself might have enjoyed the authors' own description of the output of the system, as it reinforces one of his central propositions. "The harmonisation produced by the GA has neither clear plan or intention" (Phon-Amnuaisuk and Wiggins (1999) p.5).

---

[6] Toch invites the reader to play each line, separately at first and then with the soprano, before playing the full harmonies, listening to each separately to appreciate their movements.

This is an extreme, although not atypical application of music theory to the design of algorithmic systems. Considered in the light of Cook's comments on the nature of musicological listening, and Toch's comments on the nature of musical listening some potential problems with this approach come to light. Music theory, working with a static representation of music, forms abstractions and generalisations. As Cook suggests, some of the key functional structures may be aurally imperceptable. It seems far from inevitable that in using theoretic principles as guiding principles to design systems capable of creating and playing music that we will recreate the temporal phenomenon from which the theory was derived. And without this, it may be hard to produce a sense of plan or intention that underlies the structures from which theory generalises. This could be one reason why we frequently see comments such as: "while conforming to classical triadic harmony, the music seems lifeless" (Cope (1999) p.21) or "The music often wanders with unbalanced and uncharacteristic phrase length. No musical logic is present beyond the chord-to-chord syntax" (ibid, p.22)

This is not to say that music theories are wrong, or worthless. The problem comes only in interdisciplinary settings where words have different meanings and theories do different jobs. Music theory is not aimed at providing a model of the phenomenon of music in the same was that biology does of living organisms. It offers ways of understanding music, ways of imagining music, and can undeniably alter the way that we appreciate it. But it does not aim to explicate the key phenomenon of relevance to us as listeners, or the things that composers do. It seems sensible to bear this in mind when applying it to the design of algorithmic composition systems aimed at generating material for people to listen to.

This section leaves behind scholarly academic discussions, and focuses more closely on the practical application of algorithmic processes in composition and live performance. The touch of ivory tower syndrome that dogged what we might call 'institutional algorithmic composition' for some time, whereby the standard method of evaluation was the we-think-it-sounded-quite-nice comment in a paper's conclusion is drawing to an end. Within academic circles, there are increasing numbers of conferences and dedicated workshops springing up which accompany the standard paper/ poster presentation with demos and concerts. ICMC, Ars Electronica and Generative Art have been joined by the Iteration series in Melbourne, the EvoMusArt workshop at EuroGP as well as countless other local groups. More dramatically, the increased accessibility of music programming languages mean that the days of computer music being the reserve of specialist institutes is long over. Music programming languages are not only more accessible and widely available than ever before, but open source communities such as PD and SuperCollider mean that young electronic music enthusiasts can work alongside pioneers such as Miller Puckette.

3.3 Performance Issues

In addition to those persuing electro-acoustic interactive performance, there is now a significant and rapidly expanding community of ardent laptopists, dedicated to performing live with little but their shiny machines and perhaps the odd cross-over cable uniting the smaller-ego'd in a mini laptop orchestra. This new breed of programmer-composer-performer typifies the dissolved hierarchies of contemporary culture and answers the call of many academic researchers for increased collaboration between computer scientists and composers by combining them in the same skin. The first two sections of this chapter focused on academic research in algorithmic composition, where performance exists as some kind of evaluation method. This section focuses on the practicalities of music making where algorithmic processes exist as a compositional or performance tool.

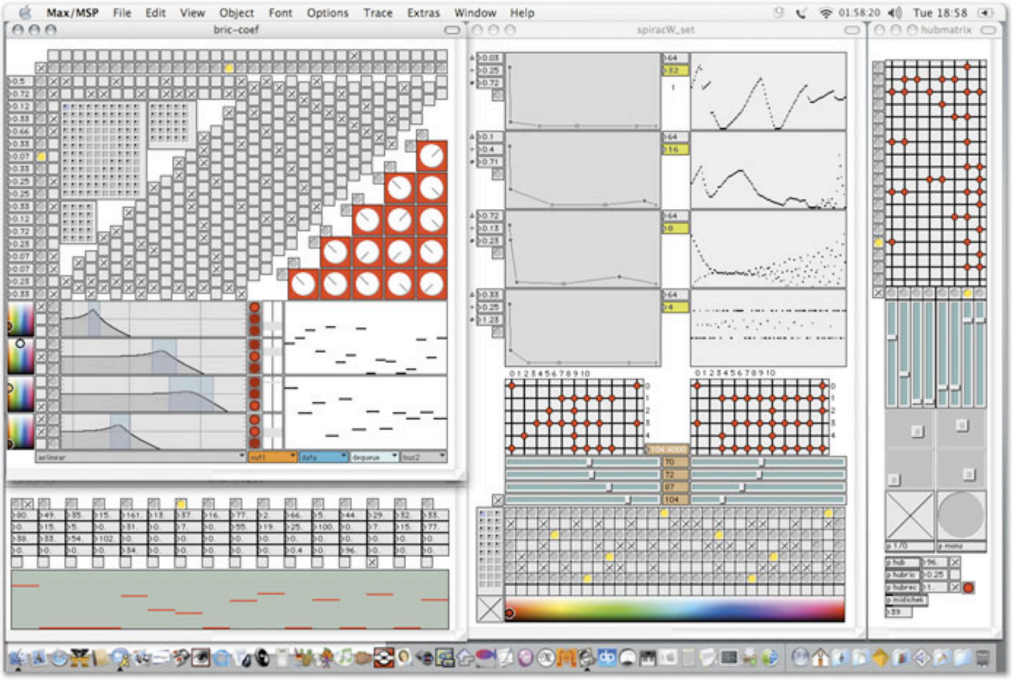


Figure 3.6: Screen shot of an Autechre Max patch.

3.3.1 Aims: Automation – Augmentation – Autonomy

Beside basic curiosity and nerdy experimentalism, there are several distinct reasons why composers in a number of fields might adopt algorithmic methods. Most pragmatically, they are very efficient, both in terms of time and storage space. Conceptually, and arguably aesthetically, the use of live processes brings back a sense of significance-to-the-now and uncertainty to electronic performance, a loss which was most famously bemoaned by Bartok with the invention of mechanical recording techniques. Creatively, algorithmic processes open up all sorts of possibilities which would be either impossible or incredibly laborious to create by hand. On the one hand of course generative processes are central to the various forms of interactive performance discussed in Chapter two in which the computer must respond to an instrumentalist or some other time-specific environmental cue. On the other hand, even in off-line composition the minutiae of the sonic/rhythmic/textural complexities afforded by the machine are invariably incredibly laborious, if not impossible, to specify by hand. Algorithms offer a crucial method of automation which is required to meet certain compositional aims. Increasingly generative processes are developed to augment these aims, pushing the user further along their chosen path, and even encouraged to wander off in new directions, suggesting new ideas.

Pragmatically, algorithmic processes can be incredibly efficient in terms of both composition time and (storage) space. This underlies the attraction of generative music for internet and games applications: indefinate amounts of never-repeating material can be specified in a definite and small number of bytes. In purely practical terms for installations which are required to run for long periods of time, algorithmic processes come into their own with respect to a composition time to runtime payoff.

In a paper discussing generative music in laptop performance, Nick Collins describes the use of live generative processes as a “tenuous hidden conceptual thrill” Collins (2003b). Tenuous in laptop performance, because, as he is the first to point out, what the audience sees is a glowing apple and a dimly lit look of concentration. It is often hard to differentiate the activites of the inventive algo-mentalist juggling hot-off-the-press SuperCollider procedures from the guy pressing 'go' on his pre-programmed Ableton Live set.

But for many the never-to-be-heard-again aspect of generative music represents a turning point in electronic music. Brian Eno has perhaps voiced this enthusiasm most loudly, demonstrating the joy of incommensurate loops in his *Music for airports*, but also raving about his experiences with other people's generative software. Describing his first encounter with Sseyo's KOAN he rejoiced:

> "I'm so thrilled by it that it is very difficult for me to listen to records anymore. Putting on a record and knowing I'm going to hear the same thing I did last time has actually become a little bit irksome. It feels quite Victorian to do that (laughter). I think this has really moved up into a new phase of music."
> - Eno (1996))

Whilst Eno may be fairly unique in sitting and listening to generative music at home, live generative music arguably brings with it a sense of uncertainty, the loss of which was feared in the age of mechanical reproduction, and again as MIDI sequencing became de rigueur in the pop music of the 1980s. Scored music is sometimes referred to with disdain by extreme proponents of the generative approach, but of course the nuances of expression and interpretation are significant enough for us to need both Pablo Casal's and Paul Tortelier's recordings of the Bach 'cello suites. And it is the freshness and uncertainties of live performance which draw us to go and see our favourite bands again and again despite having listened to their albums upteen times at home. In 1990, critic Jon Pareles disdainfully commented on the burgeoning use of sequenced MIDI material as a substitute for human players in pop music :

> "If I wanted flawlessness, I'd stay at home with the album. The spontaneity, uncertainty and ensemble coordination that automation eliminates are exactly what I go to concerts to see; the risk brings the suspense, and the sense of triumph to live pop perhaps the best we can hope for is that someone will come up with a way to program some rough edges too." - Pareles (1990)

As anyone working with live algorithmic processes will attest, uncertainty, suspense and spontaneity are algorithmic composition's middle names. The use of live algorithms certainly provides the possibility of some rough edges (though not perhaps of the sort Pareles was after) but arguably offers one way of regaining "A music that says this time is special, now is privileged " (Collins (2003b), p.71).

In live performance, but also in offline studio situations, algorithms are a welcome means of removing the laboriousness of dense or intricate composition. As Xenakis' work is often used to illustrate, computational procedures take the tedium out of calculating the millions of sonic events which the super-human capabilities of the computer can deliver. Both on and off-line, even fairly basic randomising techniques are warmly welcomed by any producers experimenting with slightly broken beats. The tedium of sequencing break-beats is removed by simple plug-ins such as Marcus Clements' ColdCutter[7] which randomly splices and/or reverses samples, creating complex dense rhythmic textures. Even the break-beat cognoscenti adopt more powerful implementations of stochastic break-beat generators such as Nick Collins' BBCut SuperCollider libraries which includes the squarepusher emulator mentioned above. In this situation, algorithms act as a method of automation, sometimes merely for saving time. In live situations, automation is the only option for achieving "…desirable but physically preposterous [effects such as] machine-gun-fire buffer stutters or micro-polyphonic granular swarming or Gaussian distribution counterpoint" (Collins (2003b) p.74)

---

[7] `http://www.brightonart.co.uk/coldcutter.shtml`

Others offer similar devices which are capable of learning the rhythmic styles of individuals. Once such project is Ollie Glass' *Breakage* system which generalises from user-supplied drum patterns. We might describe these as augmentations rather than automations, doing a bit of stylistic development on behalf of the user. Such devices were prophetically yearned for back in the 1980s by Laurie Spiegel. Bemoaning the over-use of randomness in algorithmic composition she expressed a desire for precisely this kind of algorithm:

> "A good algorithm should be a composer's amanuensis, an invisible assistant who reaches for a knob or a note for you, knowing just what you want to do and when. It should be an extra pair of hands who know their job well, and are able to execute or elaborate for you because they embody knowledge of what music is and how musicians really work. Ultimately, an algorithmic music program should be able to learn each individual user's unique personality, procedures, habits, and preferences, and to use this knowledge to take initiative and make musical suggestions when asked, to add to the individual's power by "automating" what s/he would do anyway, to extend the individual's music much further in directions it already takes. "
> - Spiegel (1989)

There is undoubtedly a calling and a place for automation and augmentation, but random number generators still feature large in the composition tools and performance patches of many using algorithmic methods. Whilst performing musicians may not pursue the world of out-of-control emergence which obsess Alife artists, there is increasing interest in semi-autonomous processes which demand answers and push the individual in entirely new directions. In the manifesto for a current EPSRC focus group 'Live algorithms for Music' (LAM), the need for algorithms which are 'not merely automatic, mechanical and predictable, but comprehensibly interactive and capable of novelty' is expressed.

3.3.2 Adaptability vs Dependability: Inspiration vs Freedom

Regardless of framework, the thrill of the unpredictable brings with it the danger of dull or irritating as well as the genius. In live performance in particular, the trade-off between adaptability and dependability is vital: if we are to make the most of the potential of live generative processes, they should be able to show variation in response to the moment; conversely if we are going to invite them onto the stage, we need confidence that they are going to behave well and not embarrass us. This is as true of tailor-made patches as it is of ready made algorithmic composition tools. Similarly any particular system should ideally provide enough structure to inspire us and prove its worth, whilst at the same time leave sufficient flexibility to allow us the creative freedom to work with it toward our personal compositional/ performance ends.

If generative processes are being unleashed live, the possibility exists that they may wander off into some dull or ear-bashing corner of its universe. Processes must be somehow constrained. But if constraints are so tight as to narrow the field to the point where differences are imperceptible or irrelevant musically, why bother writing a procedure? Why not just re-write the entire set and press play? When improvising with the digital other, there is only so long you can cover a system stuck in a stubborn silence, and there are only so many moves you can make in response to a sustained high frequency noise convulsion.

In situations when there is someone at the helm of the laptop, there are several tried-and-tested practical solutions to this problem. Assume that we adopt the 'generate and test approach', used by Hiller and Issacson in the composition of the Illiac suite, and employed in some form by many algorithmic users. The simplest solution, which works well in some situations, is to simply preview the current output - much like a DJ beat-matching before crossfading. Tasks such as tempo-matching can take place out of the audience's ear shot: algorithmic vagaries can be screened before being unleashed on the audience. Search mechanisms such as the Interactive GA technique described above could be applied to actively search large spaces for suitable candidates, a technique Palle Dahlstedt deploys in performances with *Mutasynth* (Dahlstedt (2001)).

Presets could also be made on the fly during the performances (with maybe a couple of last-resort settings made in advance in case of emergency). Max/MSP for example includes the *pattr* object and its family, which allow snapshots of parameter settings to be captured and saved. This makes it easy to capture any particularly classy behaviours which you might wish to return too. Collins suggests designing code which allows operation in 'autopilot' mode. Here, algorithmic processes are given free reign, but the procedures and interface are designed so that if a process wanders off into undesirable territory, the user can step in and take manual control.

Others not wishing to have to perform any screening or nannying adopt a more conservative approach, combining some generative mechanisms as the icing on the cake of predetermined material. For example John Eacott, working under the pseudonym *jnrtv*, performs dance sets where the rhythmic patterns and high-level structure is fixed, and surface effects are applied algorithmically to provide what he calls 'fluidity' (Eacott (2000)).

All these solutions are fine if generative process are used in solo or group laptop extemporisations, but in situations with a live instrumentalist, or when the the laptop is unmanned, we may need the algorithm itself to be a little more reliable. It needs to offer unlimited variation, but to operate within an acceptable arena: it needs to be able to do its own thing but respond to our calls.

When used live or for off-line composition, an algorithm should ideally offer a balance of structure yet freedom both in terms of control and implementation. A random number generator gives complete freedom of application, and the degrees of freedom allowed can be controlled by the composer, but it does not provide much structure. At the other extreme, many 'ready made' generative music applications leave little room for creative application on the behalf of the user. Sseyo's KOAN, which so pleased Eno was one of the most widely marketed of early systems aimed at allowing broad application, but other users don't find such joy:

> "in the case of a system like KOAN, although the documentation that accompanies it encourages you to interact with the parameters to create your own unique settings and hence, create a new piece of music, what is mostly achieved (in my experience) could be described as a remix of pieces pre-programmed by the development team." - Eacott (2000), p. 5

If mainstream commercial software is seen to be limited by the representational assumptions of the design team, aesthetic predilections of the programmer of ready-made algorithmic systems often shackle the user just as strongly. This is a common problem across much creative software, but is of particular issue for generative tools. Systems like KOAN include not only algorithmic specifications, but determine how these are mapped onto sound. As Eacott says, we are offered parameters to control certain aspects, and may be able to select between different MIDI instruments etc. but essentially we are constrained within a certain field delineated by the designers. Within generative arts more broadly, this is seen as a problem for applications offering themselves as 'creative tools', as although the process itself may offer many possibilities, decisions made by the programmer at the time of implementation introduce constraints. These issues are discussed in Chapter 7.

3.4 Discussion

In the last chapter, it was suggested that the Alife methodology was suitable for implementing software capable of supporting a rich interactive experience. Having looked at a range of approaches, potential pitfalls and pragmatic requirements, this section considers the potential for similar adaptive dynamical systems as methods of generative composition.

Section 3.2 provided a taste of the range of techniques that have been applied to the generation of musical material. Distinctions were made between scientific projects, which aim to model aspects of particular styles or the cognitive processes involved in composition, and those aimed at making practical contributions to composition methods. A further delineation was made between those aiming to develop systems with some form of representation of musical know-how derived from existing styles of music, and those aiming to explore new musical ideas employing what is often referred to as 'extra-musical' systems.

Section 3.2.2 aimed to illustrate some of the problems in the way that algorithmic composition is both framed and implemented: imbuing a system with representations of musicological thought may enable you to maximise your fitness function and help you achieve 'correct harmonisations' but it does not necessarily guarantee a decent musical output. The problems endemic in some areas of research are perhaps summed up by presentations of systems which are set the task of 'harmonising a Bach Chorale'. From a composer's perspective, this itself is a misnomer as it suggests a vertical rather than a linear approach that would have been entirely foreign to Bach's own aesthetic.

If we are not necessarily tied to the emulation of particular existing styles, we have more freedom in algorithm design; at the same time, without the trappings of familiar harmonies or forms we may have to work harder to engage the listener. The fact that the most popular mathematical models to date have been models of growth, pattern formation and population evolution suggest that one alternative is to focus on achieving a strong linear impulse. Several of these seem to be useful in sound design situations, emulating the temporal evolution of complex sound progressions. These successes suggest that at least in some circumstances, formal properties of these dynamical systems can be used to create comparably dynamic sonic effects.

In contrast, applications of Alife and dynamical models at higher levels of musical organisation do not seem to have been so successful. Whilst systems such as CAs, L-systems and chaos models are capable of generating some interesting fragments, they have not been so successfully applied to the generation of any significant musical works. There are at least three possibilities which can be investigated toward developing upon existing approaches.

Part of the reason perhaps is that they do not present immediately obvious hierarchical structures, generally being mapped onto a stream of monophonic notes. Agent-based ecosystem models as used by Dahlstedt in Living Melodies, do perhaps offer richer structural possibilities, but their system was sonified using mappings that stuck to low level pitch determinations. One obvious move then, is to explore other approaches to mapping.

Another alternative is to examine some other models. In Alife music in particular, 'Alife' has become synonymous with GAs, CAs and multi-agent models. There are many, many other possibilities: the existing artillery employed within Alife research is itself extensive, but also provides illustrations from which bespoke models can be developed. Using off-the shelf models is a starting point, but as suggested above in the application of AI learning and search algorithms, real headway may only be made once these techniques are tailored to meet our specific purposes. Of the existing techniques used in Alife research, there are many which offer an attractive middle-ground between the single dimensional chaos model and the complex co-evolving ecosystems examined in the last chapter. In particular, the sub-field of autonomous systems research develops many models for investigating adaptive control of single agents.

A third alternative is to combine different models with different characteristics. As noted above, new offerings in this area are, perhaps inevitably, given in the form of single models. Generative artists talk about composing processes, but there has been little exploration of the potential for composing *sets* of processes. These possibilities will be explored in Chapter 6.

Selection and development of these models can most sensibly proceed by taking into consideration some of the desirable properties discussed in Section 3.3. Most crucially, it was suggested that if we are going to risk relinquishing control to a live algorithm on stage, it needs to be adaptable, but dependable. Further that if it is going to have more than a one-off idiosyncratic appeal, it is desirable that aspects of its global behaviour can be tuned, giving both a vast range of possibilities, and some degree of control over where in that space it travels. CAs and the chaos models that have been used begin to address these balances. Both are sensitive to initial conditions, meaning that there is a pragmatically indefinite variation in their output. On the other hand, as deterministic systems, this initial seed can be saved, and the same pattern recreated. In the case of CAs it is also possible to define particular classes of behaviour - i.e. chaotic, ordered or complex - according to defined rules or experimental experience. Chapter 6 presents explorations of models with at once greater adaptive potential, and richer possibilities of control.

Adaptive dynamical systems open a potentially interesting pathway for computer music composers. Carving a line between the relentless unfolding of uninterruptible mathematical equations and the offline search for new musical forms illustrated by GAs. Collins suggested that algorithms could be released live on stage given the implementation of GUI controls which allow the human to take over. The use of adaptive systems makes for a more collaborative form of man-machine performance in allowing the possibility to influence or coerce behaviours, rather than grab the reigns back.

Discussions in the last two chapters suggested that the adaptive, dynamical framework adopted by Alife research offers the potential to meet the latent wants of some in the interactive and generative music communities. The initial aesthetic motivation for this whole project stemmed from observing the artificial agency apparent in visualisations of Alife style systems and wanting to bring it onto the stage. But it is perhaps easy to get wooed by this artificial agency and assume that this can be readily transported into the auditory domain. Before exploring the musical potential of adaptive systems, the next chapter takes a step back and considers the source of this apparent agency: is it something inherent in the system architecture ? or an illusion created by their physical presence or the visual trappings which allude to biological creatures?

Chapter 4

Mimesis, Alife Art and Music

The fascination with creating 'cybernatures' evident in Alife Art is not just an artistic spin on some hard science. Simon Penny (1995) has suggested that Alife research itself shares an underlying motivation with many past art practices. He proffers that Alife research can be seen as a modern-day technologically enabled expression of deep and ancient drives to imitate nature and animal qualities. These same impulses, he suggests, drove the Greek expression of human form in classical sculpture and the Georgian fascination with automata, epitomised in curiosities such as the mechanical duck made by de Vaucanson (1742) or Kaufmann's mechanical trumpeter (1810) (Häfner and Krätz (1978)). In this respect, Penny compares core concerns of contemporary Alife researchers with those of artists such as Cézanne, who at the turn of the century proffered: 'Art is a harmony parallel to nature'.

The current enthusiasm for Alife models in the generative and interactive arts can be seen as an incarnation of the same compulsions. The last fifteen years has seen an abundance of Alife-inspired art, not only on the web and at specialist events, but at major art institutions around the world: Karl Sims' Genetic Images was shown at the Pompidou centre, Paris (1993) and Sommerer and Mignonneau have recently had shows at both the Victoria and Albert Museum London (Touch me, 2005), and the Van Gogh Museum, Amsterdam (Fierce Friends: Artists & Animals in the Industrial Age, 2005). As noted in the last chapter, there has been some musical exploration of Alife techniques, but these have seen nothing of this kind of success in the public domain, either on stage, on record, or on air. This might be because the music world is not as easily ingratiated by progressive art forms. But that seems unlikely. It might be because musicians aren't as good at self promotion, or simply that the ideas have taken longer to enter into the music scene. But there is little evidence for this. Could it be due in part to the fact that the predominantly *visual* nature of Alife research is more easily transformed into visual art ? Are these sorts of processes somehow less amenable to representation in the sonic medium ? Does the sense of artificial agency not carry in the sound world ? Or have we just not yet found suitable models and mappings ?

This chapter considers the close relations between the visualisation techniques used in Alife research and the forms presented as Alife art, and questions how strongly the appearance of agency in these systems relies on their visual presentation. Are the abstract critters we see wandering about perceived as intentional just because of clever representational tricks? Or is there potential to use such systems to invite a comparable attribution of intentionality in the sound world? As a first step, all aesthetic considerations are dropped and we take a step back and question whether those formal systems that have conceptual and aesthetic appeal visually can create a similar affect in audio.

This is an important question to raise, not only for the current project, but for the use of extra-musical algorithms in general. As discussed in Chapter 3, many practitioners adopt formal models on the basis on some perceived analogy between the structural dynamics, or behaviour of the model and some musical morphology or phenomenon. As Truax noted in his scathing comment on the use of non-linear systems, a programme note explaining the rationale can capture the imagination of the audience for a little while, but the conceptual interest may be rather short-lived if the musical effect is empty. The implication is that whilst it might be a nice idea, in practice the particular algorithm may not be any more effective musically than a random number generator. Well dressed noise is a powerful tool, but in promoting a particular class of models as being useful for various musical activities, it seems important to check that they can do more than a noise function. The first step in this is to check that it is at least possible for the formal properties of an exemplary model to be perceived from a sonification of its numerical outputs.

Section 4.3 therefore presents the results of an experimental psychology study which was run to investigate whether people could perceive the states of a one dimensional (1D) binary CA from an audio representation. The results of this study lead to a deeper consideration of mapping.

4.1 Seeing Artificial Life

The Alife roots of many generative and interactive artworks are vividly apparent. There is a veritable dynasty of ecosystem-based visual installations in which abstract virtual creatures scoot about a virtual space, feeding, mating, competing and morphogenically diversifying whose ancestral origin in research such as Tom Ray's Tierra (Ray (1991)) and John Holland's ECHO (Forrest and Jones (1994)) system is unmistakable. As mentioned in Chapter 2, Richard Dawkins' BioMorph (Dawkins (1986)), which breeds insect-like forms using evolutionary computation driven by aesthetic selection, was rapidly and very directly applied to on-line and interactive art by William Latham and Karl Simms in the form of *Mutator*, (Todd and Latham (1991)) and *Genetic Images* (Sims (1991)). Similarly the graphical demonstration of the power of a handful of simple rules to coordinate flocking behaviour by Craig Reynolds (1987) has spawned an entire genre of Swarm Art within the Processing community[1]. In each of these cases not only have conceptual and formal models have been directly appropriated but also the method of visualisation.

The inherently visual basis of Alife as a research programme may be one reason for the predominance of visual over sonic application in the art world. Conway's Game of Life (Gardner (1970)) was of fascination partly because it demonstrated the emergence of complex behaviour from simple rules in silico. But if we accept the verity of claims such as that since 1970, more computer time worldwide has been devoted to the Game of Life than any other single activity (Chennamangalam (2003)) one might be tempted to attribute at least some of its appeal to its graphical interface. Examination of streams of zeros and ones would ultimately reveal the same information, but the fact that you can literally *see* the little critters flashing and blinking and gliding across the screen, undeniably increases its appeal and accessibility, and even perhaps its power of persuassion. Graphic visualisations can comprehension of complex models, but in some cases also shorten the phenomenological distance between the behaviours of these formal systems and the real-world phenomenon which they model. The same could be said for the majority of Alife simulations. Graphs of global fitness measures or line plots of ecosystem diversity provide us with the information necessary to judge the success of a simulation, but it is *seeing* the agent successfully avoid the falling object or freakish forms emerging

---

[1] `http://www.processing.org`

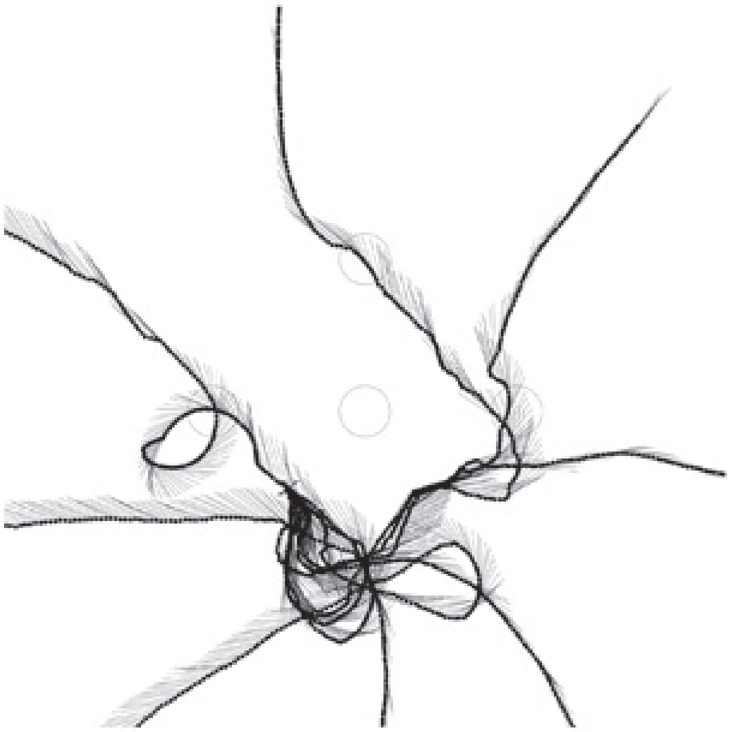


Figure 4.1: Visualisation of simulated foraging behaviour. Dale (2000).

in the silicon graphic soup which get us excited. Even before they have got into the hands of artists then, the behaviour of a great many Alife simulations, or more accurately, the visual representation of the behaviour of a great many Alife simulations, pull many of the same strings in us that artists aim to tug.

The effect can be seen in static 2D plots as well as animated graphics. In 2000 Paul Brown was artist in residence at the Centre for Computational Neuroscience and Robotics at the University of Sussex, UK. At a research seminar given by then DPhil student Kyran Dale, he saw a plot of the paths taken by an evolved animat navigating toward a food source from eight different locations in a 2D plain. The animat was controlled by a continuous time recurrent neural network (CTRNN) which had been evolved for this navigation task. The plot is shown in Figure 4.1 where the five faint circles represent land marks, and the cross (through which all paths pass) symbolises the food source.

Prior to his visit to the CCNR, Paul had been a fine art tutor at various tertiary establishments for twenty years. His response to the visualisation of the CTRNN's behaviour was that any student producing a drawing similar to that shown in Figure 4.1 ″would have been assessed by their mentors as 'showing talent' ″ (Brown (2005) p.5). It is precisely the appearance of agency, the mark of motivation or goalseeking behaviour which is evident in this drawing which appeals to the Alife artist, and indeed could be said to be one of the aesthetics of Alife art – it is also of course the intention of the Alife researcher to create systems which exhibit these life-like behaviours. It was experimenting with such systems that inspired me to try and listen to them, and indeed Paul himself cites this image as convincing him that it would be possible to create a drawing robot, a three year AHRC funded research programme which he subsequently embarked upon.

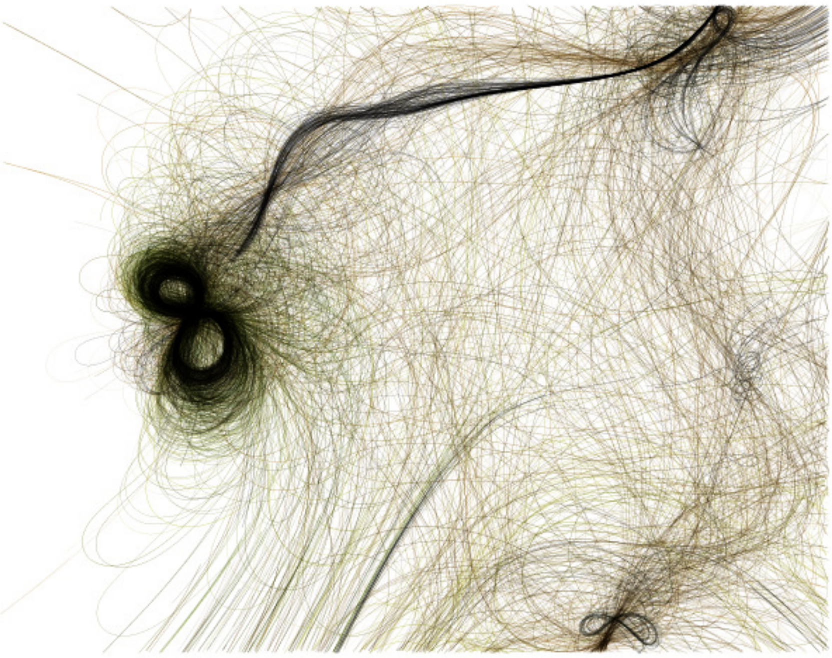


Figure 4.2: One of a set of 28 prints made with the *Tissue* software. Casey Reas 2002.

A similar, although much simpler, model has been used by artist Casey Reas. *Tissue* [2] is a body of work, based on the visualisation of the paths of thousands of Braitenberg Vehicles. Braitenberg's original vehicles (Braitenberg (1986)) were a thought experiment, but can be conceived, and modelled, as simple robotic agents with light sensors, wheels and variable speed motors. According to which way the light sensors are wired up to the motors, the vehicles can be made to seek or avoid obstacles or light sources. Reas uses thousands of similar agents, programming each one to leave a trace that shows the paths it has taken as it follows, or avoids, other agents or obstacles. The work has been exhibited both as an installation and as a set of twenty-eight prints made by Reas[3]. As an installation piece, users can interact with a 2D graphical environment through which the agents navigate. By positioning stimuli around the environment they can indirectly affect the behaviour, and thus the traces left by the agents. Reas creates beautiful organic-looking images using fine pencil-like lines and carefully selected colour schemes, however the main effect is not dissimilar to that experienced when playing with a khepsim simulator. Certainly Reas' prints and Dale's visualisation bear more than a family resemblance. This is in no way meant to belittle the work of Reas, or any other visual Alife artist, but reminds us of the close relations between Alife research and visual Alife art.

Whilst the Alife roots of these art works are vividly apparent, what is less obvious is how easily these artificial agencies can survive outside of the visual worlds in which they are presented. In visual Alife art, these virtual critters are often presented in a frame of familiar environmental structures, a ground, a sky, a familiar spatiality. These provide a context which encourages our zoomorphic attributions. As Whitelaw puts it:

---

[2] `http://www.reas.com/iperimage.php?section=works&work=tissue_p&id=0`

[3] `http://www.reas.com/texts/tissue.html`

> “Our cultural familiarity with screen-based representation and the ubiquity of this form (essentially a view of a landscape) leave us well equipped to take up these cues, however scarce or marginal, and construct a stable analogy. Once this artificial landscape is established, we read the represented events, however crude, according to the same analogy. When two computer-graphic blobs meet, and a third smaller blob appears, we understand that a birth has occurred. When two forms meet and one vanishes, we see a predator and its prey. - Whitelaw (2004), p.79

But these cues are not always present. And in nearly all cases there are definite behavioural resemblances in both their movement and their response to encounters with other objects and agents.

4.2 Hearing Alife

Many people have of course explored the potential for Alife-type models in purely musical applications as mentioned in the first two chapters (Impett (2000), Bilotta and Pantano (2002), Miranda (2000a), Blackwell (2003) etc.). However, very few of these have received public attention of the level enjoyed by their compatriates in the visual domain. It is of note that some of the highest calibre works have come from composers who have chosen to transcribe the output of the system for human performers. Both *Entre o Absurdo e o Mist´rio*, and the second movement of *Wee Batucada Scotica* by Eduardo Miranda were composed using material generated by a CA and performed by chamber orchestra and string quartet respectively (Miranda (2000b)). Similarly Rodney Waschka’s (2001) opera *Sappho’s breath* (see Section 3.1.2) was performed to large audiences by soprano Beth Griffith[4]. Is part of the difficulty in capturing public interest associated with the *digital* delivery of the music rather than the material itself ?

Writing on the aesthetics of computer music, Guy Garnett (2001) suggests that the two go hand-in hand. There are certain constraints on the compositional possibilities associated with human instrumentalists which are removed when the performer is a machine. Most obvious is the lack of physical constraints: a machine can play faster, more precisely, for longer etc., and is not constrained in pitch or amplitude of acoustic signal as is an acoustic instrument. But as Garnett notes the constraints on ‘performability’ associated with writing music for human instrumentalists impose not only *physical* restrictions, but *cognitive* limits on the musical material as well. A player must be able to get not only their hands (and maybe lungs) around compositional gestures and structures, but in order to perform music, they arguably need to be able to get their *mind* around it. Escaping the physical constraints of acoustic instruments is a major attraction for computer music composers, and arguably essential to the current project, but Garnett suggests that these restraints may well also serve to keep the material within a frame which potential listeners may be able to digest. Remove these limitations and the possibility arises for the composer to get so carried away with the formal elegance of a particular model that the results are incomprehensible to the audience:

> “The composer, without physical limitations of performance, can more easily convince himself or herself that they have created something real and comprehensible, whereas what they have may be an unhearable ideal. It is relatively easy to create algorithms that generate sounds whose qualities as music are inscrutable, beyond the cognitive or perceptive abilities of listeners. And

---

[4] Although it should also be noted that neither of these pieces were composed entirely using Alife methods, but generated fragments were recomposed by hand.

> with computer programs, it is not only possible but becomes a rather frequent occurrence." - Garnett (2001), p.26

It is easy to see how Alife music composers may suffer similar seductions. The key aesthetic element here is perhaps 'behaviour' (complex, adaptive, emergent, life-like etc.) rather than precision, but it is only too easy to be enchanted by the conceptual charm of a model of growth, evolution or self-organisation, and forget to question whether these processes which are mathematically - and visually - compelling have any psychological reality for listeners when the numerical outputs are mapped into sound.

Composers and researchers working in this area generally select models according to a perceived analogy between the structural dynamics, or behaviour of the model and some musical morphology or phenomenon. Tim Blackwell for example, uses swarm models in his various interactive performance systems as he suggests that a similar process of self organisation occurs within free improvisation (Blackwell (2004)). Bilotta and Pantano (2002) choose CAs to generate music, suggesting that the CA's "capacity to mimic both evolution and growth in biological life seem to have some basic peculiarities in common with natural human languages (and thus with music)." (Bilotta and Pantano (2002), p.1)

The self-organising 'swarm' is impellingly present in visual depictions of the algorithm, and the emergence of complexities of CAs are readily observed in their graphical representations. But can we necessarily manage to create equivalent phenomological realites in the sound world ?

4.3 Testing the Auditory Perception of CA States

One could argue that it doesn't matter. If it works, it works. If the outcomes are effective musically then why does it matter how closely success is tied to formal aspects of the algorithm? Well on the one hand if they *aren't* effective then it might be useful to know whether the delivery, implementation, or central concept was flawed. If this is carried out in a research setting, there is perhaps some onus on the author to bolster their motivational assumptions. More pragmatically, one of the central tenants of this thesis is that the use of Alife and adaptive models offer an exciting new compendium of toys for the digital composer. The development of these tools would benefit from some basic understanding of their potentials and effects. Assessment of their musical value is perhaps best left to audience reaction, but if we want to explore these types of models as compositional and performance tools, it seems sensible to stop and check that we can hear them.

This section describes a study which was run to ascertain whether people could identify distinct classes of CA rule sets from an auditory representation. This is not to suggest that it is necessary, or even perhaps desirable, for the audience to be able to fully comprehend the state dynamics of a particular model. However, if it is not possible sonify such systems such that their conceptually attractive properties can be appreciated, we might as well just write our ideas down in Truax's program note and spend our time finding clever ways to use noise generators.

4.3.1 Auditory Perception and Auditory Display

Whilst there may not have been any work explicitly addressing the issue of how formal structures are perceived in sound within the algorithmic composition literature, there is extensive work of relevance being done within the emerging field of auditory display. An established International Community of Auditory Display[5] hosts discussion of design approaches and applications for auditory display in a range of disciplines. Much of this

---

[5] `http://www.icad.org`

research is in applied settings such as assistive technologies for the visually impaired (Lunney and Morrison (1990), Kennel (1996)), mobile computing (Brewster (2002)) and virtual reality systems. Although there has been little work done in the area Alife directly, there is an increasing interest in the use of auditory display for scientific visualisation in general (e.g. Hayward (1994), Dombois (2001)) and in medical settings in particular (Fitch and Kramer (1994)) which is of relevance.

Existing research into auditory perception suggests that certain types of data may be particularly amenable to aural comprehension. Speech-based evidence of selective-attention (e.g. Handel (1989)) suggests that the auditory system may be capable of monitoring data structures embedded in other more static signals which would be too noisy to apprehend visually. A nice anecdotal example of this comes from the Voyager 2 space mission. As the craft approached Saturn it started experiencing severe problems, the cause of which could not be diagnosed from on-board graphical meters which depicted pure noise. The data was sent back to earth and played back through a synthesiser, revealing a machine gun effect at the critical period, which led to the realisation that the craft was being bombarded with electromagnetically charged micrometeoroids (Kramer (1994b)).

Other basic properties of acoustic perception suggest that sound may be a particularly good medium for presenting and understanding the sorts of complex dynamic behaviours of interest to musicians. For example it has been suggested that the ear is particularly good at resolving multidimensional data in general (Bly (1982), Gaver (1989)) and logarithmic or time-varying data in particular (Bly (1982)). The superior temporal resolution of the acoustic system (e.g. Poppel (1994)), suggests that fast changing or transient events that may be blurred or entirely missed visually can easily be heard. Sensitivity to temporal characteristics also enables discrimination between periodic and aperiodic events. We are able to detect salient patterns, even when subject to radical transformation. Again, this is supported by anecdotal evidence from the lab in which the quantum whistle[6] was discovered. The oscillations predicted by quantum theory could not be detected using a visual oscilloscope, however, transformation of the data into an acoustic signal created a faint whistle, providing the first experimental support for theoretical predictions (Pereverez et al. (1997)).

Of key interest in the current context is the ease with which complex dynamics can be appreciated in an audio signal. Consider for example that doctors' principle tool for analysing ailments in the human respiratory, digestive or circulatory system is the stethoscope: medical students learn to *listen* to irregularities in blood pumping through veins, oxygen osmosing through alveoli, or gases bubbling in the intestines. Experimental results show that in a simulated operation, medical students provided with eight dynamic variables describing the health of a patient presented in audio, out-performed those given visual, and even audio-visual displays (Fitch and Kramer (1994)). Results from other medical and engineering investigations into auditory display support the idea that cycles, rhythms, patterns and short events are particularly amenable to acoustic analysis, McCabe and Rangwalla (1994). Whilst there has been no direct investigation into our ability to perceive the state dynamics of complex systems, all these findings suggest that our hearing system is well attuned to be able to do so.

Research in the field of auditory display also suggests that data describing natural processes such as seismic readings can be more easily appreciated than other data such as stock market figures, due to a shared physics:

---

[6] A quantum whistle is a peculiar characteristic of supercold condensed fluids which vibrate when you try to push them through a tiny hole. This has potential for developing incredibly sensitive rotation detectors which could be used for example to measure rotational signals from earthquakes or very precise gyroscopes for submarines.

> "A seismic recording will sound like a recording of natural environmental sounds, because sounds transmitted through air (acoustic waves) have a similar physics to seismic vibrations transmitted through the earth (elastic waves). The direct, physically consistent, playback can take advantage of human experience with natural sounds" - Hayward (1994), p.93

This might be seen as a benefit for those working with musical applications of biologically-inspired models, on the premise that many of the forms and dynamics modelled share certain characteristics with phenomenon in the natural world and so from an evolutionary perspective may be more comprehensible than other formal processes. At a basic level of perceptual comprehension, it seems that there is no reason why we shouldn't be able to hear types of processes typical of Alife-like models. Infact it seems like our hearing system might be *better* than our visual systems at taking them in.

4.3.2 The Effect of Musical Experience on Perceptual Accuity

Many of these research findings tarry with our experiences of listening to music (which itself could be conceived as a complex dynamic system): we can pick out a rock bassline plastered in heavily distorted guitar riffs; we can differentiate between, and simultaneously attend to, the vocal, keyboard and guitar parts; and we can recognise familiar tunes even when key or tempo are dramatically altered. Musicians can do even better than this. Music students learn to not only monitor and separate individual musical lines, but even to dictate four or five part harmonies, transcribing the individual pitches and rhythms of parts even for instruments of similar timbres. They can recognise not just familiar tunes, but pick out novel motivic fragments even in complex orchestrations and when subject to radical transformations in rhythm or pitch. These feats are impressive illustrations of our ability to hone our perceptual accuity, but also represent quite considerable individual differences in listening ability which could be of relevance for artists (or scientists) wishing to represent formal systems in sound. In particular, it suggests that it is highly possible for a composer to appreciate the abstract processes he is sonifying as he sits for hours on end listening to incremental changes during the development of his system, but that by the time it gets a public hearing, the layers of complexity render the central propositions utterly irredeemable to the first-time listener.

The significant effect of musical training on acoustic perception is illustrated by a range of studies. Physiological and psychological differences between musicians and non-musicians have been demonstrated (Petsche et al. (1988)), and differences in EEG dimensionality between classical and popular music listeners point to the psychophysiological nature of this difference (Birmbaumer et al. (1996)). Musical expertise has been shown to affect simple perceptual, as well as conceptual judgments of pitch. For example, in a controlled experiment, Neuhoff and Wayand (2002) tested participants of varying levels of musical experience and found that musicians reported significantly greater pitch changes than non-musicians for the same interval. In addition, errors in judgements of direction of frequency change were significantly greater for non-musicians (i.e. they said note *a* was higher in pitch than note *b* when it was in fact lower). These findings have obvious implications for the development and application of auditory displays, but may be useful considerations for algorithmic composers, especially those sonifying complex dynamic systems.

4.3.3 Design Rationale

The sorts of characteristics of relevance to musical Alife applications are things such as general trends in the population dynamics of a GA or ecology model, the dynamic organisation of a swarm system, whether the outputs of a neural network have settled to a stable state or are still evolving: general classes of behaviour for complex dynamic systems. Pilot work investigating the comprehension of a homeostatic network (described in Chapters 5 and 6) suggested that people could readily hear whether the continuous-time outputs of a multi-node network had settled into a stable converged or oscillatory state, or were oscillating wildly out of equilibrium. As noted above, it is well known that the auditory system is capable of monitoring multi-dimensional data, and we are adept at recognising periodic patterns, so this task is relatively easy. To create a perceptually more challenging task which would allow examination of auditory recognition of state dynamics, and also enable the investigation of differences according to musical experience, a 1D CA was chosen as the model to be sonified.

CAs are one of the most explored models in Alife music (Bilotta and Pantano (2002), Miranda (2000a), Brown et al. (2000), Burraston et al. (2004)). They are discrete models which are generally conceived (and visualised) as a regular grid of cells, which can each take on one of a finite number of states. The model is described by a set of update rules which operate in discrete time steps and determine the state of each cell at time $t + 1$ according to the state of its neighbourhood at time t. Rules and neighbourhoods are usually fixed. One of the simplest CA models, which is used here, is a 1D model where each cell takes on a binary value. The system is usually visualised by plotting the state of each successive iteration as horizontal lines, one below the other (Figure 4.3).

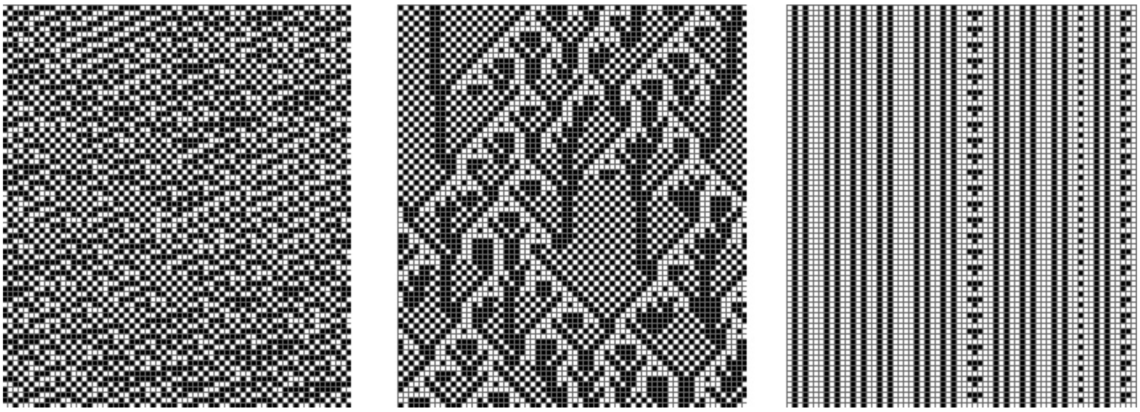


Figure 4.3: Examples of visual stimuli for chaotic (left) complex (middle) and ordered (right) rule sets.

Some rules produce static, ordered patterns of cells, such as those shown in Figure 4.3 (right). Others produce chaotic distributions of cell states, much like the noise on an untuned television set (Figure 4.3 (left)). Computationally the most interesting is the third class of rules which produce complex patterns (Figure 4.3 (centre)). Areas of high order suddenly give way to areas of chaos and then re-order. These patterns are easily observed in graphical depictions, which provides a control with which to compare auditory recognition. Because these are discrete time systems, and the recognition of rule class requires consideration of the current state in the context of its history, it is likely that the global state of a CA will be harder to hear than that of continuous time models. However the pattern detecting powers of the auditory system suggest that it should be possible to represent these patterns in sound such that these three classes can be differentiated.

4.3.4 Method

A categorisation task was designed in which participants had to classify the outputs of a 1D binary CA as one of three classes: complex, chaotic or ordered. This was done using graphic, audio and audio-visual displays and carried out by music and science students of comparable ages.

Participants and apparatus
Twenty music students from Northbrook College Music Technology course and Twenty non-music students from the Informatics department at the University of Sussex, UK were each paid £5 to take part in the study. All reported normal or corrected to normal vision and hearing. Participants were screened to ensure that music students all spent at least 10 hours per week engaged in active listening (playing an instrument with others, dj-ing or producing music) and had done so for at least three years. Non-music students were screened to ensure they did not have similar experience. It was assumed that they were all familiar with graphical displays.

The task required the classification of 1D binary CA into one of three qualitative states (ordered, chaotic or complex). These are equivalent to the four classes described by Wolfram (1982) where classes one and two are conflated. Rules from each class were taken from (Wuensche (1997) (K = 5)). Three blocks of twenty-one trials were presented, across which mode was manipulated, creating 63 trials in all. Visual stimuli were presented on a 15 inch LCD display. Auditory stimuli were presented via Sennheiser stereo headphones. The experiment was run on purpose-built software, using MIDI to trigger native instruments FM7 virtual synth, (preset bank 1 ALL, no 23 'native percussion').

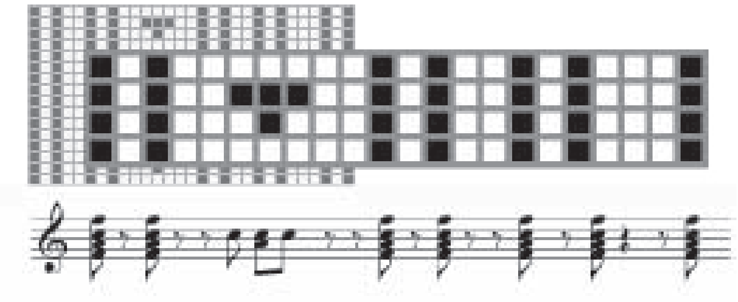


Figure 4.4: Rhythmic mapping: cell states are transformed to musical events: 1 = play, 0 = rest. Four lines are voiced simultaneously

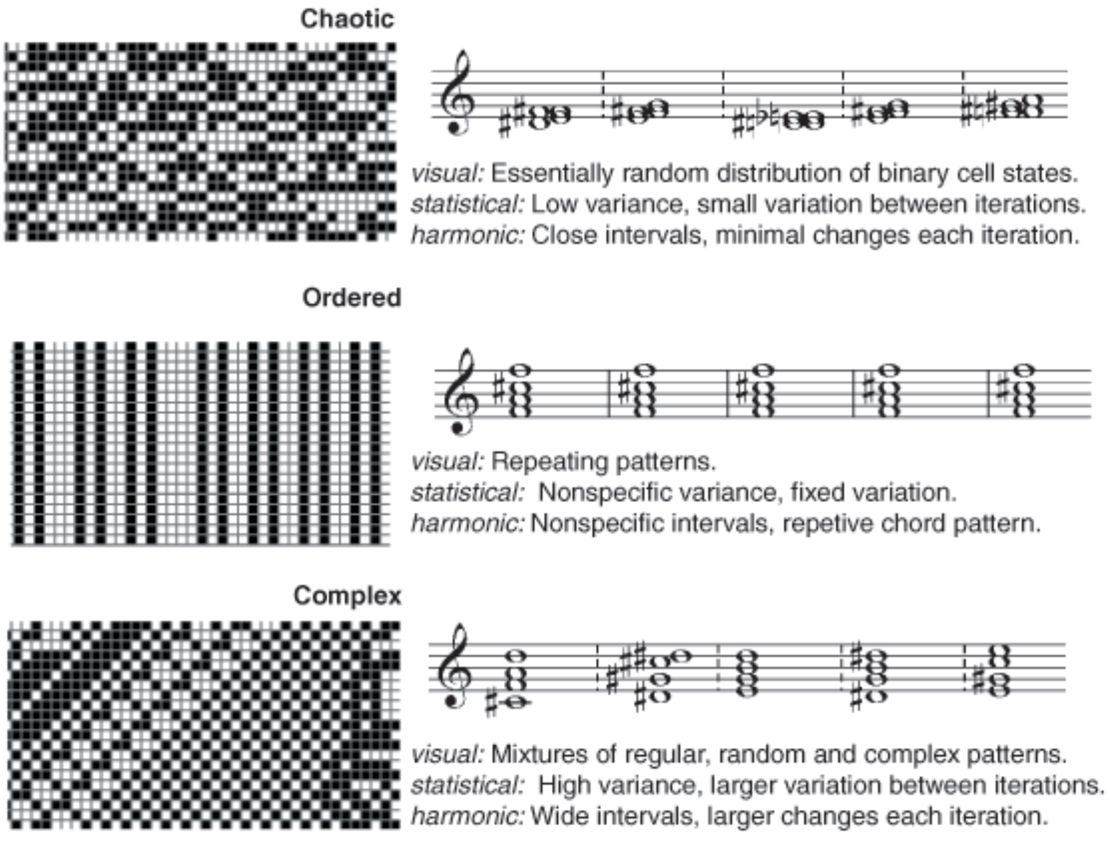
Stimuli
CA Rules were taken from those described by Andrew Wuensche (1997) which are categorised according to the entropy variance of the rule look-up tables. They were implemented on a grid 66 cells wide with wrap-around and initialised randomly with 20% set to one. Prior to presentation, each rule was run until it achieved its characteristic state.

Visual stimuli were presented on computer monitors placed 50cm from participants. CA states were represented graphically in a black and white grid with a grey background. Example stimuli from each class are presented in Figure 4.3. Initially four lines were presented, and then automatically updated at the same rate as the audio representation progressed.

Auditory stimuli were created using a sonification scheme which employs two sets of mappings. One transforms the familiar spatial patterns of the CA into temporal patterns, creating distinct types of rhythms for each class. The other converts statistical properties of the rule look up table into pitch values, creating harmonic progressions which vary characteristically for each class.

The rhythmic mapping, which is shown in Figure 4.4, transforms spatial patterns into temporal patterns by mapping cell state to note status: 1 = play, 0 = rest. The 1D array of cell states is read left to right, producing 66 timesteps per iteration of CA rules. In order to preserve the context or history available in the visual display, four lines were voiced simultaneously at different pitches. Note that although this mapping changes the temporal characteristics, producing a continuous sequential rhythmic development in contrast to the discrete synchronous graphical update, the spatio-temporal mappings preserves Gestalt properties that are thought to be key to pattern perception such a grouping by proximity.

The harmonic mapping determines the pitch of the each note according to the frequency distribution of the rule look-up table which is updated each iteration. At each time step, the number of times each possible rule is used is recorded. The mean of this frequency distribution is used to determine the pitch of the bass note. Any cells that are alive in the current array are voiced at this bass pitch. Live cells from the previous three iterations are voiced at successively higher pitches at intervals equal to the variance of the frequency distribution. Because the statistical distributions vary qualitatively with each rule type[7], this mapping produces chords, and chord sequences that differ characteristically: ordered rules produced fixed progressions that are repeated, chaotic rules produce close, dissonant chords that vary minimally and complex rules produce wider chords with more significant changes (see Figure 4.5).



visual: Essentially random distribution of binary cell states.
statistical: Low variance, small variation between iterations.
harmonic: Close intervals, minimal changes each iteration.

visual: Repeating patterns.
statistical: Nonspecific variance, fixed variation.
harmonic: Nonspecific intervals, repetive chord pattern.

visual: Mixtures of regular, random and complex patterns.
statistical: High variance, larger variation between iterations.
harmonic: Wide intervals, larger changes each iteration.

Figure 4.5: Harmonic mapping for the CA.

---

[7] This same pattern underlies the discrimination by entropy variance used by Weunsche. This measure was used here as for the current purposes it provided the same differentiation, but was less expensive computationally

Design

In order to familiarise participants with the task, each completed a practice phase before undertaking the main task. In the practice phase, participants were required to categorise CA states using an audio-visual display, and given feedback on their choice. They were instructed to view sufficient examples until they felt 'comfortable' with the task.

In the main task, all participants in both groups were subject to all three conditions across which presentation mode was manipulated. Each classifed the same three sets of twenty-one rules presented in three blocks according to presentation mode (audio, visual, audio-visual). Condition order was counter balanced across participants, and presentation order of rules was randomised. Classifications as well as response times were recorded.

Procedure

Participants were first given written instructions and explainations of the task and then presented with visual and audio examples of each of the three classes. In the practice phase the auditory and visual representations were displayed simultaneously (equivalent to the audio-visual condition). There were initially four lines of visual display which updated in time with the auditory display. For the first six examples the class type was displayed on the screen. Subsequently, participants practiced classification by clicking one of three labelled buttons, and received on-screen feedback as to the correct response.

In the test phase, participants no longer received feedback, and were instructed to attend each stimuli until they felt confident of their classification choice. Responses were made via one of three labelled buttons, and the next stimuli was presented 75ms after the 'next stimulus' button was clicked. They were encouraged to have a short break between conditions if necessary.

4.3.5 Results

Raw percentage accuracy scores were taken as the performance measure. These are summarised in Figure 4.6.

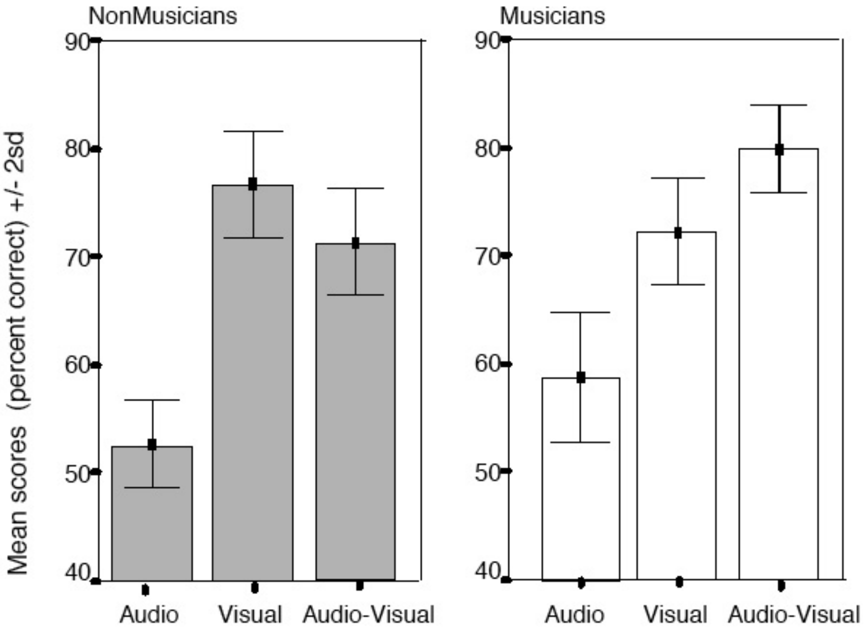


Figure 4.6: Mean scores and standard deviations for each group across all conditions

Presentation Mode Effects In every condition, for both groups, percentage accuracy was significantly above chance, suggesting that participants were able to make correct classifications based on the available information in all modalities. Although participants apparently found the patterns hardest to discern when presented in audio, they were still able to classify CA states correctly from an audio representation.

There was a significant main effect of presentation mode on accuracy scores for both groups (non-musicians, df = 1.419, F = 66.915, p = 0.000; musicians, df = 1.269, F = 22.583, p = 0.000). Pairwise comparisons (using bonferoni adjustment) show that for non-musicians, scores in the visual-only condition were significantly higher than those of both other conditions ($p < 0.001$), and that audio-visual scores were greater than those of the audio-only ($p < 0.001$). For musicians, scores in the audio-visual condition were significantly greater than in both visual (p = 0.002) and audio (p = 0.000) conditions. Visual scores were also higher than audio (p = 0.006). This suggests that musical experience does have some effect on preferred display modality: the non-musicians performed best in the visual-only condition (suggesting that the audio actually put them off!), whilst the musicians performed best in the audio-visual condition.

4.3.6 Musical Experience Effects

The performance of the two groups suggests that they were able to discriminate CA classes from the auditory representation, but that experience and/or perceptual skills may affect the clarity with which patterns were perceived. In this instance, the audio mapping produced was not particularly straight forward, and it is of no surprise that both groups found the audio displays hardest to classify. As noted above, the traditional 2D graphic representation of the CA is effective as the recent history of system can be seen at a glance. The transient nature of sound means that the immediate history of the CA system is not as comprehensible in audio as it is in a 2D graphical display.

Reinforcement and interference in multi-modal displays

Of greater interest is the differential performance in the audio-visual condition (the musicians performed best in this condition, but the non-musicians scores were better in the absence of any audio cues). That the musicians performed best when presented with a multimodal display fits with research suggesting that redundant, or complimentary representations facilitate comprehension. Although this remains a contentious issue, benefits of redundancy in multimodal displays, principally in the education literature, have been made on the basis that multiple encoding, or cue-summation improves retention, recall, and understanding of contents (Findahl (1981), Drew and Grimes (1987), Severin (1967)).

Why then did the additional information available in the audio display decrease the classification accuracy for non-musicians ? One possibility is that comprehension of the audio display demanded recognition of harmonic and rhythmic patterns which were too complex for them to perceive accurately. Recall that the mapping produced not only a fairly straightforward rhythmic pattern, but also harmonic pattern that varied both in terms of intervalic structure and harmonic progression. Given the findings cited in section 4.3.2, it seems possible that the harmonic patterns in particular may not have been perceptually clear to an untrained ear. Confusion over the audio clues may have meant that the combined audio-visual display produced sets of contradictory, rather than complementary cues. There is strong evidence for the interference effects which arise when contradictory information is presented simultaneously to different senses. Perhaps most famous is the McGurk effect (McGurk and MacDonald (1976)), where perception of a speech phoneme is altered by dubbing it onto a video of a speaker saying a different phoneme. More recently, conflicting audio-visual cues have been shown to create perceptual bias (Sekuler et al. (1997)), illusions (Shams et al. (2000)) and even cross-modal after effects (Kitagawa and Ichihara (2002)).

Further trials are needed to make conclusive remarks but these findings suggest that it is possible to perceive high-order characteristics of complex systems when the system outputs are represented only in sound. However, findings also highlight the importance of considering musical experience when designing any mapping which aims to render data listenable. For those interested in auditory display as a visualisation tool, the impact could be enough to render the tool useless. In artistic application these results perhaps serve to remind composers that the musical patterns and morphologies which they aim to create and may be able to perceive, may not be so evident on first hearing or to an untrained ear.

4.4 On Mapping and Model Selection

The mapping used undoubtedly had an effect on the ease with which both groups could make a classification. The central importance of mapping design is well recognised within the field of algorithmic composition, and is also a significant area of investigation within auditory display research.

For sonifications developed within the field of auditory display, the main focus is on the development of intuitive and unambiguous mappings. In musical applications, we may not want to be so literal, but research findings in this area raise some issues worthy of consideration. Design of auditory displays for data analysis focuses on the psychological meaningfulness of the resulting signal. Currently, most mappings reflect subjective preference, at best evoking common metaphor - such as increases in frequency with temperature - in an attempt to produce mappings that are compelling (Kramer et al. (1997)). Such metaphors are limited however and the mapping procedure for most variables is far from intuitive (Walker and Kramer (1996)). Differences in specific data-sound mappings have been shown to affect reaction time and accuracy in monitoring tasks (ibid). However even for common physical dimensions, there seems to be little consensus over preference for particular mappings or their direction (Walker et al. (2000)).

Perceptual Interactions Within Display Dimensions.

Even when intuitive mappings are developed, the limited number of orthogonal dimensions in sound space potentially create perceptual interactions which can distort the way relations within the data are perceived. Numerous studies have demonstrated that the auditory dimensions of pitch, loudness and timbre interact perceptually (e.g. Melara and Marks (1990)). Even within one dimension, there appear to be perceptual asymmetries for rising and falling intensities of equal magnitude, e.g. subjects report larger absolute changes in volume when it is getting louder than when it is getting quieter (Neuhoff (1998)). Research has shown that these same interactions and asymmetries occur even when mapped onto data dimensions (Neuhoff et al. (2000)). Values of stock prices and trading volumes were mapped onto pitch and intensity of an audio signal, and participants were instructed to make judgments of relative changes in trading figures according to perceived changes in the sounds. When both auditory dimensions changed in the same direction, perceived variation in the target variable was reported to be greater than for incongruent changes of the same magnitude.

Timbral parameters are similarly susceptible to interaction, such that linear changes can have unpredictable, non-linear perceptual effects. For example, our perception of the brightness of a sound is determined by several factors including the attack time, and spectral evolution. This means that a bivariate display, in which one variable is mapped to the position of the spectral peak and another to the attack time of a static harmonic tone will not be heard as a simple 2D space, as many different combinations of these two variables can create a perceptually equivalent level of brightness. Indeed it has been suggested that a true balanced multivariate parameter mapping may not be possible in practice (Kramer (1994b)).

Although these interactions may cause problems if data is mapped to continuous parameters, the use of discrete timbral variations can be effective. Using contrasting acoustic textures, much like employing different colors in a graphical display, increases the number of dimensions that can be represented by high level audio dimensions and if carefully designed can prevent masking effects, allowing attention to be equally divided.

Preservation of key characteristics

Despite insights from auditory psychology studies, we are far from any comprehensive 'theory' of mapping. Currently the community operates by rules of thumb such as "relevant changes in the data should ensure a change in what is perceived. Changes in what is perceived should signify meaningful changes in the data." (Barrass and Kramer (1999), p.25). Although this may sound like a truism, it serves as a useful reminder to any composers using sonification methods to consider which dimension of sound can best carry the structures they wish to present, or conversely, which types of systems produce dynamics most suited to the domain they are interested in structuring.

For example a more effective means of representing the evolution of the patterns in the 1D CA used in the study above may be to map each element in the array to a pitch value, and present each row synchronously at audio rate (i.e. greater than 20 Hz). Patterns in the data would then be perceived as timbral, rather than rhythmic and melodic variations. The periodic patterns arising from ordered rules, would produce a more harmonic tone, chaotic patterns producing a more noise-like signal. Such a mapping would preserve the inherent synchronicity of the system and go some way in overcoming the lack of persistence of sound. Other researchers exploring musical application of CAs similarly report that they are more successfully applied in the synthesis domain.

Perhaps the most published CA-based music and sound applications are those of Eduardo Miranda. He used different 2D CAs to create both harmonic fragments (*CAMus*), and as a granular synthesis engine (*ChaosSynth*) (e.g. Miranda (2000b)).

In *CAMus*, two different CA rule sets running on separate grids are used to define the orchestration and placement of notes in pitch and time. One set of rules, Conway's *Game of Life*, consists of binary cells, which form characteristic discrete configurations. For example blinking crosses, static boxes or the infamous glider, a set of five cells which traverses the grid. In the other rule set, *Demon cyclic space*, cells can take one of seven states. From initially random configurations the system settles to produce stable patchwork patterns (shown in Figure 4.8).

The Game of life rules are used to determine a three note chord by transforming the cartesian coordinates of any given live cell into successive intervals above a user defined root. This is shown in Figure 4.7. In this example, the user has chosen the note G2 as the root note and cell at location (19, 7) was alive, the other two notes are D4 (19 semitones above G2) and A4 (the note 7 semitones above G2). The time intervals between these notes are determined by the states of neighbouring cells. The three notes are then voiced on (MIDI) instruments defined by the state of the corresponding cell in another 2D CA described by the rule set demon cyclic space. If the user had defined an oboe to orange, and the cell at position (19,7) on the demon cyclic space grid was orange, then the triple G2, D4, A4 would be voiced as an oboe.

In *chaosSynth* a CA rule which mimics chemical oscillations is used to parameterise a granular synthesis engine. These cyclic CAs evolve from a random state to produce spatial oscillations, mimicking the pattern formation seen in some chemical reactions. The granular engine consists of a bank of oscillators each of which are associated with specific groups of cells. Each cell can takes a continuous value, which determines its

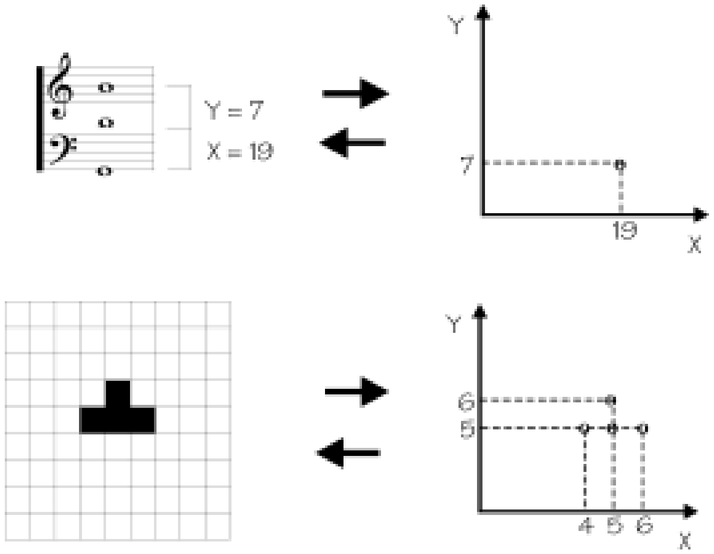


Figure 4.7: Mappings used in CAMus. The cartesian coordinates of a live cell are mapped to a triple (top). Each iteration of the rule set produces a number of such chords (bottom).

state as quiescent, depolarising or burned according to whether it is below, between or above certain minimum or maximum thresholds. Cell values are mapped to frequencies, and the amplitude and frequency of each oscillator is determined by the arithmetic mean of the associated cell group. The duration of each sound is determined by the number of configurations produced by the automata and the (hand set) grain length.

Comparing the results of the two systems, Miranda has concluded that CAs are more effective as a tool for sound synthesis rather than operating at the higher note level (Miranda (2000b)).

> "In general, we found that *Chaosynth* produced more interesting results that *CAMus*. We think that this might be due to the very nature of the phenomena in question. The inner structures of sounds seem more susceptible to CA modelling than large musical structures." - Miranda (2000b), p.5

The dynamics of the chemical oscillator CA rule, as it evolves from a random state to sustained oscillation, bear strong resemblance to the morphological evolution of many acoustic instruments where partials converge from a random distribution to oscillatory patterns (see Figure 4.9). The mappings used to parameterise the granular engine preserve these characteristics, so the sounds produced similarly bear these morphological features. However, Miranda himself writes that the mapping used in CAMus is arbitrary. Even if we saw some musical relevance to the blinking and gliding characters in the Game of Life, the mapping does not preserve these dynamics in a way that the listener can comprehend. It is not necessarily true then that the inner structures of sounds in general are more susceptible to CA modelling than larger musical structures. Just that in ChaoSynth, the model used captured key characteristics of the musical phenomenon it was applied to, and the mapping used preserved these characteristics.

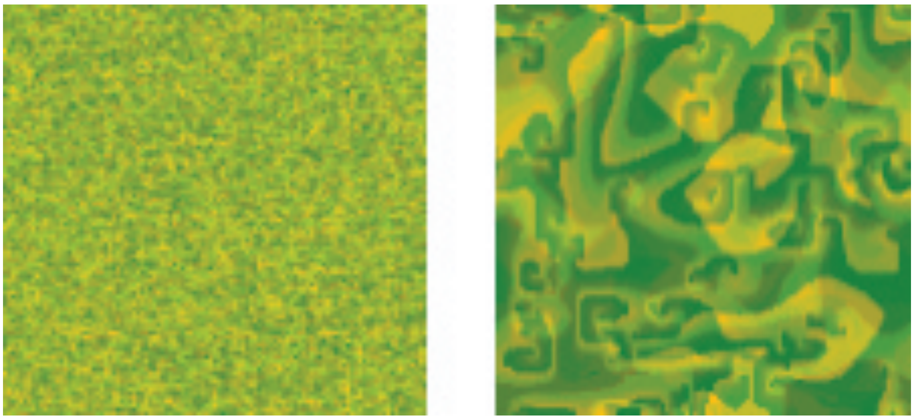


Figure 4.8: Evolution of CA used in Chaosynth: initial random distribution of cells (right) evolves to an oscillatory pattern (left).

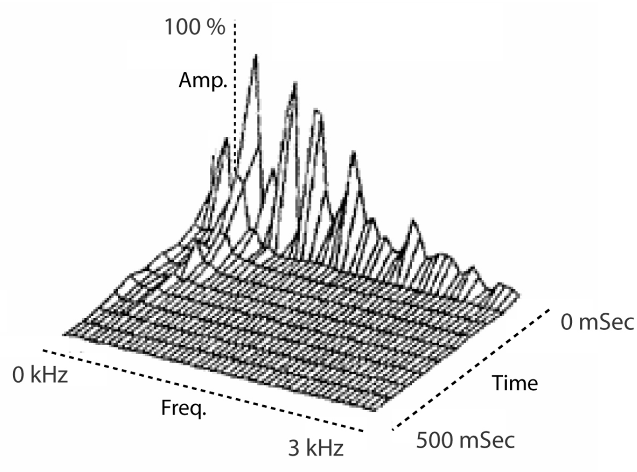


Figure 4.9: 3D wire sonogram showing evolution of spectra from initial white noise to 2nd, 3rd and 4th harmonic for a Mridangam (east Indian drum) stroke.

4.5 Summary

This chapter considered the degree to which the attributions of intentionality which Alife art invites are bound up in the visual presentation. Just as the quality of algorithmic music is determined in part by the mappings used, so successful Alife art may be due to visual cunning on behalf of the artist. The relative success of visual work in this area, in terms of high profile public appearances, suggests that there is some kind of inequality between visual and musical applications of Alife techniques.

The results of the study presented in this chapter represent a first systematic step into exploring the most basic source of this inequality: that Alife-type systems simply can't be heard. Of course the results of this study can't be generalised. Just because these people could identify these particular CAs states under this particular mapping, it doesn't mean that all aspects of Alife phenomenon can be perceived in audio. And just because something can be recognised, it in no way guarantees its resplendence as a musical device. However it seems important to perform such basic tests to ascertain at least that the complex dynamics of some Alife systems *can* have any phenomenological reality in sound.

The results of this experiment led to a discussion on the importance of mapping and model selection. There are various cues we can take from literature in auditory perception concerning perceptual interactions etc. Almost all discussion of algorithmic composition includes somewhere a line saying how mapping is the key. This is of course important, but as these examples from Miranda were aimed to illustrate, before we think about mapping, we need to think carefully about the peculiarities of the model we are using and the musical effect we wish to make. Different models are suitable for different jobs: some may not be suitable for anything, some may be suitable for lots of things, others may need adjusting slightly. The next chapter presents a set of 'studies', exploring a range of mappings for a variety of simple adaptive dynamical systems.

Chapter 5

Studies in Simple Adaptive Dynamical Systems

This chapter introduces some specific examples of simple adaptive dynamical systems. The term 'simple adaptive dynamical system' is far from elegant but aims to mark out a class of system that exhibit characteristics for generative and interactive activities. The *dynamical* approach was introduced in Chapter 2 where it was suggested that it was highly suited to subserving a rich continuous flow between human and machine components of a performance network. Reviewing algorithmic composition in Chapter 3, it was also suggested that formal dynamical systems could be successfully applied to create a strong sense of linear impulse and, at least at the timbral level, a perceptually strong dynamic morphology, making them useful as compositional algorithms. Consideration of the characteristics necessary to realise the richer conversational model of interaction also ear-marked the importance of *adaptation*, in providing a coherent, internally generated response. Following the examination of Alife installation art in Chapter 2, it was suggested that the many layers of adaptation in the evolving ecosystems created a system too complex and unresponsive for live music purposes, and it was proposed that the simpler, single agent systems offered a more suitable model. The *simple* qualifier here then aims to reign in the boundaries, these will become clearer with the illustrations presented below.

- The audio examples discussed in the text can be found on the accompanying DVD, tracks 1- 17.
- Max/MSP of objects for most of these models are also available on the DVD, along with help files that illustrate the basic mappings described here.

5.0.1 Models

One of the propositions of this thesis is that adaptation is not only something to be considered in an interactive context, but that adaptive responses present a useful device in generative *composition* practice. Generative art is typically discussed in terms of designing a process. Processes like L-systems or CAs can be specified, and unfold to generate a particular structure. At the other extreme, part of the fascination with evolutionary processes, is that they can create outcomes which exceed the expectations of the artist, surging off into the computational sublime. This project aims to carve a middleground, retaining the coherent unity of the L-system as it develops through time, but introducing multiple sets of parallel processes which influence each others' path. The focus of interest lies between the generation of 'structures' (as in Xenakis' interest in 'out of time structures') and 'composing interactions' with a sonic by-product (as in Di Scipio's AESI) and pushes toward composing interdependent, reactive structures 'in time', or more appropriately, 'behaviours'.

A handful of models have been selected, each of which are dynamical systems of some form, and each of which exhibit some level of adaptation to environmental input which is observable as a characteristic pattern of behavioural response. These have been appropriated directly from, or inspired by, cybernetics, Alife and ecology. The systems include models of homeostasis, entrainment, pattern propagation and population distributions. Similar techniques may have application for learning, searching, or problem solving within engineering AI or AI music approaches. Here however their potential as adaptive pattern generators is explored.

Models are considered individually in this chapter and a range of mappings are explored. Consideration is given to both specific compositional goals, and the ways in which their adaptive characteristics can be employed in both compositional and performance situations. In all instances, the aim was not to replicate any specific musical style or idiom directly, but to attempt to create a sense of musical life and coherence by using systems which can be seen to exhibit some degree of goal-directedness and/or adaptation to their environment. This goal-directedness seems a first important step in creating digital systems which exhibit some degree of independence, and ultimately a musical 'personality'.

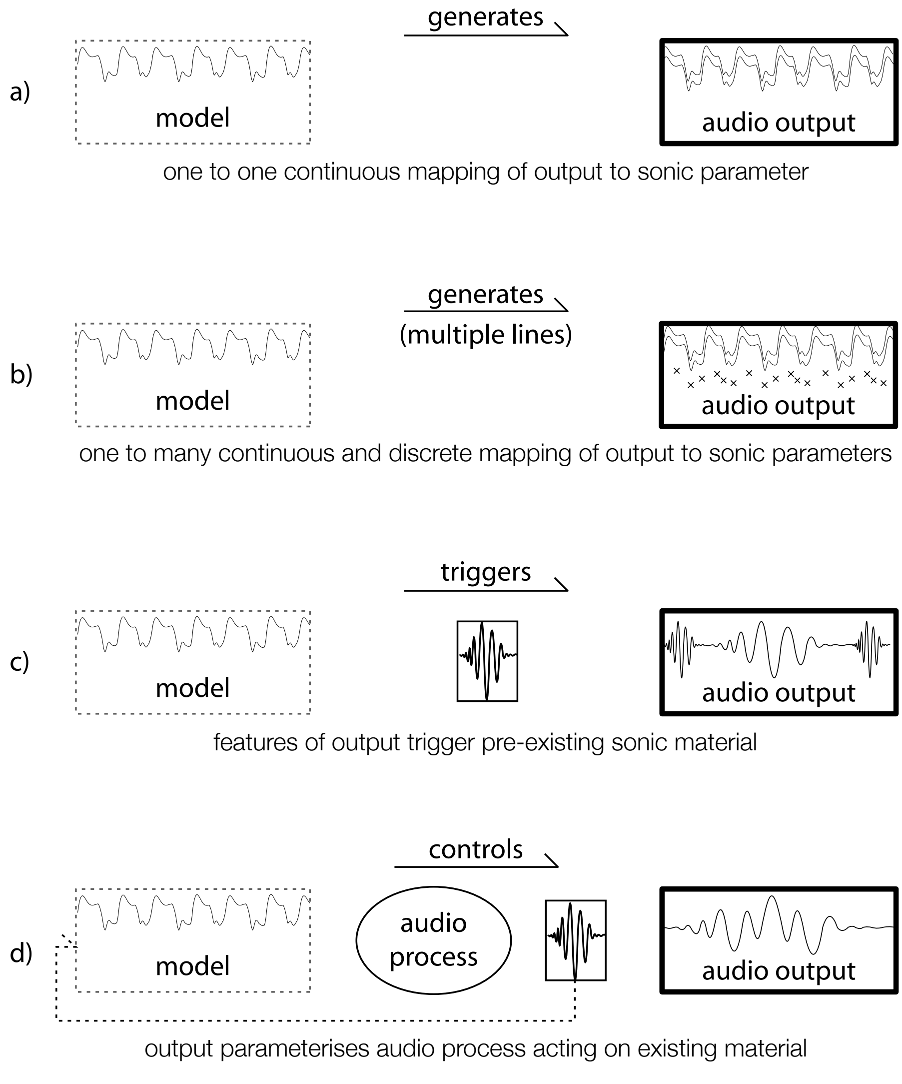
5.0.2 Mappings

In previous chapters it has been suggested that the responsive characteristics of such models make them attractive as interactive mechanisms in live generative performance and may also provide rich dynamics that are potentially capable of generative interesting musical material. In other words that adaptation has potential as a compositional as well as an interactive mechanism. In order to explore whether this is true, a number of different mappings were examined for each model. This project departs from the approach taken by champions of algorithmic composition such as Xenakis or Roads.

In some of the most successful examples of algorithmic composition (e.g. Xenakis (1971b), Roads (2001)) formal processes were developed for specific compositional situations. In these cases the algorithm and the mapping are tightly intertwined. We could almost say that with stochastic systems as Xenakis' *GENDYN*, what we hear is the process itself: a direct sonification of the stochastic models. In a very different way Di Scipio's AESI also presents the process itself, although in this case it makes little sense to talk of a process distinct from its sonification. In both cases, the process has been designed with a very specific compositional aim and this aim defines the mapping.

The current project is motivated by a broader aesthetic aim: a desire to create a form of *behavioural* generative system for performance and composition. The proposal is that the dynamics of simple adaptive systems are capable of evoking a minimal sense of agency or goal directedness that invites an attribution of intentionality, or personality. As discussed in Chapter 4, certain algorithms may be more or less suited to structuring particular levels of musical material. The aim of this chapter then is to explore some different ways of mapping a range of models in order to ascertain primarily, whether any of the models are effective at all.

In order to structure the explorations, mappings were explored at different levels of complexity and at varying degrees of remove. These are summarised in Figure 5.1. In the simplest case (Figure 5.1.a), the numerical outputs are directly sonified, for example being used to specify the pitch of an oscillator. In this case the model is used directly to *generate* musical material. This approach tarries with that of a data visualisation exercise and allows immediate appreciation of the basic form of a model's dynamics. Rather than mapping every data point, certain characteristic features can be used to generate short events. Combinations of continuous and feature-based mappings can be also be used in conjunction to create *multiple mappings* (Figure 5.1.b). This is an effective way of producing multiple parts that are closely related. These two approaches can also be applied to sample-based sonification: features can be used to *trigger* short samples, or the full data stream can be used to continuously manipulate some aspect of pre-specified sound material – for example continuously altering the playback speed of a sample (Figure 5.1.c). Alternatively the outputs can be used to *control* some other audio process acting on existing (or generated) audio, such as a filter or other effect (Figure 5.1.d).



one to one continuous mapping of output to sonic parameter

one to many continuous and discrete mapping of output to sonic parameters

features of output trigger pre-existing sonic material

output parameterises audio process acting on existing material

Figure 5.1: Outline of the mapping techniques explored. Outputs of the model are used to: generate material directly creating either single lines (a), multiple different but related lines (b); trigger pre-existing sonic material (c); or to parameterise some other DSP process (d).

5.1 Adaptation and Homeostasis

In the 1950s, Cybernetician Ross Ashby built an electro-mechanical machine called the homeostat. By all accounts the thing itself was an engaging machine, but its notoriety in certain circles is due to the theoretical ideas that it incarnated. One of the conundrums that preoccupied Ashby, was how a system (biological or mechanical) could be at once state determined, and yet adapt to a changing environment and learn. Ashby (1952) proposed that one of the key mechanisms underlying adaptive behaviour is homeostasis, and like all good cyberneticians, provided a concrete, physical device to demonstrate his theoretical notion of *ultrastability*.

Adaptive behaviour is a major research topic in contemporary cognitive science, and indeed the importance of homeostatic adaptation is re-emerging in philosophical circles as a key aspect for understanding of life, mind, autonomy etc. (e.g. Di Paolo (2005)). Basic homeostatic adaptation is the starting point for the current exploration of adaptive systems for interactive and generative music. Iconically and practically then, Ashby's homeostat provides inspiration for one of the central conceptual and algorithmic devices used throughout the projects presented here. The term homeostasis was coined by Canon to describe the internal self-regulating mechanisms of biological organisms which maintain essential variables such as blood temperature, pressure and sugar levels in a dynamic balance. Cyberneticians such as Wiener (1948) and Rosenbleuth (1943) provided us with a systemic understanding of the patterns of organisation subserving adaptation and homeostasis- i.e. self-correcting negative feedback loops. The process is illustrated by every day examples such as thermostatically controlled heating systems or lavatory stopcocks and was expressed by Wiener (1948) in a characteristically wordy statement:

Figure 5.2: Ashby's electro-mechanical homeostat.

> "When we desire a motion to follow a given pattern the difference between this pattern and the actually performed motion is used as a new input to cause the part regulated to move in such a way as to bring its motion closer to that given by the pattern" - Wiener (1948), p.6

Ashby advanced the concept of a self-correcting feedback system in his theory of self-regulating *ultrastability*. He defines an ultrastable systems as one that is able to reconfigure plastically in response to any of its essential variables going out of bounds. In a self-correcting system the relation between the input carrying the signal error and the regulation device is fixed (the ballcock in a cistern is attached to a stiff rod connected securely to the valve). An ultrastable system exhibits a higher order stability which allows self-regulation of the regulatory mechanism itself (a cistern which could change the position of the nut on its rod, or even invert the relationship between the angle of the ballcock and the valve). Ashby illustrates the difference by inviting us to consider the mechanisms controlling an autopilot. A standard autopilot might consist of a gyroscope connected to the airelons on the aircraft wing: if the craft banks in one direction, the gyroscope measurement induces the necessary change in the airelons to roll the craft back to horizontal. If the connections between the gyroscope and airelon were reversed, the smallest bank in either direction would be amplified: the autopilot would implement positive rather than negative feedback and this would continue until the craft crashed.

The higher-order stability central to Ashby's concept of ultrastability refers to a system which would be able to adapt to, and compensate for, this reversal of connections. In this case, once the roll reached a certain critical magnitude, the connections between gyroscope and airelon would themselves invert until the roll was corrected and the aircraft restabilised. In order to achieve this Ashby argued that a system *necessarily* requires a mechanism consisting of a primary direct feedback between sensorimotor system and the environment, and a *secondary* feedback, operating intermittently at a longer timescale, between the essential variables and the sensorimotor system. It is this secondary feedback system which reconfigures the sensorimotor connections when the essential variables exceed their limits. Ashby's mechanical homeostat was a physical proof of concept for this theory of ultrastability.

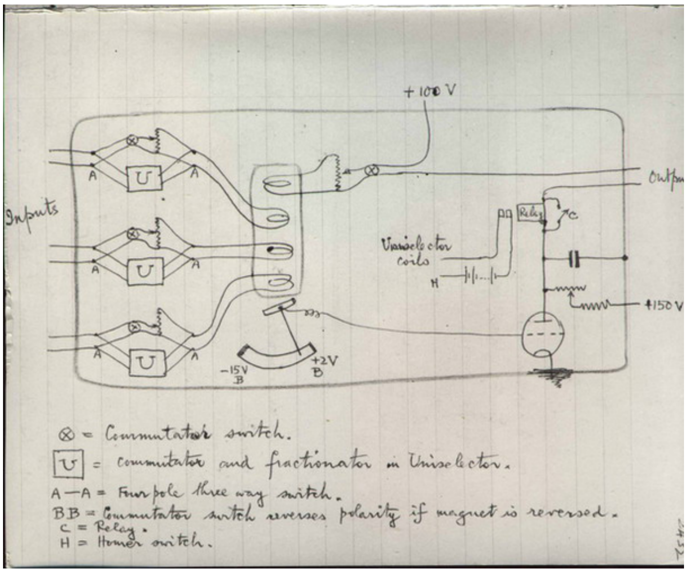


Figure 5.3: Diagram of part of the homeostat circuitry from Ashby's notebook.

The machine consisted of four units with a pivoted magnet on top of each. The angular deviation of each magnet's position representing the essential variables which were to be maintained within 45°. Each unit sends a current proportional to the deviation of its magnet from the centre (no current being sent when it is centred). This was achieved by dropping a wire from each magnet into a trough of liquid with electrodes at each end, so providing a potential gradient. The wire therefore picks up a graded potential depending upon the position of the magnet. The viscosity of the liquid in the troughs affects the behaviour of the homeostat: highly viscous liquids creating a turgid, stable system, more fluid liquids producing wilder, more fluctuating behaviours which take longer, if at all to stabilse. These electrical connections model the primary feedback, where any one unit can be conceptualised (arbitrarily) as representing either the environment, or the sensorimotor system of an agent in that environment.

The units were joined with connections between each magnet. The connections operated via coils where the torque on each was proportional to the sum of the currents in connected units. Each unit also had a recurrent connection. The current on each was modified by passing it through a commutator and potentiometer which determined the polarity and proportion of each input which is passed. These act as parameters to the system, implementing a secondary feedback which was controlled by a uniselector on each unit. The uniselector has twenty-five discrete states, each consisting of a three random values derived from a standard statistical table. Each uniselector checks the value of the outputs of its daughter unit, assigning new values to the commutator and potentiometer if the magnet's angle of deviation exceeds the critical value of 45°. The new values affect the movement of a unit's magnet, and so change the potential that is passed to connected magnets.

> "When these parameters are given a definite set of values, the magnets show some definite pattern of behaviour; for the parameters determine the field, and thus the lines of behaviour. If the field is stable, the four magnets move to the central position, where they actively resist any attempt to displace them. If displaced, *co-ordinated* activity brings them back to the centre. Other parameter-settings may, however, give instability; in which case, a 'runaway' occurs and the magnets diverge from the central positions with increasing velocity - till they hit the ends of the troughs" - Ashby (1952), pp.102-103

By a process of trial and error, the machine is able to maintain its essential variables within specified limits. Ashby also demonstrated that the machine could exhibit basic reinforcement learning, adapting to alternate environments and presented it as an example of basic self-organisation.

Wiener (1967) described the homeostat as "one of the greatest philosophical contributions of the present day" (p.54), but it was not without critics. Grey Walter (1953) dubbed it the *Machina Sopora*, suggesting that if it were to be judged entirely by its behavior, the naturalist would classify it as a plant (p.124). Another fair criticism which has been raised is that the mechanism used to achieve homeostasis (i.e. random search) is incredibly inefficient and unpredictable. In Ashby's electro-mechanical device, there are 25^4 $(390,625)$ different combinations of uniselector parameter values that a four unit homeostat can randomly explore in order to find a combination that leads to stablility. This prompted Singh's (1966) critical description of the homeostat as a 'permutational orgy'. As well as taking an incalculable length of time to stabilise, the system is incapable of accumulating adaptations, i.e. once it has achieved a certain behaviour, the stochastic nature of adaptation makes it likely to be lost irretrievably as Ashby puts it:

> "In general, if the Homeostat is given a problem A, then a problem B, and then A again, it treats A as if it had never encountered A before; the activities during the adaptation to B have totally destroyed the previous adaptation to A." Ashby (1952)

The ultimate goal of the device was to maintain consistency in the face of change, which may not seem like a very interesting musical attribute. Its indeterminacies revoke consideration of its employment as a robust learning device, particularly in a real-time situation. But the basic adaptive and dynamical process by which it achieves homeostasis is appealing.

The system illustrates the appearance of unpredictably complex behaviour arising from the interactions of simple devices. The internal adjustments made provide a minimal form of goal directed behaviour: the homeostat behaves *as though* it were seeking to keep its magnets in central positions. Despite its basic mechanical, deterministic substrate, the system exhibits open ended and unpredictable, yet coherent behaviour.

> " . . . but what strikes me about them is their singular liveliness. I can't actually think of any prior example of a real machine that would randomly - open-endedly as I would say - reconfigure itself in response to its inputs. When I think of 1950s machines, I think of lathes, drilling machines and whatever - deterministic devices that either respond predictably to commands or just break down and never work again. It seems reasonable, then, to speak of the homeostat as having a kind of agency - it did things in the world that sprang, as it were, from inside itself, rather than having to be fully specified from outside in advance." - Pickering (2002)

From a practical creative perspective, the system offers an attractive balance of autonomy and controllability. System behaviour arises from an internally controlled, open-ended configuration, but is parameterised by the degree of viscosity. Although it is 'doing its own thing', we can induce it to operate within a given field. The characteristically different responses to different forms of input displayed also provide a form of global control. Finally as will be discussed below, as a modular system, the size of the network and degree of interconnectivity have significant impact on its behaviour, and can be engineered for specific tasks.

5.1.1 A Model of the Homeostat

The key aspects of the machine were simulated in a neural-network style model. The machine is conceived as a network of I units, each connected to J other units (shown schematically in Figure 5.4) where the output of each unit is updated according to the weighted sum of the output of all other nodes as shown in Equation 5.1 (these weights modelling the potentiometers and commutators described by Ashby). In this simulation if the output of any node exceeds a prespecified value, weights connecting units in the network are re-randomised, simulating the role of the uniselectors in assigning the system parameters. As in Ashby's machine, the recurrent connection is held constant. Investigation showed that the frequency of uniselector action (i.e. testing outputs) did not have any effect on the major properties so it was held constant and outputs were checked at every iteration. Viscosity was implemented by constraining the amount by which any one unit could move between iterations.

$$\mathrm{O}_{i(t+1)} = \sum_{j=0}^{j} I_{ij(t)} \times W_{ij(t)} \quad \text{where} \quad I_{ij}(t) = \sum_{j=0}^{j-1} O_{j(t-1)} + O_{i(t-1)} \quad . \tag{5.1}$$

Where $O_{i(t+1)}$ is the Output of the i_{th} unit at time $_{t+1}$, $I_{ij(t)}$is the input to the i_{th} unit from the j_{th} and $W_{ij(t)}$ is the weight from unit j to unit i.

5.1.2 Homeostat Behaviour

This basic model is capable of replicating the principle characteristics of Ashby's homeostat. Primarily, once stable it will actively resist small interferences (the primary feedback mechanism bringing all outputs back into line), large perturbations trigger weight changes representing the secondary feedback mechanism which reconfigures the uniselector action in Ashby's machine. This is shown in Figure 5.5. Once stable, the system

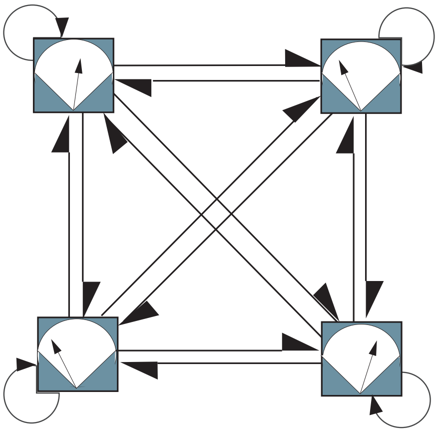


Figure 5.4: Schematic of a fully connected four unit homeostat. Each unit is represented as a square box, its output being the deviation of the small arrow from the centre. Weighted connections between units are represented by the uni-directional arrows which link each unit.

exhibits a minor transient response to perturbation below critical limits (marked **a** in Figure 5.5). At point **A**, the output of unit one was forced *outside* its critical limit. This causes weights changes, and the system enters a different stable state. Note also that when stable the system sometimes converges to a point attractor (iterations 0 - 500 in Figure 5.5), or oscillates in limit cycles, (as in iterations 500 - 1000 after critical perturbation) often each node entering a cycle of different lengths. This can be used to generate basic poly-rhythmic patterns.

In the original physical machine, the degree to which the system state was historically determined was controlled by the viscosity of the liquid in the troughs in which the outputs trailed. This damping effect was modeled by restricting the variation in outputs in any one unit from one iteration to the next. The effect of changing the value of this variable proved similar to the assumed effect of varying the viscosity of a liquid: low values (representing high viscosity) result in turgid, stable behaviour; high values produce more exploratory ‘run-away’ behaviour as each unit does not have time to achieve stable parameter settings before other units transgress the critical limits. This is demonstrated in Figure 5.6 (left) which shows stability as a function of viscosity. Here stability is measured as the time taken for all units to stabilise from an initially random weight selection. In a later paper, Gardner and Ashby (1970) also discussed the effect of network size and connectivity on the stability. Figure 5.6 (right) replicates his results, showing the inverse relationship between stability and either size or connectivity of network.

5.1.3 Example Mappings from the Homeostat

Simple pitch control

The basic behaviour of the homeostat can be heard clearly if the outputs are mapped directly into pitch deviations as in Figure 5.1.a.

- In Track 1 the outputs of a ten unit homeostat control the frequency of ten sine wave oscillators, offset by a small amount to increase clarity. The initially unstable network settles with each input entering a limit cycle of a different length. This produces a minimal poly-rhythmic loop.

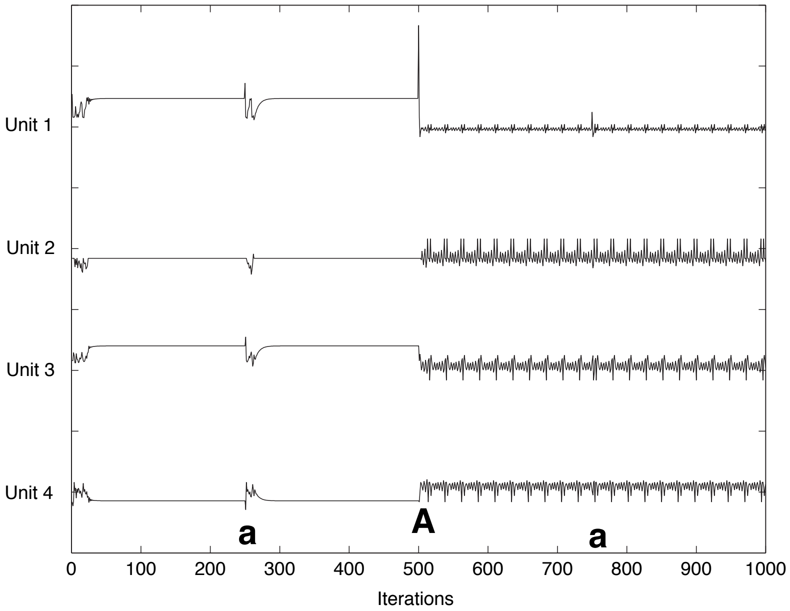


Figure 5.5: Outputs of a four unit homeostat demonstrating stability to minor perturbation (**a**) and re-stability after critical perturbation (**A**).

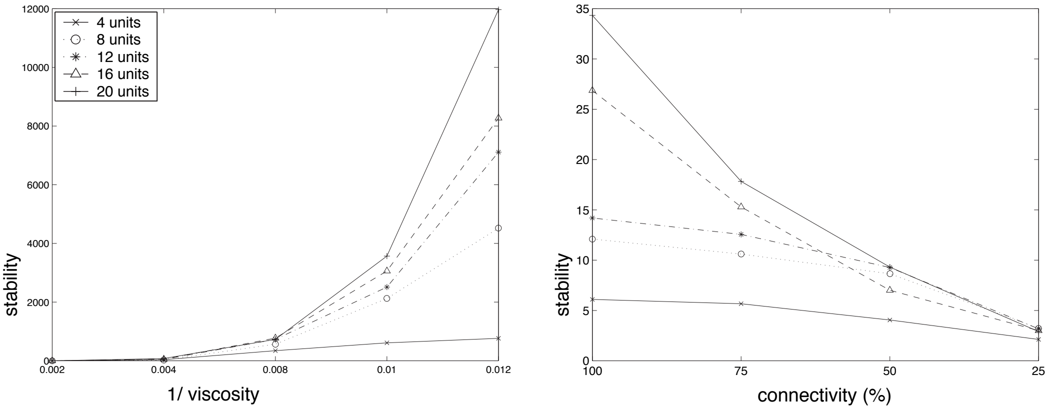


Figure 5.6: Change in stability as a function of number of units and 1/viscosity (left) and as a function of connectivity (right). Stability is taken as the point at which all units remain inside limits (and therefore weights remain constant), and measured as the number of iterations taken to achieve this state, averaged over 200 runs.

- Track 2 illustrates a similar mapping made using MIDI. Here the outputs of a four unit network are mapped to pitch bend, producing microtones of $\frac{1}{32_{th}}$ tone. The effect of applying a small input to a stable network can be heard: at around 10″, the regular pattern deviates for a few cycles and is then reinstated. This track also illustrates the effect of employing multiple mappings. As well as mapping outputs to a continuous pitch variation, a 'melody' line is created by using the output of each unit to specify the pitch of a percussion instrument. The timing for each unit is determined by selecting random number N for each unit in the range (2,10) and voicing its pitch every n_i beats. This creates a strange harmonised melody line.

This approach was used in the *AdSyMII* installation described in Chapter 6, and also forms the basis of a track *Sines* which was commissioned by generative film makes Iain Helliwell for the LUX Open 2002, a festival of experimental film at the Royal Art College. Here the homeostat is used to simply control the pitch of a set of four sine oscillators, giving an organic feel. This is given on Track 3.

Splicing and remixing audio samples
As well as determining low level musical attributes, the homeostat works well as a method of re-mixing existing audio material. In this example, the output range of the homeostat is scaled to the length of an audio track. The original piece is *Planting trees, creating beauty* by Norweigan trumpeter Arve Henriksen, an excerpt of which is given on Track 4. At each update, a short sample from the source material is triggered by each unit, the position being determined according to the value of each output. Rather than each output playing its selection on every beat, the sound is thinned out by specifying that some outputs only play when negative, some positive. This creates changes in density as well as changes in content. A similar mapping process was used in the Self-karaoke system and is described in more detail in Chapter 8, Section 8.2.2.

- Track 5 gives an example. The network is initially stable, each unit in a fixed limit cycle. This causes each to repeatedly play back the same few sections of the original file. At 15" a large input is applied to unit one, triggering weight changes and causing the network to rapidly settle into a fixed point attractor. The network is perturbed again, once more settling to a limit cycle. Over the next minute, a series of small perturbations cause a sequence of deviations from a repetitive cycle until around 1'15 the viscosity is turned right up. This causes all units to rest at a similar value, all triggering the same quiet section of bowed metal. The network is perturbed once more, and the viscosity turned down, making the system more excitable, and causing it to take longer to stabilise. At 1'45 you can hear the units converge, this time reiterating a vocal sample, until the network is perturbed a final time just before the end. The homeostat is iterated at 160ms intervals giving the rhythmic pulse which can be heard.

Spectral Filter Automation
Even when in 'Machina Sopora' mode when the homeostat settles quickly to a point attractor its dynamic response can be put to good effect. Track 6 gives an example where the outputs are used as an 'automated effects' device. Here the change in the outputs of the four units are scaled, and used to control the amplitude of the first 30 bins of a spectral filter, the remaining set at zero. The filter works by performing a Fast Fourier Transform (FFT) on an incoming audio signal and splitting the signal into a number of bins. The amplitude of each can be individually controlled[1]. In this example, the audio input is Morton Feldman's *Piano Piece for Three Hands*, which is provided dry in the example along with the filtered output. Rather than applying an input by hand as in the case of the examples above, the amplitude of each attack in the piano part is analysed and used as the input to unit one of the homeostat.

- In track 6, the viscosity is set high so the system settles quickly whenever perturbed. Each attack therefore triggers a very brief period of oscillation, heard here as spectral fluctuations after each note which die out between chords as the homeostat settles. Once settled, the entire spectrum of the filter is at zero, meaning that just the dry signal is heard. Notice also that quieter notes are insufficient to trigger

---

[1] This is essentially a fine grained graphic equaliser like you might have on your stereo to allow you to boost or cut bass, treble, mid etc.

the homeostat: as there is no change, the amplitude of each filter bin remains at zero here also, giving no wet signal. At around 1′10 the viscosity of the system is increased, resulting in larger spectral fluctuations which continue between notes. The increased activity of the homeostat means more bins have higher values at any one time, also increasing the overall volume level.

Rhythmic generation
The incommensurate lengths of the cycles into which the outputs often settle can be used to generate regular, if lopsided, rhythms.

- Tracks 7 and 8 give examples of a rhythm generator made for a live performance at Wrong Music[2], an organisation dedicated to experimental noise music. Here each output of an eight unit homeostat is used to trigger a different drum sample. The actual value of the output being used to determine the playback speed, giving rise to the variation in pitch which can be heard. Throughout both tracks, the system was repeatedly perturbed, producing both small variations from the principle beat, and larger changes in texture. In track 8, the update time was also manipulated creating gaps and dense sputters.

5.1.4 Summary of Homeostat Features

The homeostat exhibits a number of behaviours and features which make it attractive as a generative system for music. It exhibits a range of dynamics and characteristics which can be used to generate novel but arguably evocative material. As can be seen from Figure 5.5, when it stabilises it either converges to a single point, or to limit cycles, with each output often settling of a different length cycle. This in itself can be used to create complex polyrhythms. As also shown in Figure 5.5 it exhibits different responses to perturbation: small changes causing a temporary deviation from the current attractor, which is usually returned to after brief deviation, large inputs triggering weight changes which invariably lead to the system settling on a new attractor. The viscosity variable also enables global control over the nature of its dynamics: high values creating turgid, repetitive systems, and low values creating wild searching behaviour.

The exact output therefore can never be known, but the behavioural dynamics can be controlled on a qualitative level. Weights on the recurrent connections have a strong effect on the nature of the system's response to perturbation, and general behaviour. In this implementation, these are set when a new instance of the object is made. Again, although the effect of any one set of weights cannot be predicted, the idiosynchracies of any one configuration can be learnt in a more performative way by playing with the system. As these are randomised on initialisation, the random number generator seed can be saved so that 'favourite' configurations can be returned to. Despite Grey Walter's suggestion that this machina sopora is closer to plant than animal life, these characteristics provide a balance of autonomy and responsiveness which seems appropriate for the development of interactive and generative music systems. This basic homeostat is explored within a generative music system in Chapter 6, in an interactive installation in Chapter 7 and in a performance system in Chapter 8.

5.2 Entrainment in Neural Oscillators

Since the early 1980s, neural networks have been used in algorithmic composition, but invariably employed as pattern matching or learning mechanisms. Here a continuous time model was developed and used for the generation of musical material.

---

[2] `http://www.wrongmusic.co.uk/`

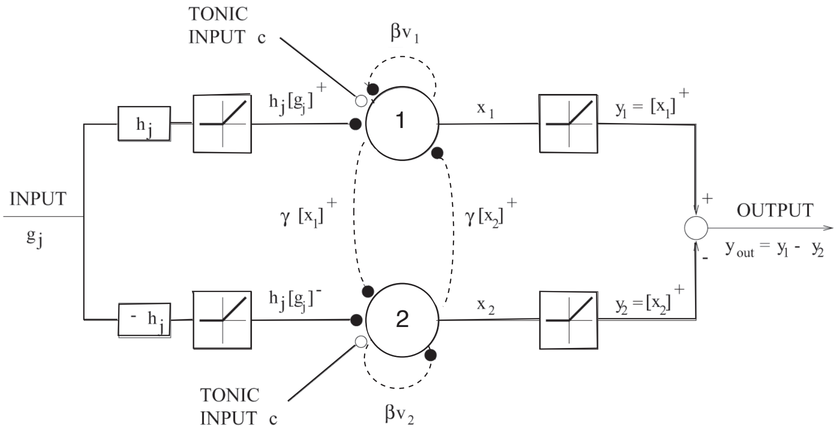


Figure 5.7: Schematic of a neural oscillator node. The oscillator equations simulate two neurons in mutual inhibition as shown here. Black circles correspond to inhibitory connections, open to excitatory. The mutual inhibition is through the $\gamma[x_i]^+$ connections ($[x]^+ = \max(x,0)$), and the βv_i connections correspond to self-inhibition. The input g_j is weighted by a gain h_j, and then split into positive and negative parts. The positive part inhibits neuron 1, and the negative part neuron 2. The output of each neuron y_i is taken to be the positive part of the firing rate x_i, and the output of the oscillator as a whole is the difference of the two outputs.

5.2.1 A Neural Oscillator Model

Neural oscillators are continuous time, real valued neuron models, arranged in pairs such that the output of one inhibits the activity of the other, creating an oscillatory output at a fundamental frequency. If a periodic input signal is applied to the pair, it will entrain the input frequency. When nodes are arranged such that the output of one node acts as input for other nodes, the frequency of oscillation across the network will be identical, although the phase and exact shape may vary. Using simple mappings into sound, this produces musical material that shares a common pulse or metre, but varies rhythmically. This property also means that the basic pulse can be set by an external (user controlled) input signal.

Neural Oscillators have been used in robotics tasks that require rhythmic movement such as sawing (Williamson (2002)), and drumming (Kotosaka and Schaal (2001)), and in models of rhythmic entrainment (Thaut (2003)). Here, a small network of simple neural oscillators was built, based on the model described by Matsuoko (1985).

The oscillator system consists of two simulated neurons arranged in mutual inhibition, as shown in Figure 5.7. The time evolution of the oscillator is given by equations 5.2 to 5.6, where $[x]^+ = \max(x,0)$. The output of the oscillator is y_{out}, β and γ are constants (here set to 2.5). c is a constant that determines the amplitude of the oscillation and τ_1 and τ_2 are the time constants that determine the natural frequency (in the absence of input), and shape of the output signal. Inputs (g_j) to the oscillator are weighted by gains h_j.

$$\tau_1\dot{x}_1 = c - x_1 - \beta v_1 - \gamma[x_2]^+ - \sum_{j=0}^{j} h_j[g_j]^+ \tag{5.2}$$

$$\tau_2\dot{v}_1 = [x_1]^+ - v_1 \tag{5.3}$$

$$\tau_1\dot{x}_2 = c - x_2 - \beta v_2 - \gamma[x_1]^+ - \sum_{j=0}^{j} h_j[g_j]^+ \tag{5.4}$$

$$\tau_2\dot{v}_2 = [x_2]^+ - v_2 \tag{5.5}$$

$$y_{out} = [x_1]^+ - [x_2]^+ \tag{5.6}$$

5.2.2 Neural Oscillator Behaviour

If an oscillatory input is applied, the node will entrain the input frequency i.e. it will produce an output of equal frequency, but not necessarily the same phase, as the input. This can be shown to be true over a wide range of input amplitudes and frequencies (see Appendix A, Figure A.1 for an illustration).

The fundamentally dynamic nature and specific behaviours associated with its entrainment properties make this model an attractive resource. The input driving signal can be given either by an external source, or from another software system, making it a useful component for the modular approach adopted here. Networks of oscillators exhibit a range of musically-relevant behaviours which are parameterised by a handful of variables. The sonic effects of changing these parameters is of course determined in part by the mapping and is discussed below in a simple case. In general, the fundamental frequency and form of the output can be controlled by the two time constants (τ_1 and τ_2). If run at audio rate, this can be used to generate audio signals directly. Iterating the model at slower speeds enables the generation of either melodic or rhythmic lines according to mapping scheme adopted. The entrainment property means that networks of these modules can create material of chosen degree of density, where each part bears a global relation to the whole. This creates parallel streams of data which retain their individual identity over time, but move in relation to each other.

5.2.3 Example Mappings from the Neural Oscillator

Pitch control

One of the simplest, and perhaps most effective, methods of sonifying this system is to simply map the output value of each unit onto a pitch value. When the bias of each oscillator node is between zero and one, the output will always be in the range (-1,1). This means the output can be easily mapped onto pitches in a chosen audible range. Figure 5.8 shows an example where the pitch has been quantised to semitones. The scored notes represent the waveform within the dotted box above.

The periodic oscillation of the node produces a basic arpeggiated effect. Under this mapping, changing the constant c varies the amplitude, and so pitch range of the line. Quantising the continuous output means that small changes in output, as well as fixed values result in a constant pitch. In the example shown in Figure 5.8 these repeated values were excluded, automatically introducing some rhythmic variation. The time constants affect the fundamental frequency of oscillation as well as its form, so can be used to alter the melodic contour of the output. Changing the absolute value of the weight between nodes as well as its sign determines the extent and nature of the influence of each node on connected nodes, changing the relations between parts.

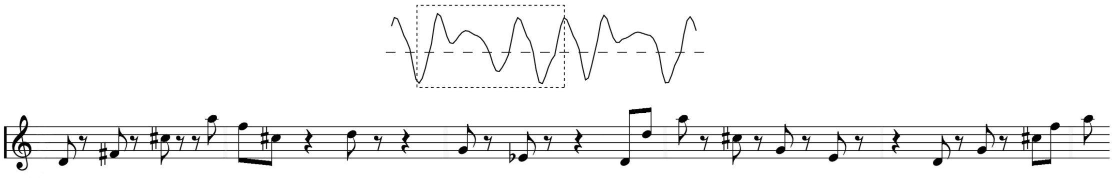


Figure 5.8: Mapping from a continuous output to quantised notes. The section of score represents the graphed output within the box and *above* the horizontal line only. Notes are only re-voiced if they have changed by more than a semi-tone across timesteps, creating the spaces shown here as rests.

- Track 9 gives an example of the basic arpeggiated line generated as well as the effect of inverting the weights between nodes. In the example here, the outputs of two nodes with the same bias but slightly different time constants are played on two pianos. Initially node two is played alone after five cycles (20″) the second piano enters. The weights are negative, causing the outputs to be in opposite phase, creating a sense of turn taking. The weights are then inverted at 55″ causing both parts to play in unison.

.

Applying an external input can have several musically useful effects. Primarily of course, if above a certain amplitude, it will determine the overall frequency of the system output. Continuous periodic input (such as a sinusoidal function) of low frequencies clamps the outputs of strongly connected nodes during the positive or negative parts (depending on the polarity of the weight). This causes the output to freeze at a particular value, being 'released' when the amplitude of the input drops. Sonically this creates the effect of a line pausing, or resting on a pitch, then 'coming back to life'. Finally although the external input entrains the overall frequency of output, characteristics of the fundamental oscillation are preserved. This produces an inner pattern which is modulated at the period marked by the main input.

- Examples of these effects can be heard on track 10. Again there are two voices here, a piano and a sustained synth sound. Initially the synth is clamped, repeating the same note. Once it comes in it takes a simple descending four note motif, which is modulated by the input frequency, altering the pitch of some of the notes in the internal structure. Here the synth sound is triggered only at local minima rather than continously, creating a bass line feel.
- Track 11 gives an example with four parts playing and demonstrates the effect of altering the input amplitude and frequency. At the start, four nodes are connected with different time constants and biases, giving each a characteristic shape. There is no input signal, so the frequency is determined internally by the nodes. From 20″ - 60″ the amplitude of the input signal is gradually increased. This has a differential effect on individual units depending upon how closely they are connected to it, and how strong their weights are. At 1′10, the frequency of the input signal is decreased, the longer period clamping the outputs. Here repeated notes are omitted so this audibly this thins out the parts. Finally at 1′50, the input is removed and the ensemble returns to its initial repetitive cycle.

This melodic mapping was used in the installation *Organised Entry* which is described in Chapter 6.

Rhythmic mappings
The network can also be used to generate rhythmic patterns by defining certain points in each oscillation to trigger a percussive voice. If a number of nodes are arranged in series, with an external input, the frequency of oscillation is constant, but the oscillations may vary in shape or phase. This provides a means of generating layers of rhythmic patterns with a constant metre, or pulse, but with a much greater freedom in the placement of individual beats than is common in most computer music. These discrepancies in timings can bring a human feel to the output, akin to expressive deviations from the beat. Equally however, deviations can make the output simply sound 'out of time'. In these examples, standard GM percussion instruments are triggered at either local minima, local maxima or at zero-crossings. An example is shown in Figure 5.9.

- A simple rhythmic example is given on track 12 with successive nodes voiced, demonstrating the possibility for generating conventional rhythmic patterns.

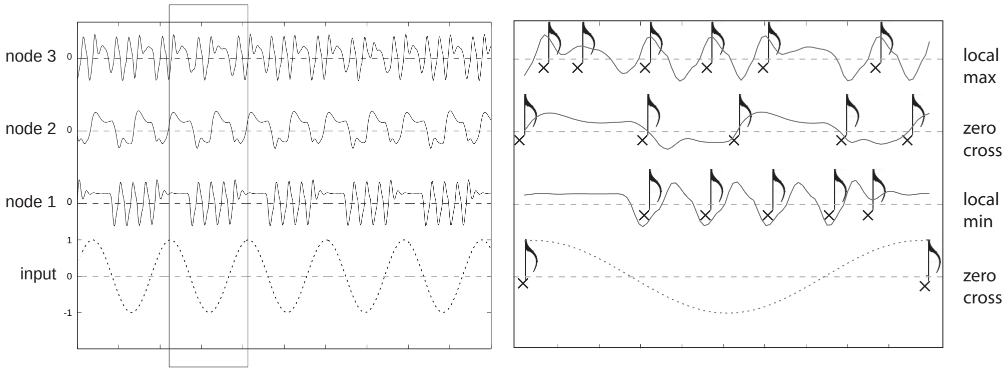


Figure 5.9: Outputs of three nodes in series (left) and detail, showing beats triggered (right): sinusoidal input and node three are triggered at local maxima, node two at zero-crossings (falling and rising) and node one at local minima.

- A more interesting example is given on track 13. There are four nodes in series, weights between each node, and biases are held constant. The different rhythms are produced by changing the time constants of individual nodes. A change is made half way through to one of the connecting weights, demonstrating how the basic beat is preserved, whilst varying the ornamentation and altering the stress.

5.2.4 Summary of Neural Oscillator Features

The basic oscillatory patterns of the nodes mimic the wave-like structures of many melodic and phrasal structures in instrumental music. The continuous nature of the outputs provides scope for mapping to a range of musical domains. Control of individual parts is made possible by altering the gains and time constants, creating variations across components in a network operating at a unified frequency. Altering the weights between the nodes obviously also gives control over the relationship between constituent parts. The same mechanism could be applied in interactive system with a performer - positive weights on the input causing the system to spring into life when the performer plays, negative weights causing their playing to inhibit the system which would only play when they are silent. In artificial neural networks, these weights model basic mechanisms of inhibition and excitation which are fundamental to neuronal communication. Musically this process can be used to mimic basic modes of interaction between musical parts or players i.e. unison or contrary motion

In the examples above, the entrainment property of the neural oscillator is utilised to provide a metrical unity across each rhythmic part. This property also provides an implicit beat detection mechanism that can be used to set the network outputs to a user-defined pulse. A beat interval supplied via a MIDI (or ASCII) keyboard, or analysis of audio signal can be used to set the frequency of the input signal, to which the rest of the network entrains. For certain settings, changes in the input frequency change the shape of the output signal. The result is a system that can keep time with a human player, but will produce novel, unpredictable rhythmic variation.

This simple network, even with hand set parameters can be used to generate intriguingly musical outputs, with connected nodes creating a sense of ensemble. One of the immediate draws backs of this implementation is the incessant nature of the output. The melodic mapping described above was used for the installation *Organised Entry*, presented in Chapter 6, but combined with another system which acted as a mixer, controlling the entries of individual units in the network.

5.3 Pattern Propagation in Cellular Automata

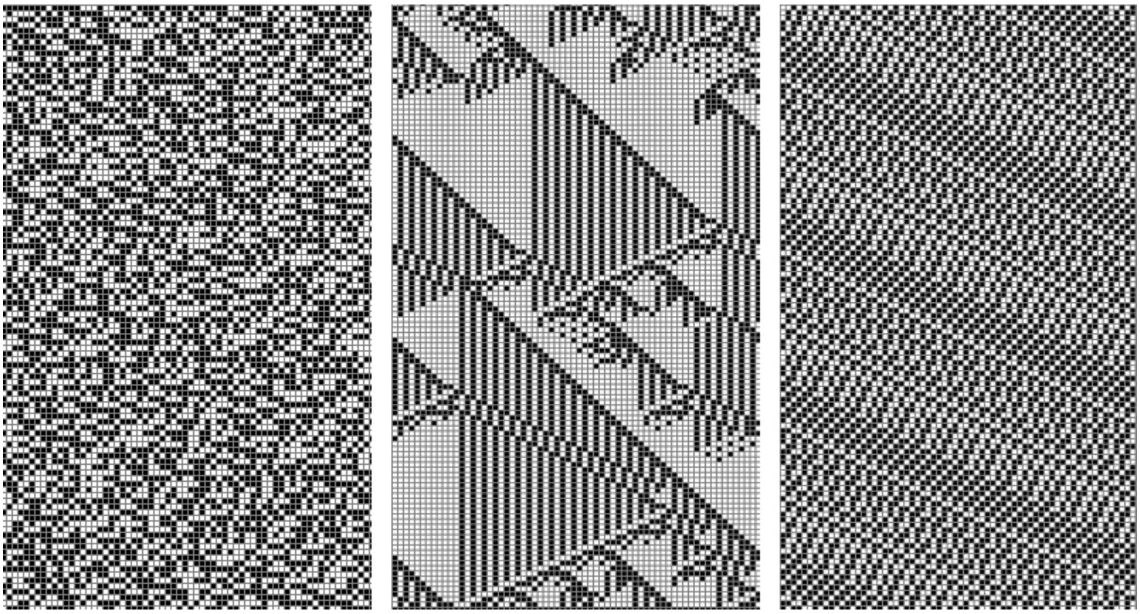


Figure 5.10: Graphical representation of 1D CAs: chaotic (left), complex (middle) and ordered (right).

Cellular automata are amongst the most used Alife system, their pattern propagation properties being an attractive means of generating low level structures. As described elsewhere, they have been used to specify pitch information, as well as to control signal-level parameters for sound synthesis. A basic description of CA is given in Chapter 3, Section 3.1.2. In the current project, the different forms of patterns generated by CAs have been used to generate different rhythmic textures. The main mapping employed is that described in Chapter 4, Section 4.3.4 which creates pitched rhythmic patterns from their output. Further examples are given on tracks 14, 15 and 16 which correspond to the graphical representations of the rules (chaotic, complex and ordered) shown in Figure 5.10.

- Track 14 gives an example of the chaotic rule set shown in Figure 5.10 (left). The random distribution of black and white cells creates an almost continuous spattering rhythmic pattern. Recall that the mean and variance of the frequency distribution of the look up table are used to determine the root pitch, and the size of the intervals of the triad above this pitch respectively. In chaotic rules, each configuration is equally likely, meaning that the frequency distribution of the look up table is flat. This creates small intervals both between iterations and a low variance across each iteration. This is heard as sets of close chords which vary minimally

- Track 15 gives an example of the complex rule set shown Figure 5.10 (middle). Here the localised patterns evident in Figure 5.10 can be heard as broken blocks of regular pulses. The mixture of local areas of order and higher level more complex forms mean that the same rules are used repeatedly for a period, then change. This means that the frequency distribution of the use of rules is skewed in any one iteration, and varies over iterations. These high variances result in the wider chords which can be heard, as well as the larger changes in root note which signifies the start of each line.

- Track 16 gives an example of the ordered rule set shown Figure 5.10 (right). The mapping employed creates short rising phrases from the diagonal stripes with a regular rhythm. As an ordered rule set, the same individual rules are used over and over again, preserving the same set of pitches. Here only two different rules are used on alternate iterations, producing the alternating root note which can be heard.

5.3.1 Summary of CA Features

The discrete patterns formed by 1D CAs provide a mechanism for generating strongly rhythmic patterns. Under the mapping used here, although there is no metre imposed, the patterns propagated produce distinct patterns of stress which punctuate the low level events with structured accentuations. From a compositional perspective, the distinct rule classes provide a means of varying the rhythmic complexity or accessibility. Rather than mapping only the immediate state changes onto musical events, the use of changes in statistical properties of the process can be used to relate different musical dimensions.

A CA was used in conjunction with a homeostatic network in *AdSyMII* described in Chapter 6. CAs are usually seeded and left to run, but it is also possible to interfere with the state flow by changing the state of cells in current influential neighbourhood. This could for example cause an ordered rule to diverge, an interruption from which it may or may not recover. A similar principle was explored in the individual based ecology model described below.

5.4 Ecology Models

The models presented above predominantly have application in generating material. Other models taken directly from, or inspired by population modelling have also been explored as a means of controlling parameters in, or orchestrating, other systems. Two different classes of model were examined, an individual-based model and a set of coupled differential equations.

5.4.1 N-species Lotka-Volterra Model

The Lotka-Volterra model (Lotka (1925), Volterra (1926)) appears in all undergraduate textbooks as the simplest ecology model that describes predator-prey relationships. It consists of two coupled differential equations as shown in Equation 5.7

$$\frac{dF}{dt} = F(a - bS) \text{ and } \frac{dS}{dt} = S(cF - d) \tag{5.7}$$

Where F is the number of prey (rabbits, small fish, flies, etc.) and S is the number of predators (foxes, sharks, spiders etc.), $a =$ reproduction rate of prey, b = predation rate, c = reproduction rate of predators (per prey eaten) and d = death rate of the predator.

For any positive values of a, b, c and d the system oscillates in a limit cycle. In ecological terms this is incredibly over simplistic, as no ecology consists of only two species, but is made up of numerous trophic levels connected in a complex food web. For the current purposes the ecological validity can be ignored and the potential dynamics of the system increased by creating a model for N species.

There are many ways of generalising the basic Lotka-Volterra equation. The one employed here was developed by Arneodo et al. (1980). In contrast to the simple limit cycle exhibited by the two species Lotka-Volterra model, the n-species model used here exhibits a broader range of dynamics for a larger number of species which is readily parameterised. The system of ordinary differential equations for n-species can be re-written as:

$$\frac{dx_i}{dt} = x_i \sum_{j=1}^{n} A_{ij}(1 - x_j) \tag{5.8}$$

where x_i represents the ith species and A_{ij} represents the effect that species j has on species i. The A_{ij} terms can then be represented as a matrix. For three species the values can be defined as:

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0.1 \\ -0.5 & -0.1 & 0.1 \\ \alpha & 0.1 & 0.1 \end{bmatrix} \tag{5.9}$$

where α parameterises the whole system.

As shown in Figure 5.11, the system exhibits a range of dynamics which are controlled by the α-value. Low values ($\alpha \leq 0.75$) cause the system to converge on a fixed point attractor, at higher values simple periodic behaviour emerges. Increasing the value beyond this causes period doublings until at around $\alpha = 1.5$, the system exhibits chaotic dynamics.

.

- Track 17 provides a simple example where each of the three outputs are mapped to the playback speed of three different versions of the same sample. The recordings are of Inuit caribou ladies babbling. At an initial α value of 0.75 the system is converged on a point attractor, and each sample plays back at normal speed. As α is increased to 1.2 you can hear the simple periodic behaviour emerge as uniform oscillations in the playback speed; the period doubling evoked at $\alpha = 1.4$ gives a double loop, and at $\alpha = 1.5$ the chaotic dynamics create chaotic pitch changes.

In itself the behavioural repertoire of the GLV model is perhaps a little limited, but mechanisms like this are a useful addition to the compendium of objects. This model was used as a mixing device in the *Organised Entry* installation described in Chapter 6.

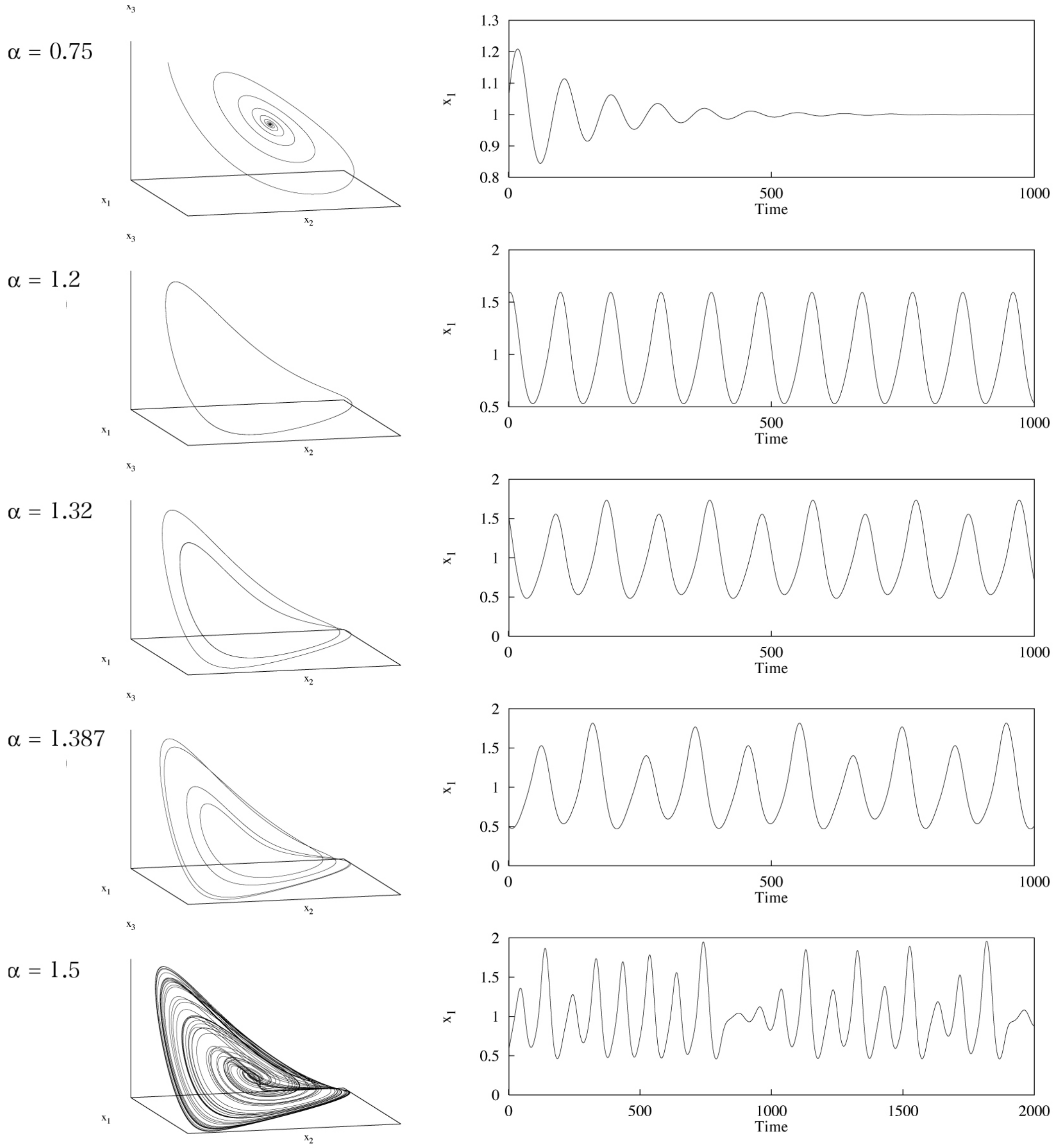


Figure 5.11: Period doublings in a three-species Lotka-Volterra system: phase space on the left and x_i on the right.

5.4.2 A Simple Agent-based Model of Spectral-temporal Organisation

Evolutionary agent-based models are often used in Alife art and music as a means of generating diversity, and exploring more open ended behaviours. This simple model serves to illustrate how systems can be contrived to fulfil specific methods of control and organisation. Specifically, the model aims to provide a mechanism for distributing spectral and temporal features of events into unique niches. In contrast to the traditional application of GAs as a means of achieving a single honed individual, this model aims to achieve specific properties at the population level.

Within music, whether a symphony score, R and B track or improvisation group, the function, value and significance of each part or player only makes sense relative to every other voice. A central aspect of composition is in balancing the various lines such that they each occupy their own unique space. Many composers, particularly of the acousmatic tradition, draw inspiration from the organisation of sound in the natural world. Bio-acoustic studies of natural habitats suggest that each organism occupies its own sonic niche both in frequency/ spectral domain and in time. Results from studies performed in Sequoia national park (Krause (1993)) support this hypothesis [3]. If one creature stops vocalising, another joins the chorus, keeping the bio-spectrum intact. This idea is supported by reports that point to the disruptive effect of human industrial noise on populations of local wildlife. For example the population decline of birds living in areas of motorway development has been attributed to the noise of the traffic preventing communication and therefore mating (Barot (1999)).

Drawing from these observations a simple model was implemented to investigate whether a self-organising mechanism, based on the premise that sound objects could only persist if they occupied a unique spectral/temporal niche, could be used to organise a randomised set of pitch-time values into unique and stable spectro-temporal niches.

The Model

The model is a simplified version of those used in individual-based ecology models (e.g. Epstein (1996), Forrest and Jones (1994)). The system consists of a population of agents which are defined by their pitch (P) and vocalisation time (VT) values. Based on the premise that individuals in any one species can only reproduce if they can hear each others' mating calls, reproduction can only occur between individuals of the same pitch if they share the same VT value AND no other individuals of any other species hold this value.

A population of agents is initialised with pitch and VT values selected from a uniform random distribution over the intervals $[1, 10]$ and $[1, 100]$ respectively, and an energy level. There are currently no spatial dimensions, and no resources. Time is discrete, and each iteration consists of N timeslots, during which individuals vocalise. At each timeslot t_n any agent with vt value n, produces note p. If that timeslot is uniquely occupied by agents with the same pitch value, reproduction occurs. Half the number of agents with coincidental values are produced. Offspring inherit the parental P value which remains fixed. VT values are inherited and mutated, using creep mutation with wrap around, with a probability of 0.1. Energy levels are reduced for all agents on every iteration according to whether or not they reproduced: taxes for those that did not reproduce are twice those that did. When energy levels reach zero, the agent dies.

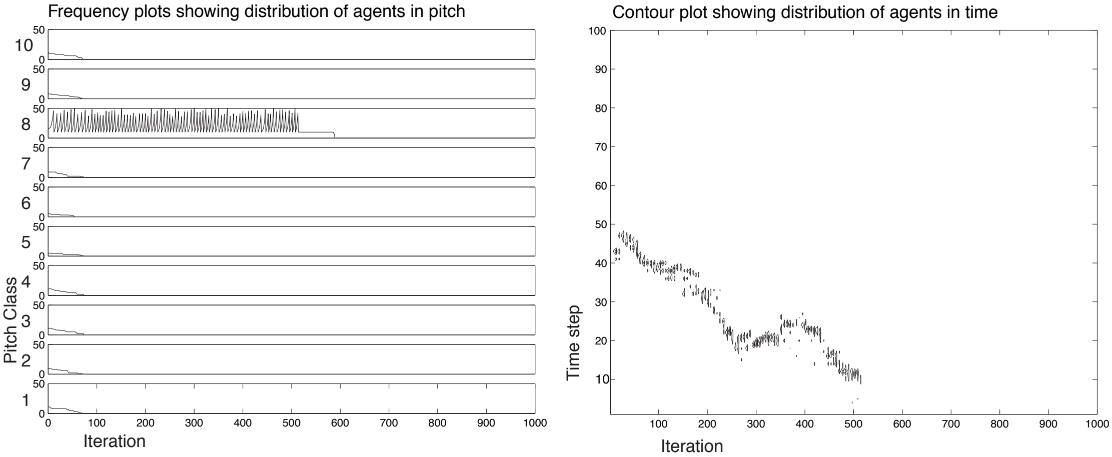
This mechanism alone was sufficient to produce populations which inhabited unique pitch-time spaces, but in the absence of any external resources, additional factors were required to curb the population and introduce novelty. A global population maximum

---

[3] The team suggest that the biophonies of natural habitats can be used as a measure of the health, or stability of an environment: the more clearly demarcated each species is in spectro-temporal map, the more stable the system.

was set. When this is reached, no reproduction can occur until some agents die out. This produces periods of stasis. A maximum is also set for each pitch class. When this is reached, the number of agents of that pitch is reduced to X% of the remaining population by randomly culling individuals.

In the absence of any external resources, or reproductive mutation of pitch values, the system is extremely sensitive to the initial distribution in terms of the number of agents that can reproduce. Once a pitch class dies out, there is no possibility for it to re-enter the population. An extreme example, with only one pitch class is shown in Figure 5.12. Drawing from the observation that in natural acoustic ecologies, when one spectral niche is freed, another organism adapts its call to fill the gap, here when the number of pitch classes (species) drops below a certain threshold, pitch values of the remaining members of the population are mutated with a low probability at each iteration for the remainder of their lives, introducing life-time variability in pitch.


Frequency plots showing distribution of agents in pitch

Contour plot showing distribution of agents in time

Figure 5.12: Figures showing distribution of agents in each pitch class (left) and across time slots (right) for system with no lifetime pitch variation.

The system described above is capable of organising an initial random population into subgroups that occupy unique areas of pitch-time space. Figure 5.13 shows the initial and final distributions of a population in pitch-time space in a typical run. Figure 5.14 shows the movement of the population over 200 iterations in pitch (left) and time (right). Even in this simple model, it seems that the reproduction criterion (in conjunction with the restraint thresholds) is sufficient to produce populations that are stable - in terms of neither dying out nor overcrowding - yet dynamic in producing movement of sub-groups through pitch-time space.

The model also enables external manipulation of the population dynamics. In Figure 5.15, four agents of pitch class five, onset time 80 were introduced are iteration 250. The system was started from the same initial seed as that shown in Figure 5.14, demonstrating the potential for a user to change the course of the evolution of the system.

Summary

The model presented here is extremely simple, and in its current form, the sounds produced are far from interesting musically. However, it suggests that population distributions can be controlled according to simple reproduction restrictions. The reproductive success of each agent is a function of the global environment, which comprises the behaviour of every other agent. This produces a unity between musical output (which is

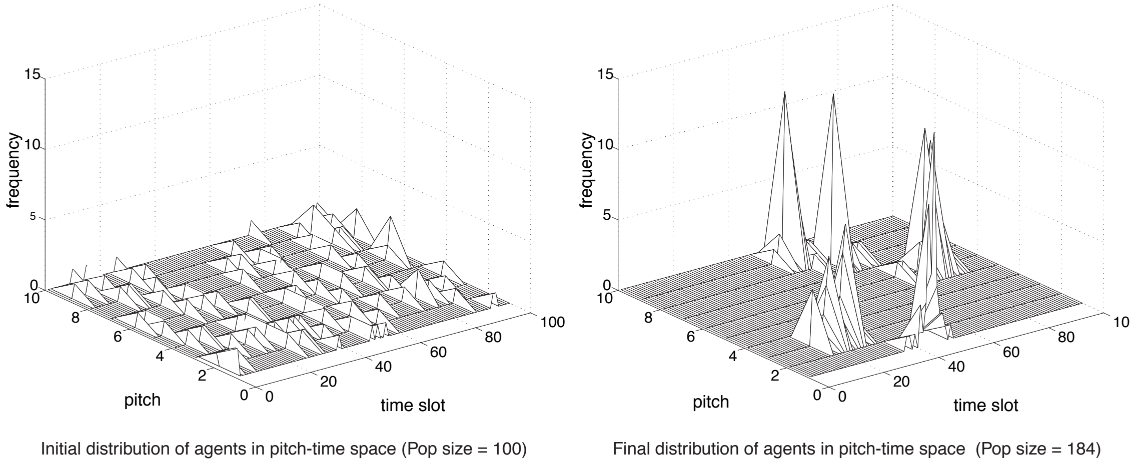


Figure 5.13: Figures showing initial distribution of agents in pitch-time space (left) and final distribution (right).

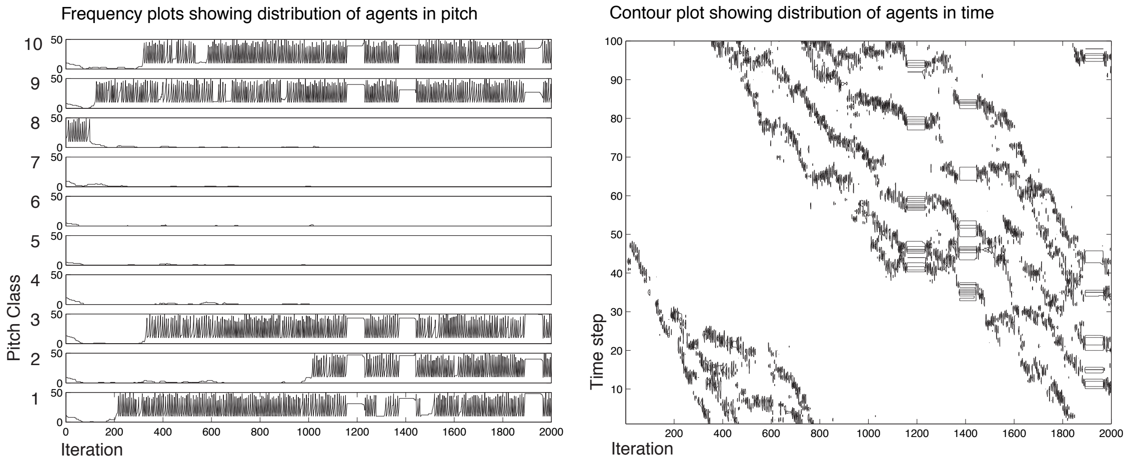


Figure 5.14: Figures showing distribution of agents in each pitch class (left) and across time slots (right) for system with lifetime pitch variation.

the collective behaviors of all elements), and system state. The evaluation of all parts is inherently a dynamic function. This possibility for establishing a coherence *across* a population offers an interesting approach to generative music systems that contrasts with existing evolutionary approaches where the focus is on getting a small subpopulation to achieve a certain criterion, or the pairwise testing of coevolutionary models.

5.5 Implementation

All these models were first developed in C++, and their basic behaviours examined. Where appropriate their response was compared with previous implementations. The homeostat, neural oscillator and GLV equation were then developed as Max/MSP externals so that they could be used within this environment. This makes the exploration of different mappings very swift compared to coding the equivalent DSP or MIDI mappings from scratch. The CA and agent-based model were developed as stand-alone Windows applications.

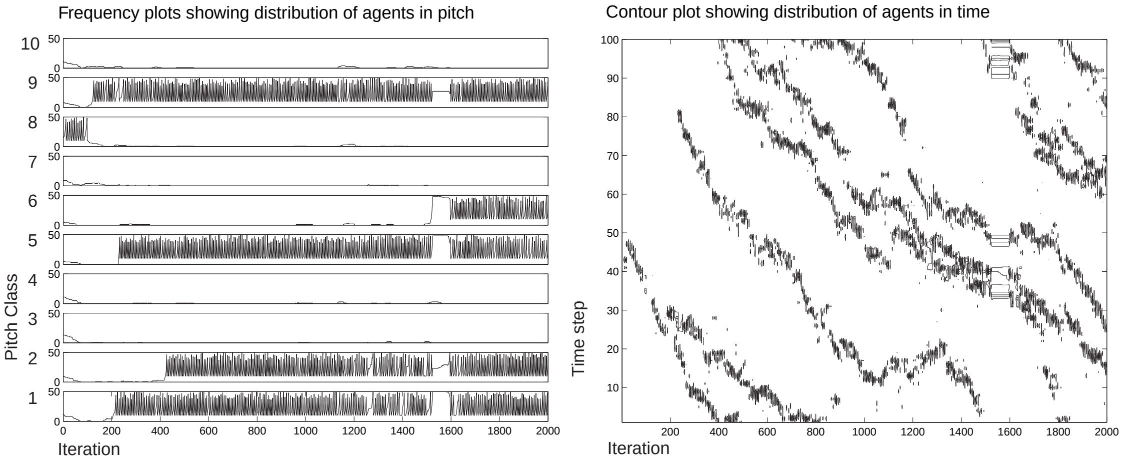


Figure 5.15: Figures showing distribution of agents in each pitch slot (left) and across time slots (right) for system with lifetime pitch variation. At iteration 250, four agents of pitch class five, onset time 80 were introduced

5.6 Discussion

Although very simple, these studies demonstrate the compositional potential of simple adaptive systems in terms of both the sonic effect of their dynamics under different types of mappings, and the practical impact of adopting algorithms that we can only influence rather than directly control.

The homeostat and neural oscillator models in particular both generate a set of evocative behaviours. The interdependencies between the outputs of separate nodes in the networks arguably creates a sense of dynamic cohesion under continuous and multiple mappings in which each part is audibly related and influence is decentralised. Under quantised, rhythmic mappings textures are created that have a strong sense of pulse in the absence of any rigid metrical constraints.

Both systems are parameterised by a handful of variables which allow the user to shape their behaviours whilst retaining the generative independence of the model. This is useful in both compositional and live situations. In addition, both respond to external influence which can be applied manually or algorithmically, evoking a contingency which goes beyond button pressing. The response of the homeostat to perturbation provides an interesting form of control by which we can suggest that 'something' happens, leaving the details of what that 'something' is to the algorithm. When mappings are designed to take this into account, this can create some enthrallingly organic deviations from, and recapitulations to, previous material.

This rhythmic interpretation of the CA takes advantage of its inherent pattern propagation properties, and the use of multiple mappings here, as well as in track 2 of the homeostat demonstrates how single models can be used to generate sets of independent but related musical lines.

Many of the mappings used here act to quantise the continuous outputs of some of the algorithms. For example in using the NOSC outputs to trigger MIDI notes, much information is being thrown away. In some respects, the true, continuous, dynamic nature of the models is only preserved under mappings such as that used for the Lotka-Volterra system. This is an example of the *control* mapping outlined in Figure 5.1.d – in this case the algorithm's three outputs were used to continuously alter the playback speed of three versions of the same sample. These sorts of mappings are perhaps most typical in Sonic

Arts domain where algorithmic composition is popular. The use of neural models to generate arpeggiated forms typical of classical or early electronic music may seem to be a mixing of worlds, but it is precisely this synthesis of difference that characterises the cyber-nature aesthetic of Alife visual art, a synthesis which I am interested to evoke in the sonic domain.

From a systemic perspective, perhaps the most interesting mappings are those that go beyond a simple one-way number-to-note formula and feedback into the system. This was demonstrated in a very simple case in Section 5.1.3. In this case the algorithm is controlling a filter process operating on an existing sample, but the sample is itself affecting the homeostat. This stitching-together of algorithm and implementation is a promising direction for a more collaborative approach to interactive and generative composition and performance and will be pursued a little further in Chapter 8.

The mappings used here have been developed for illustrative purposes. The main thrust of this thesis is to lay the ground work for a more collaborative form of man-machine musicianship, a collaboration in both systemic terms – such that human and algorithm are mutually influential – and a collaboration in aesthetic terms – such that the vagaries of algorithmic composition play out alongside the established acoustic traditions. The implementation of the algorithms in the form of software objects that do *not* impose any restrictions on the way in which they are mapped is also quite intentional. This is in line with the modular approach central to the musical communities that are developing around software such as Max/MSP. In developing algorithms in this way, it is hoped that other musicians can adopt these context-free algorithms for their own compositional ends.

Chapter 6

Composing Generative Systems

Many of the models described in the last chapter have been brought together within larger generative systems in two sound installations. The composition of different models which exhibit different characteristics and operate at different time scales has not been extensively explored previously. Whilst it would be possible to contrive a model that was capable of generating dynamics across a range of timescales, this is no trivial task. The modular approach adopted here is much more tractable and therefore attractive for the generative composer.

As may be evident by now, the aesthetic intention of these systems is to create music that lies at the edge of digital generative and traditional/natural forms, analogous if you like to the 'cyber-natures' of Alife installation art which juxtapose the real and the artificial, and create an illusion of biological convention within silicon graphics. These two installations explore both the use of composite systems and the deployment of multiple mappings from one system into separate musical lines.

- Audio examples of the outputs from these systems are given on the accompanying DVD tracks 18 and 19.

6.1 AdSyMII

AdSyMII was first played publicly at one of the inaugural Blip events in Brighton, as well as at the Generative Art 2002 meeting in Milan. The mappings used in this system aimed to explicitly reference elements of traditional musics: algorithms were selected for their potential to generate 'harmony-like', 'melody-like' and rhythmic forms. Underlying the adoption of these coupled dynamical systems as compositional tools is an implicit assumption that the logical coherence of the models can appreciated in audio.

By combining several models with different characteristics operating at different temporal resolutions, the hope is that structures and coherences will develop that go beyond the initially-interesting but ultimately-insipid streams of musical events that any one model generates alone, and begin to evoke perceptions of higher level form.

The design and implementation of this system aimed to analogise several fundamental music characteristics. A homeostatic network was combined with a CA and a simple stochastic algorithm for generating note lengths. The homeostat outputs were used to create microtonal harmonies, a (loose) analogy being drawn between the vertical and horizontal structures of traditional harmony and the result of the primary feedback mechanism of the homeostat which orchestrates discrete units in relation to each other over time. Direct mappings were used to make the dynamics of the homeostat evident and create sonically effective contrasts between periods of stability and exploration.

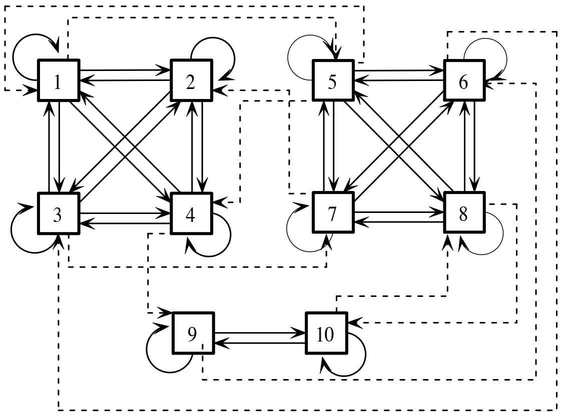


Figure 6.1: Diagram of homeostat network showing full intraconnections and 10% interconnectivity

The pattern propagation properties of a CA are used to generate temporal patterns which aim to capture a basic rhythmic feel. The pitch of these are also determined by the homeostatic outputs, as are the pitch values of events in a single monophonic 'melody' line. These multiple mappings were implemented in an attempt to achieve a higher degree of harmonic coherence to counterbalance the unfamiliar microtonal space.

- Track 18 gives example output of the system.

6.1.1 A Self-regulating Homeostatic Network

The model of the Ashbian homeostat, as described in Chapter 5 exhibits some useful properties but as here it was to be used in a standalone generative system, designed to run over a stretch of time, the model was adapted to provide dynamic structure on a longer timescale. In order to do so, two modifications were made. Firstly the network size and connectivity were altered. Secondly, several possibilities for implementing an internal feedback for self-regulation of key parameters were explored.

In Chapter 5 it was noted that the stability of the network was an inverse function of both size and connectivity. This is a well recognised property of networks in general, as studied for example in graph theory. Equally recognised is the fact that larger networks can be made more stable if they contain sub-networks which are densely intra- connected, but sparsely inter-connected. This potentially allows the development of more complex dynamics. A network diagram is given in Figure 6.1. The network can be conceived as two four-unit homeostats and a two-unit homeostat[1] Each are fully *intra*connected but the *inter*connections can be varied randomly according to a user specified percentage connectivity.

Various methods of network self-regulation were explored. The two most influential parameters that were fixed (by hand or circumstance) in Ashby's model were the critical deviation (which determines the point at which the network weights are reconfigured) and viscosity (which determines the stability of the system). Various methods of direct and proportional control of these parameters by the outputs of certain units were investigated.

Initially, the two parameters were simply set to the output values of two different units. This was rejected as it meant that as soon as viscosity reached a high value, the

---

[1] The number of parts and their sizes were selected arbitrarily: the effects of these variables and their ratios deserves investigation.

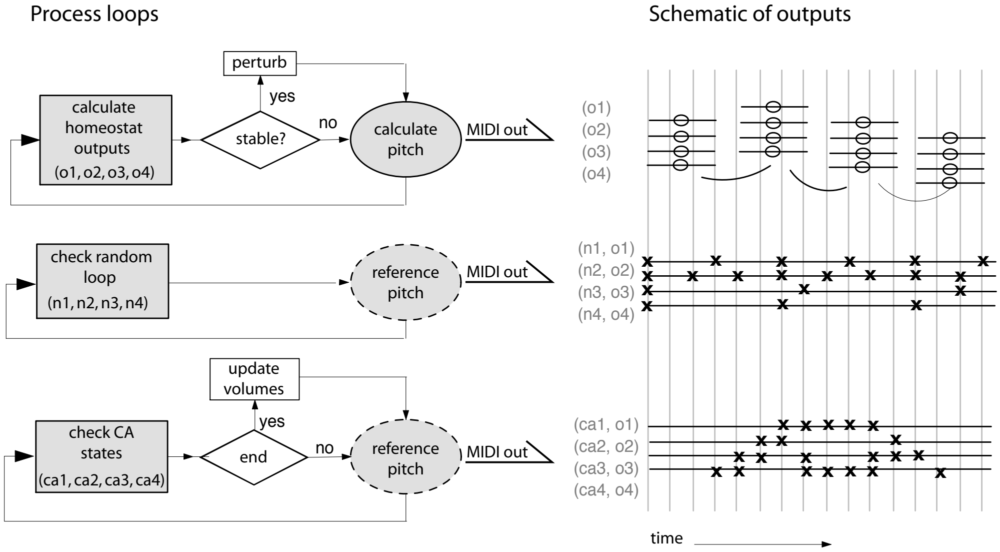


Figure 6.2: Schematic of the three process loops in AdSyMII and outputs produced.

system froze, with many units stuck on this value. Aesthetically, this produced uninteresting behaviour. More seriously if the unit controlling the value of the critical deviation was stuck on the same or lower value, all other units at or above this value would be out of limits. Thus continual weight changes occur, but the system cannot move to self-correct. Homeostasis cannot occur.

After further experimentation, self-regulation of the critical deviation was rejected, and proportional control of viscosity was implemented using Equation 6.1. The equation was derived to invoke a relationship, in discrete time, analogous to the differential equation governing the damping effect of viscosity on a moving body.

$$\frac{1}{viscosity_{(t+1)}} = \frac{O_{i(t)} + CV}{2} \tag{6.1}$$

where O_{it} is the output from unit i at time t and CV is the critical deviation.

To ensure that the essential qualities were preserved in the modified network, resistance to perturbation and effect of connectivity on stability were examined. The results of these are given in the Appendix A, Figures A.2 and A.3.

6.1.2 System overview

Musically, the system produces three separate parts, voiced on different instruments. These can be conceived as a sustained four part harmony, a pitched rhythmic part, and a melodic part. The pitch values of all parts are derived from the outputs of a self-regulating homeostatic network, the timings of the rhythmic part are defined according to the states of a 1D CA, and the note placements of the melodic part are determined using a simple stochastic method, all described below. A schematic of the three processes involved and their outputs is given in Figure 6.2 and a transcription of a few seconds of the output is given in Figure 6.5.

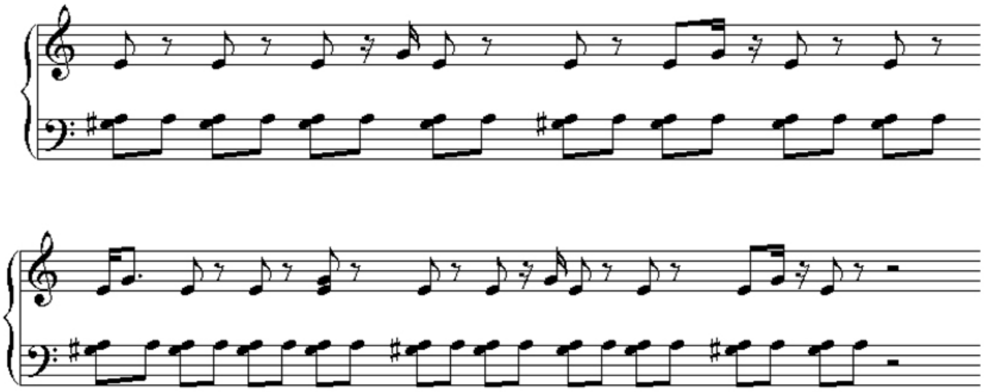



Figure 6.3: Transcript of the simple random loop. In this example the notes G♯, E, A and G are voiced every 4,4,2 and 11 semiquavers respectively. In the actual system these pitches are defined by specific homeostat units.

At each iteration, four pitch values are generated from a subnet of the homeostat, and the rhythmic and melodic parts voiced at these pitches according to the CA states and a set of randomly selected numbers respectively. A fixed CA is used and the random values remain constant. Each part therefore repeats the same rhythmic loop with variation in pitch. When the homeostatic network stabilises (i.e. all outputs converge to a point or limit cycle) it is perturbed, pushing the system onto a new path, and creating new harmonies.

6.1.3 AdSyMII System Components

Stochastic melody line

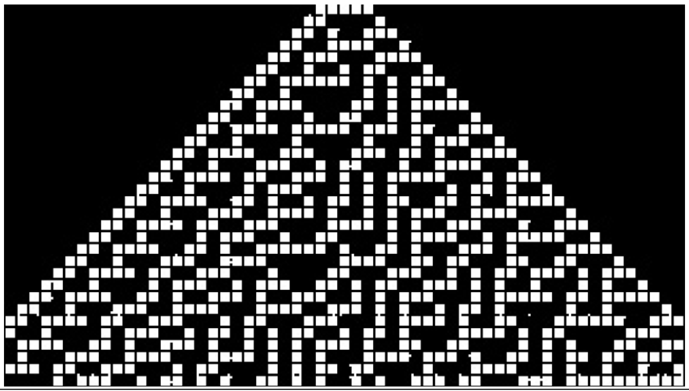
A simple stochastic method was used to determine the timings of the 'melody' line. N numbers $n_0, n_1 ... n_n$ are randomly selected. N defines the 'density', and can be any number, but somewhere between 3 and 10 works well. The value of n determines the frequency of each note i.e. it will be played every n beats. For example if three numbers 3, 5 and 9 are selected, notes will be played on beats 3, 5, 6, 9, 10, 12, 15 etc. The period of the whole loop is therefore the lowest common multiple of the entire set. In this case the loop will repeat every 45 beats. Where one value of n is a multiple of another, or a common multiple is shared by a subset, an accent is created, or if different pitch values are used, a chord.

A transcription of a simple example is shown in Figure 6.3. In this example, the selected numbers are 4, 4, 2, and 11. These are played on the notes G♯, E, A and G respectively with the semi-quaver as the smallest unit of time. The 2 and 4 beat notes are shown as quavers for clarity.

Rather than being fixed as in the example shown here, the pitch of each note is determined by one of four homeostat outputs. When the network is stable and all units are converged on single values, the melody is made up of just four notes. During unstable or oscillatory periods however, more complex melodies are created, as the period of oscillation rarely coincides with value of the random number with which it is associated.

CA rhythms

The rule set and a graphical representation of the CA used are given in Figure 6.4. The production rules are used to generate the CA pattern in a 13 x 22 grid. This is done in order to preserve the changes in rhythmic density (arising from the triangular shape). The states of individual cells in the CA are read left to right and interpreted very simply



neighbourhood states (t):	111	110	101	100	011	010	001	000
cell state ($t+1$):	0	1	1	0	1	1	1	0

Figure 6.4: Graphical depiction of the 1D CA used (above) and rule set (below)

as a rhythmic score: on = play, off = rest with four lines being concurrently voiced, played at pitches designated by the homeostat outputs. This mapping is illustrated in Chapter 4, Figures 4.4 and 4.5.

Homeostat network and automations

The self-regulating homeostatic network described in Section 6.1.1 lies at the centre of the system, defining the pitch of the sustained harmonies, melody line and CA rhythms. The stability of the network is monitored and when stable, a large input is automatically applied to unit one, triggering weight changes, and a new set of dynamics. Outputs of this network are used to control not only the pitch of the four chords, but the volumes of the melody and harmony parts, the transposition of the chord, and the value of viscosity (as described above). Whenever the system stabilises, it is perturbed after a short period. The time step between iterations of the whole system is also automatically controlled, changing the pace of the output.

The outputs of units one to four are scaled and mapped to MIDI pitch bend to create microtonal harmonies of $\frac{1}{32}^{th}$ tone. For the melodic and rhythmic parts, these are used as absolute values. The sustained harmonies are additionally transposed at the end of each iteration of the CA grid.

The amplitudes of the melody and rhythm parts are determined by the outputs of units nine and ten respectively. Unit eight controls the shift in root value of the chord (transposition). Unit seven controls the value of maxchange as described in 6.1.1). These updates are made at the repetition of the rhythmic loop. At the end of the CA loop, the output values of units nine and ten are scaled to control the volume settings for the instrumental parts. The change in root values affect only the sustained chords and not the melody and rhythm parts.

A gradual accelerando is implemented by reducing the time between iterations logarithmically (i.e. reducing the time step by one each iteration) to a minimum value of 180ms. This reset to its original value of 300ms when the homeostat stabilises. When stability is achieved, a critical perturbation of the output of unit one is induced, causing randomisation of connections and creating a new harmonic path. Using the outputs of the units to control the musical parameters effectively normalises their distribution: outputs are most frequently within the limits of the critical deviation, but occasionally lie at extremes. Thus extreme behaviour in the system is reflected in the music produced.

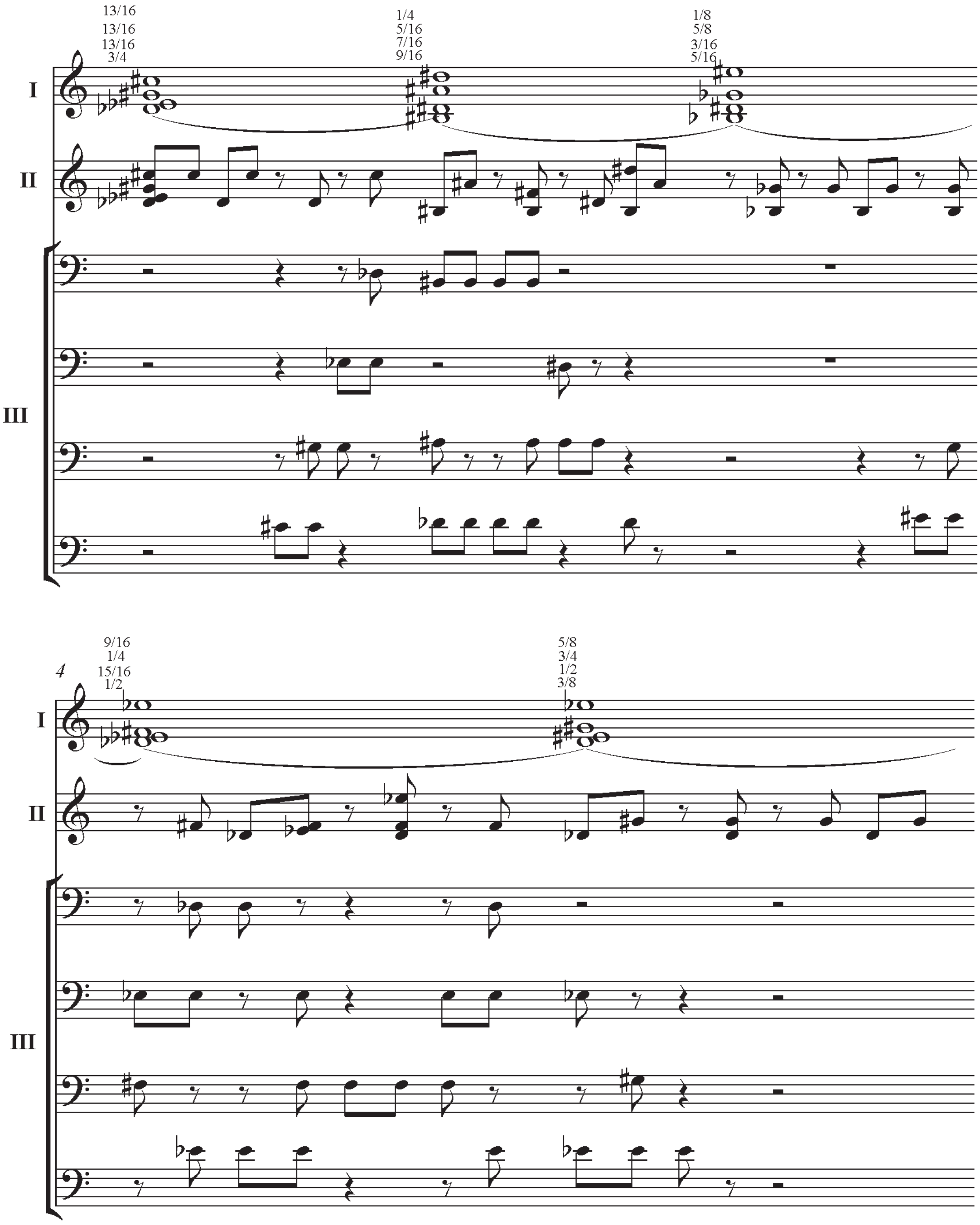



Figure 6.5: Transcription of the three separate lines generated by the system. The top lines shows the sustained microtonal chords generated directly by the outputs of four units the homeostat network. The second line show the 'melody' line picked out of these chords at fixed time intervals. The bottom four staves show the rhythm part generated from four lines of the CA shown in Figure 6.4

6.1.4 AdSyMII Reactions and Discussion

AdSyMII was very well received in two different contexts. It was first shown at a forum for sci-art where it was played for several hours in a bar. The setting provided an ideal stage for the system, as listeners could experience the large scale structure of the system in terms of its progression through various states over an extended time.

All those attending were 'very impressed' and commented that the system was 'exceptionally musical' compared to other, even knowledge based, generative systems. Generative artist Paul Brown, the evenings' main speaker, showed similar enthsiasm, as did a representative from a renowned interactive multimedia company 'audiorom.com', who expressed an interest in commissioning the work for an installation project. John Petigrew, one of the founding members of SSeyo also expressed an interest in the system for use on his planned 'evolved art' cable channel. In other private presentations, comments such as, 'ooh is that you on the cello?', were made by naive listeners, who were all surprised to learn that what they were listening to was entirely algorithmically generated.

In a more rigourous setting, twenty volunteers, all unaware of the nature of the compositional process, listened to an example of the output of the system. There was no imposed listening time, but on average each listened for just under fifteen minutes. Each filled in a survey consisting of scaled and open questions (details can be found in Eldridge (2002)). All agreed that the audio produced was "interesting", "musical", and "would be described as music". 95% of listeners agreed that the audio example they had listened to bore "qualities that they normally associated with music", elaborating their statement with common descriptions of musical forms and structures: "*sense of melody ... driving sense of rhythm*", "*there were definite harmonies if unusual at times*", "*sense of harmonic and rhythmic structure and melodic progression*". Reference to structure was made by several listeners: "*structure and development on different timescales/ resolutions*" , "*certainly if not composed by a person it must have been restricted in scale, structure etc*". Many listeners also made comments pertaining to the emotive qualities of the output: '*tension building and resolution of tension*', '*It had the ability to generate mood ...*', '*It oozed atmosphere ... of a crazy graveyard in Iceland*'. Other comments suggested it succeeded in achieving a 'cybernatural' balance: "*weird and surprising yet strangely familiar*". Details are given in Eldridge (2002)

Reports of perceived structure are encouraging in suggesting that the bottom up approach can be successfully applied to achieve definite musical aims. The main algorithms employed controlled musical events at the lowest level - simply defining positions in pitch-time space, however listeners consistently described their perception of harmonic and/or rhythmic structure. It is not clear *which* factors promote the perception of sounds in time as music. In this instance, it could be simply the familiarity of the timbre of the MIDI instruments with which the lines are voiced, or the somewhat arbitrary presence of dynamic and tempo changes. However, listener responses suggest that any musical success may be attributable to the internal *structures* of the music which reflect the dynamics of the algorithmic processes.

The basic stable-runaway-stable (S-R-S) pattern characteristic of the simple homeostat when perturbed produces a basic balance of repetition and novelty. (A similar balance is present in the CA rhythm and stochastic melody, as the timings are repeated, but played at changing pitches). In the homeostatic network used here, the S-R-S pattern of the individual unit is also manifest at a higher level in the overall cycle through stability, oscillation and runaway behaviour. It seems possible that the higher-level dynamics of the network provide an internal structure that promotes 'musicalness', perhaps by engaging analytic processes in the listener, despite the absence of traditional musical conventions of form.

6.2 Organised Entry

Organised Entry is a generative sound installation developed for The BigBlip 05, a festival of generative art. The piece was designed to emanate from below a set of metal stairs that led down into the basement exhibition space. Whereas AdSyMII used several different algorithms to create separate parts, this system is based on one principal generative engine – a neural oscillator network – the outputs of which are mapped in a number of different ways. The Lotka-Volterra equations are then used to control the amplitude of each layer, effectively acting as a mixer. The piece moves away from the parody of note-based music of AdSyMII into the world of more atmospheric Sonic Art and aims to evoke a (comical) sense of frustration – of there being something trapped under the stairs.

- Track 19 gives an example of the output.

6.2.1 Organised Entry: System Overview

Organised Entry uses a neural oscillator network driven by a sine oscillator and the Lotka-Volterra models described in Chapter 5. All sonic outputs were created by mapping the outputs of an eight unit neural oscillator network in different ways to different sets of samples and MIDI instruments. The Lotka-Volterra model is used to control the amplitude of these layers and also the amplitude of the driving oscillator.

As shown in Figure 6.6, the neural oscillator network was used to create three principle parts. Firstly the continuous outputs of each node in the network were quantised and used to generate arpeggiated lines realised on a MIDI control synthesiser. A 'bass' line was also created by sustaining the minima of nodes. A second layer was created by triggering pre-recorded rhythmic samples at local minima or maxima where playback speed was determined by the value of the corresponding node. Finally a set of textural samples that played continuously were modulated according to the output values of the network. These layers were then controlled by a three species Lotka-Volterra system, the three outputs of the model being used to determine the amplitudes of one of the three sonic outputs.

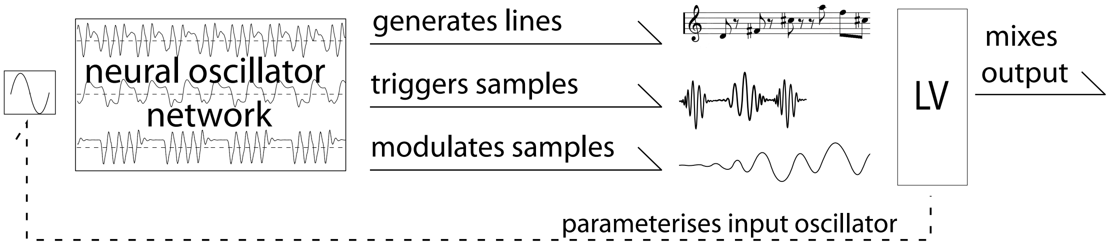


Figure 6.6: Schematic of the organisation of models and mappings in Organised Entry.

6.2.2 Organised Entry: System Components

The basic engine consists of a network of eight neural oscillator nodes each with two inputs. Nodes one and five receive input from a single sine oscilator, providing global control of the principle period of oscillation. Weights between nodes, gain and time constants were hand-set and updated randomly within fixed ranges according to the output of the Lotka-Volterra model which was described in Section 5.4.1. The outputs of each of these eight nodes were used to generate three different types of musical material.

Melodic lines were generated using the mapping described in Section 5.2.3, creating eight different lines voiced on eight different MIDI instruments. This is shown graphically in Figure 5.8. Between two and eight of these were heard at anyone phase, the lines being randomly selected. In addition, a sustained melodic bass line was created by voicing just the minima of each node, this was not controlled by the Lotka-Volterra outputs, but remained throughout. The oscillator outputs were scaled to cover a broad pitch range (ranging from MIDI note value 12 - 136). In addition, each node has an associated cut off value, below which it does not play. This provides a means of hand-tuning the density and pitch range of individual nodes.

Rhythmic samples recorded from an mbira (thumbpiano) were triggered and pitched using outputs from the same network. Each sample was approximately three seconds in length. Using a mapping similar to the rhythm generation approach described in Section 5.2.3, four different samples were controlled by the outputs of nodes one to four. Samples one and two were triggered on local minima, and samples three and four on local maxima. In each case the absolute value of the associated node was mapped to sample playback speed, altering the pitch of the material according to network outputs, this was implemented such that pitch changes only occurred at the start of each playback. Output values were scaled such that playback speed varied by approximately +/- 10%. The range was limited so that the samples could always be recognised.

Textural samples, which were mostly recordings of bowed and struck metal were also triggered at local minima in network outputs. These were all longer in length, and modulated in pitch and amplitude during playback by their associated network outputs. The variation in playback speed of these was much larger, creating a wide variety of textures according to the behaviour of the network.

These three sets of mappings provided a great deal of musical material, all stemming from the same network, and so sharing essential structural and dynamic forms, but realised in quite distinct sonic classes. These were not played out constantly in a continuous cacophony, but mixed using the outputs of the three-species Lotka-Volterra model that was described in Section 5.4.1.

Lotka-Volterra equation as system mixer

A three species predator-prey model was used to orchestrate the three separate classes of material. Cut off points were defined for each output which determined when a class was heard. At values below this, the part was tacet, above threshold, the value was mapped to amplitude or volume of the respective class. This is shown in Figure 6.7.

The third output was also used to determine the amplitude of the main sine oscillator input to the network, switching the amplitude between zero and one as it droped below or above a threshold value respectively. Introducing an input causes the nodes to entrain the frequency of the input signal, altering the shape of each output oscillation and so changing the melodic lines that are produced. Thresholds on the Lotka-Volterra outputs were set such that all three parts rarely played in unison, and periods of silence were limited. On the occasions that all outputs *did* fall below threshold, time constants, weights and biases for the network were updated by selecting randomly form a table of preset values.

6.2.3 Organised Entry: Reactions and Discussion

The system succeeded in achieving a sense of life and local direction, several people (most notably children) were convinced that ‘there was something under there’ as they walked down the stairs. The use of multiple mappings from one system onto a number of different sound sources seems to be an effective way of achieving a sense of coherency and contingency amongst parts.

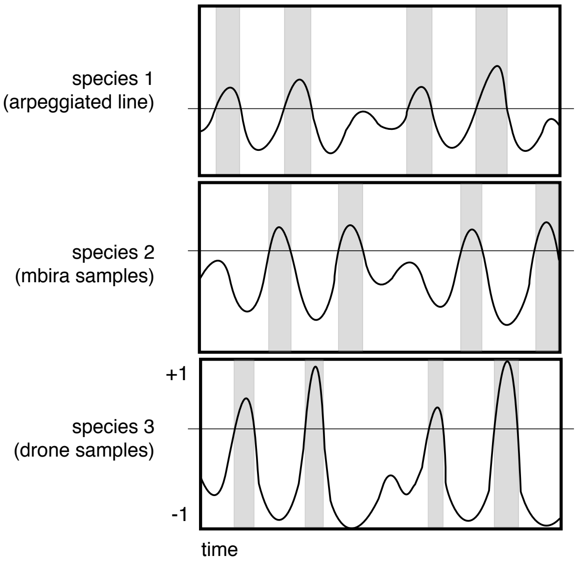


Figure 6.7: Three-species Lotk-Volterra equation as mixer. The horizontal lines represent the hand-set cut-off values. As the population density of each species exceeds the cut-off, the species-density is used to specify the amplitude of the associated part. Shaded areas signify the times when each part is audible.

The Lotka-Volterra model acted fairly successfully as a mixer, bringing parts in and out in a seemingly purposeful way and creating some range of textural variation. In this case, the α value was set at around 1.4, such that there were irregularities in the periods over which parts came in and out, but any very short bursts which would have occurred in the chaotic region were avoided.

In this case, the parameters of the neural oscillator network were varied randomly amongst a hand-selected set. Although the arpeggiated parts seemed to be going 'somewhere', it is a pretty mindless meandering toward a no-where-in-particular destination, sounding like a woodlouse stuck in a corner than an entity with any pertinent goal. An obvious next step if this model is to be used is to consider ways in which parameters could be evolved. Of course this requires some form of fitness function, which might be hard to specify for particular melodic lines, but might be a feasible means of developing specific relational attributes – e.g. setting parameters such that unit two moves in phase with unit one and unit three operates at half the natural frequency with two local minima etc.

In public settings, full advantage of the real-time generation of these systems could be taken advantage of such that events are not only contingent upon internal happenings, but the actions and movements of people in the space. Organised Entry was originally designed to receive an input from the movement of people down the stairs such that the pace of the system reflected (or reversed) the level of activity in the entrance. Due to health and safety requirements, this was sadly not possible, however in simulated tests, changing the input signal produced an interesting range of sonic variation.

6.3 Summary

The combination of multiple models in composite generative systems provides a practical means of generating outputs with a good degree of variation both synchronously and diachronically. In theory it would be possible to design single models that are capable of exhibiting similarly diverse dynamics at multiple timescales, but this is no mean feat. Moreover, as the complexity of the model increases, so to our influence over it invariably diminishes. This modular approach offers a more practical and musically intuitive approach to generative composition.

There are many ways to consider the success of explorations such as these. The individual models can be considered separately as they were in the last chapter. We could question whether one particular set of mappings was more suitable than another. The way in which the basic models are integrated into a larger system is also undoubtedly of importance. However as all of these things necessarily interact and influence each other, consideration of one aspect in isolation is not terribly meaningful.

In terms of diversity and energy, AdSyMII was perhaps more effective than Organised Entry, despite being voiced purely on GM instruments. In this case the mappings were very simple. At the note level a fairly incessant beat is created, but this simplicity meant that the longer term changes could be clearly heard. The dynamics of the self-regulated system varied substantially over twenty minutes or so.

In contrast, the neural oscillator network mapped to MIDI notes to create arpeggiated lines is arguably more immediately musical than the oscillations of the homeostat in the short term. The other samples used created a potentially much richer sound world than the GM instruments to which the homeostat and CA were mapped. And yet after just a few minutes, Organised Entry does not have much more to say.

In Chapter 4 it was noted that mapping is often seen to be equally important as algorithm design. I also suggested that one of the most important things to think about when designing mapping schemes is that the key characteristics of the model are preserved. In the case of AdSyMII, the basic model is run very slowly, so mapping to musical events at the note level works effectively, allowing the longer term dynamics to be heard over minutes of actual time. Organised Entry is run at a similar pace. The slow modulations of samples and note-based events do not contradict the dynamics of the model. In this case however, I suggest that the mapping is doing all the work it can, and that the system itself lacks any long term dynamics.

By hand-setting the parameters of the neural oscillator it is possible to find some vaguely interesting forms in the short term, but difficult to achieve anything on a longer time scale. Setting up the homeostat to be self-regulating seemed to achieve more interesting dynamics across timescales. The modular approach of composing multiple models, and the implementation of multiple sets of mapping from one output to multiple audio events seem promising. However both inevitably rely on suitably interesting dynamics in the underlying model.

Chapter 7

Ashby's Grandmother's Footsteps: an Interaction Installation

> " In the environment, the participant is confronted with a completely new kind of experience. He is stripped of his informed expectations and forced to deal with the moment in its own terms. He is actively involved, discovering that his limbs have been given new meaning and that he can express himself in new ways. He does not simply admire the work of the artist; he shares in its creation." - Krueger (1976) p.84

This chapter opens up the closed generative networks of the last chapter to the real world, experimenting with the use of simple adaptive systems as a means of mediating responsive environments in an interactive and generative sonic game.

Figure 7.1: Visitors treading carefully in Ashby's Grandmother's Footsteps at Artpool, Budapest, 2006.

Ashby's Grandmother's Footsteps is a play on the children's game *Grandmother's Footsteps* where one person stands at the end of the room facing a wall and all the others have to creep up on them. At any point (for instance if they hear movement), grandmother can turn around, whereupon any child caught moving has to return to the far wall and start again from there. This is a cybernetic version, where grandmother is a homeostat, receiving information about people's movement via video analysis and commanding control sonically.

The piece was installed at the Artpool[1] gallery, Budapest, as part of the 'Process Revealed' exhibition in conjunction with the EvoMusArt workshop at EuroGP 2006. Approaching the back of the underground room, visitors to the exhibition hear a faint drone emanating from a pair of wireless headphones hanging on a nail on the whitewashed wall. Fortunately people love being allowed to pick things up in galleries, so this is enough to get them interested. They put them on, and turn to face a previously concealed corridor - Alice-in -Wonderland what's-down-the-rabbit-hole curiosity is again enough to make them approach it. Stepping past an invisible boundary, they suddenly *hear* themselves walking as if on amplified delicate gravel. Walking the corridor, every movement is heard. As they progress down the corridor, the monophonic drone builds and differentiates, developing strange harmonies. A bold step forward triggers a SCREECH - loud enough to halt them for a moment and aversive enough to want to avoid it. Now they must creep slowly to the end of the corridor and turn off grandmother's eyes.

7.1 Power and Play in Responsive Environments

Since the early days of digital, and even electronic arts, people have explored ways of engineering situations in which the audience themselves 'bring forth' a work of art via their interaction with a space, engaging them in a process of co-creation. Often these pieces employ 'invisible' interfaces such as video cameras and sensors to create responsive environments: physical spaces with no obvious exposed technology that respond to visitors movements, gestures or vocalisations. These 'natural' interfaces, bolster a suspension of disbelief in the audience, arguably making them more open to the unlikely events that unfold within.

One of the many devices that gets played out in these environments, is a tipping of the power balance between user and system. Turning the push-button reactivity model on its head, many artists have found ways of engaging audiences in a game where they become manipulated by the system. This may be taken as a social comment of the cultural effects of technology by some, but it also presents a pertinent model of interaction for man-machine performance in demonstrating ways in which digital systems can be made to take the lead and actively induce a response from the human user. This section looks at some historical examples of power and play in responsive environments, and considers how auditory feedback can be given to maximise the audience's engagement.

7.1.1 Responsive Environments

The diffuse and intuitive interfaces of much contemporary European interactive art build on ideas laid down by players in the Art and Technology movement of the late 1960s. As early as 1969, Myron Krueger worked on multi-media 'responsive environments', such as *GlowFlow* (1969) which combined pressure sensitive floor pads with basic surround sound and reactive light elements. Phosphoresent particles were pumped through tubes attached to the walls of a darkened room in such a way as to distort the visitor's perception of the room's shape. The glowing particles and sounds were triggered by users standing on the sensors placed through the room. Others such as Seawright and Rauschenberg explored similar environments (Dinkla (1994)), but it was Krueger who developed the technology to create a more complex dialogue between user and environment.

---

[1] `http://www.artpool.hu/`

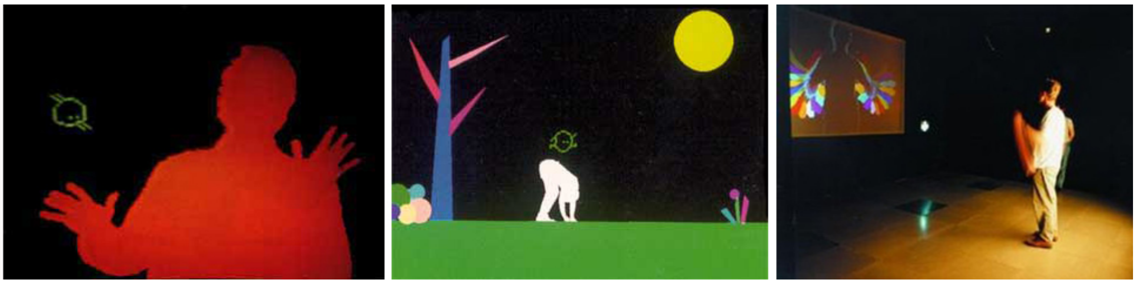


Figure 7.2: Artificial Reality space of Myron Krueger's *Critter* from *Videoplace*.

Krueger was concerned with creating what he described as 'artifical realities', aiming for a full-body participation in computer-mediated events that were so compelling that they would be accepted as real experience. In contrast to the head-mounted displays of nascent VR, or touch-screen/ mouse interfaces popular in the USA at the time, Krueger adopted the closed circuit video technologies that were fashionable amongst his contemporary video artists. The real time projections were combined with responsive, generative software to create these graphical 'artificial realities'. The viewer would enter a space dominated by a projected image which combined their own self-reflection within a world inhabited by computer generated images which were designed to respond to, and often evoke responses in, the participant. Figure 7.2 (right) shows one such projection where a simple delay on the video creates after images of their movements.

Beside his technical innovations, Krueger pioneered the notion of the artist as a 'composer' of intelligent, real-time computer-mediated spaces. Works such as *Videoplace* are presented as sets of composed interactive experiences, in which different forms of gestural analysis interpret, or even anticipate user's actions. In perhaps the most famous piece in the Videoplace collection, *Critter*, the user is taunted by a cartoon like creature. Shown in Figure 7.2 (left and middle). As someone enters the space, they see themselves as a shadowed outline on the screen into which jumps the small green critter. The critter tries to 'make contact' with the visitor, steering toward their outline, landing there and attempting to climb up their arm, shoulder, neck until it reaches their head. The user must try to outwit the critter, to move around in the real space, contorting their projected image to try and shake the Critter off. If they fail, and the Critter conquers the summit, it performs a joyful dance to signify its victory.

This simple game represents an early example of the probing of power distribution between user and system, inverting the paradigm of control and navigation common in other interactive forms. Real-time projections of the user's actual shadow are interlaced with computer-mediated graphics. The user's movements directly affect their shadow and influence the computer mediated graphics. The graphics in turn are heavily contrived to influence the user's physical movements. This sets up a simple but powerful play on the real and the artificial, on leadership and submission.

A sonic analogue of Krueger's closed-circuit video interfaces can be seen in works such as David Rokeby's *The Very Nervous System* (VNS). The various incarnations of VNS also work with video analysis as a diffuse, invisible interface, but rather than operating in a visual space, invite interaction with a synthesiser. Rather than the mixed realities of Krueger's projections, Rokeby invites the user to cooperate with this system to 'bring forth' a sonic environment. A schematic of the basic operational loop is shown in Figure 7.3 (right).

Whereas Videoplace manipulates the visitor by evoking a very precise attribution of cause and effect, Rokeby plays with a more complex feedback between movement

Figure 7.3: David Rokeby in Very Nervous System in the street in Potsdam (left) and a schematic of the basic operational loop (right).

and system response. He aims not to invert control, but to develop a relationship of "encounter and involvement". Physical movements over several cubic metres are tracked and intricately analysed with computer vision techniques. The synthesis component is heavily composed, using banks of 'instruments' with preset tendencies. For example, one instrument might be voiced on a snare drum, tend to play on off beats and double its rhythm if you move faster; another might synthesise a brass section and rise in register if you lift up your right arm. Various instantations of the piece have been exhibited, primarily in gallery environments and outdoors in public spaces (see Figure 7.3 (left)), but the systems have also been used in live performance.

Whilst Rokeby had to work hard in early VNS predecessors to achieve the real time motion capture necessary to support the transformation of thin air into a persuasive interface, this can now be easily achieved on a laptop. The use of 'invisible' interfaces in interactive arts is now very common: their immediacy and intuitiveness underlying their attraction as both installation and performance based interfaces.

Across the interactive arts, the use of 'invisible interfaces', whether focused for musical expression, or extended for installation environments creates a very intuitive and flexible mode of interaction which encourages acceptance of uncanny or complex man-machine dialogues. In both Videoplace and VNS, the artists play with the user's concept of control using carefully designed instructions which mediate between sensor input, and the graphical or sonic effectors. These rules govern the appearance or movement of the Critter, or the controls of a particular instrument specification in a VNS performance.

7.1.2 Designing Feedback

As in all digital art, the implementation and materialisation of formal processes and mechanisms is as important than the processes themselves. Just as the mapping from mathematical procedure to sound parameters defines the musical success of the formal structures, so the visual or sonic environments in which the interactive features of the composed environments are made available to perception, define the interactive experience.

Two considerations in particular drove the design of Ashby's Grandmother's Footsteps. Firstly the use of explicit rules mediating the sensory inputs (video camera/sensors etc) was replaced with the homeostat which acts as a self-modulated control system. Secondly thought was given to the importance of how the sonic feedback was delivered. As mentioned above, with responsive environments of this type, where the user 'brings forth' the content, it is vital that they are immediately and sustainably engaged. At the lowest technical level, this means ensuring that the interfaces are accurate and fast, and that they evoke meaningful perceptual experiences (even if this is confusion and frustration as in many cases). In Rokeby's system this is assured by his well developed VNS software. Arguably in his system, user's interest is sustained in part by the complexity of mappings from movement to sound. As he describes it:

> "The feedback is not simply 'negative' or 'positive', inhibitory or reinforcing; the loop is subject to constant transformation as the elements, human and computer, change in response to each other. The two interpenetrate, until the notion of control is lost and the relationship becomes encounter and involvement." - Rokeby (1990)

In VNS, addition, the sonic feedback which users get is heavily composed, so we get fragments of pentatonic panpipes, flurrying shakuhachi and bubbling brooks giving the audiences the sensation that they are composing musical works with their bodies.

In Ashby's Grandmother's Footsteps, some of the feedback material is composed to an extent (generatively), but more central is the exploration of the effect of including feedbacks on a number of timescales. Inspired by Ashby's notion on ultra-stability, in which two levels of feedback subserve an organism's ability to adapt to ongoing environmental interactions, the hope was that providing feedbacks at multiple timescales would engage the visitor more deeply in the system, creating a richer interactive experience.

Such considerations are central to installations in responsive environments, but are also of course of relevance to all forms of interaction. Several researchers have suggested that exploration of mappings in interactive sonic installations is crucial to developing understandings of interactivity which will push performance software forward.

> "I propose that the public exhibition of interactive, responsive sound installations and environments is a good platform for the investigation of mappings that may be inherent to the process of interaction. Of course, the interface design dictates the nature and the scope of all interaction to some extent, but public exhibition exposes the work to an untrained and inquisitive audience, who are prepared to invest time in the development of a relationship with the interactive system. They have no prior knowledge of the rules of engagement, and therefore set out to develop a cognitive map of possible relationships with the system, a map that deepens over time." - Paine (2002), p.298

7.2 Design and realisation

One of the aims of the installation was to play with the user's feeling of control in the space. Ultimately, they have to play the grandmother's footsteps game on the machine's terms and try and cheat it with stealth-like movement. In addition to this simple game playing, the piece contains elements of interactive and generative sound. 'Grandmother' is concerned only with whether or not they move too quickly. Others aspects of their movement through the space evoke and control several other layers of sound. This was designed to encourage exploration of small movements in the space and give the user a complementary sense of control.

- Track 21 gives an example of the output of the system

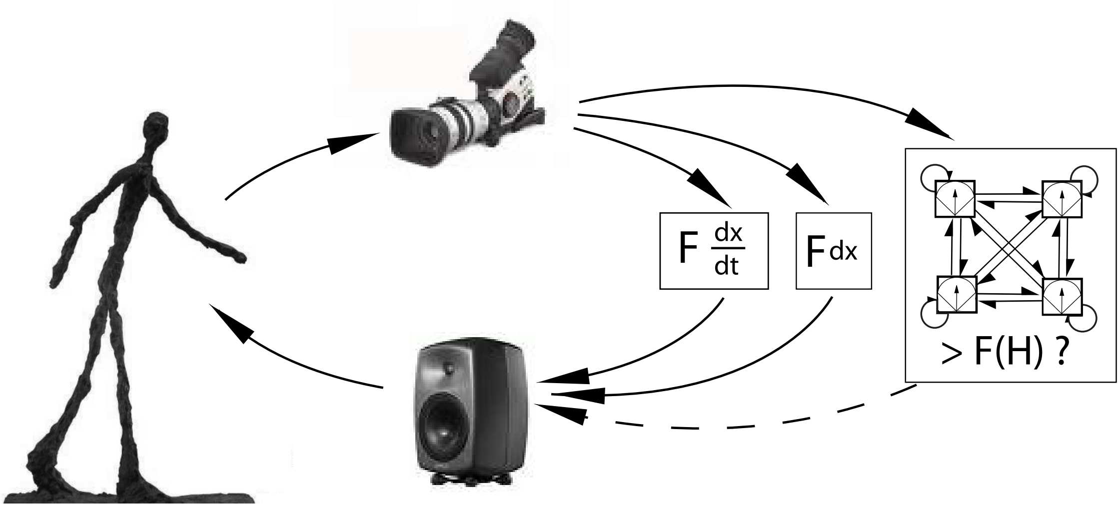


Figure 7.4: Schematic showing basic feedbacks between user, sensors, mediating devices and sound.

7.2.1 Overview

Within the space, the visitor's movements are implicated within three principle feedback loops. Each is associated with a different measure of movement and contributes to different aspects of the sound scape. The principle 'grandmother' control triggers a loud aversive metal-on-metal screeeach if the user moves too suddenly. The actual threshold is determined by the current state of the homeostat (shown as [$> F(H)$?] in Figure 7.4). A low level continuous feedback is given by directly sonifying the rate of change of movement (shown as $F(\frac{dx}{dt})$). This acts to augment the standard sensory-motor loop which engages us in the world, and aims to create a very personal and immediate sense of artificial reality by heightening awareness of movement. Finally the movement of visitors down the corridor (labelled $F(dx)$) triggers progressive changes in the harmonic drones, as well as increasing levels of some of the incidental sounds; this was designed to give a basic sense of progression in the sound world, reflecting the visitor's progression in physical space. In contrast to these sets of mappings between aspects of movement and sound, if the user stops completely still for a certain period, a contrasting set of sounds is introduced.

All three movement measures are derived from a motion detection algorithm operating on a live video feed. The DVCam was situated at the end of the corridor and covered its length. Details are given below in Section 7.2.3. All sounds were delivered via wireless headphones. Originally the installation was conceived to be set in the dark and to deliver sounds via speakers. The use of headphones rather than speakers changed the impact of the sound: the piece is very rich in bass frequencies, using sounds that have quite a strong impact physically, and adding to the sense of 'artificial reality'. On the other hand, the use of headphones makes the more delicate sound of each movement much closer and more immediate. Having only one pair of headphones also provided a much appreciated method of preventing more than one person entering the space at a time.

7.2.2 Composition and Implementation

As with all interactive generative pieces, composition must be approached as the design of sets of possibilities which a user can wander through: fields of sound, if you like which may or may not be heard in any way that you as the composer have ever experienced. In this setting, the qualities of the sounds themselves were also vital in evoking the desired combination of submission and exploration. The mappings from movement to sound are of course also key in creating a sense of contingency which is convincing enough to lure them into the game.

Sound design was approached with the aim of supporting a sense of 'artificial reality'. All material used in the piece were sampled from natural materials, and manipulated through granular synthesis, time stretching and reversing. All sounds are manipulated or processed in a Max/MSP patch which was controlled via OSC from another laptop running Jitter. There are three layers of sound which are directly influenced by the user: a sustained drone which is present as the user enters and differentiates harmonically as they travel through the space; a delicate crunching which is triggered with their every movement and the loud screech from Grandmother when they move too quickly. Other lines are more indirectly influenced as described below.

Grandmother Control in the Homeostat

In seeking a device to play the role of Grandmother in this installation, Ashby's homeostat offers an interesting mechanism for detecting change: it is richer than simply implementing a fixed cut-off point as it responds to minimal changes, but triggers a sudden change above a certain threshold, a threshold which is self-determined. This internal control over sensory-motor mappings gives the system some degree of dynamical independence.

The model homeostat used in Ashby's Grandmother's Footsteps is almost identical to the basic simulation described in Chapter 5, except that rather than issuing perturbations via a button press or feed from audio analysis, the simulated units are joined to the real world, being influenced by the user's movements via video analysis. The network in Ashby's original description represents the couplings between an organism and its environment. Each unit, or collection of units, in the mechanical device then, can be arbitrarily conceived as an organism, or its environment (which could be another organism). The full homeostat can be seen as a formal implementation of the concept of a 'responsive environment'.

The homeostat receives input into one unit via a motion-detection algorithm working on the camera feed. As this is invariably connected to the rest of the network the user's movements impact on the state of the homeostat as a whole. When stable, the homeostat can adjust to small fluctuating values, but if the input is large enough to cause any of the units to exceed their critical values, the network reconfigures as described previously. This triggers a loud screeching sound, but also of course means that as the weights have almost certainly changed the effective sensitivity is altered. The critical threshold for movement is therefore determined by the current internal state of the homeostat.

The sound delivered is a loud metal-on-metal screeeach: a recording of two pieces of resonant metal being scraped against each other[2]. This is simply triggered as a sample in Max/MSP, but modulated incidently each time it is played in an attempt to preserve its aversive properties.

Reconfiguration of the network also triggers updates in the processes controlling the drone harmony (described below). The output of one of the units of the principle homeostat is also used to control the playback speed of an indian bell sample. This is modulated over a very large range producing very different sounds according to the current state of the network. For some states it is inaudible - for example when the outputs oscillate at high speed. At times it appears as a deep resonant bell, at others a delicate tinkle. Other states trigger a slow modulation creating a flanging effect on its natural harmonics.

---

[2] This is a sample taken with kind permission from Arve Henriksen's Planting Trees, Creating Beauty. From the album Sakuteiki (2001) (Rune Grammofon)

Sonified Movements

The sensitivity of the video analysis was set such that the tiniest movement - even of a finger being bent - can be detected. This was crucial in this setting. In many applications motion detection is applied to fairly large movements, aiming to capture gross physical gestures of expression. In this situation, large or sudden movements are banned. In order to provide adequate room for exploration, the space of tiny movements needed to be augmented.

This was achieved by using the motion detection output to trigger and control the playback speed of a sample of some empty snail shells being rolled around in a tea-towel. This source of the sound isn't making any artistic statement, its just a good noise that is both very delicate and also incredibly rich harmonically: at low playback speeds it gives reverberating clunks, at close to original pitch it sounds like someone walking on eggshells, and at high speeds gives an electro schraaaunch. Two different samples are looped at playback speeds that differ by a factor of ten to give a richer sound. Although this is looping, it is rarely recognised as such due to the continuous changes in playback speed.

Harmonic Progressions

Underneath the intricate snail-shell movements and metalic screeches is a bed of drones which differentiate, aiming to draw the visitor along the corridor. These shift harmonically and increase in volume as the visitor moves through the space. A slow constant pounding is drawn through the whole space, increasing from an almost imperceptible breath to a fairly unnerving thud toward the end.

Originally the intention was to use infrared or ultrasonic sensors to give information regarding the person's position along the corridor. This would have given the possibility of creating more sophisticated compositions in the space. However determination with wishing to work with low level components and lack of time to tune what turned out to be a rather erratic device meant that an alternative method of estimating their location had to be devised. After testing several possibilities, it was found that a sufficiently accurate estimate could be made by taking a cumulative reading of movements. [3]

The sound itself is a sample of a bowed cymbal, processed with a purpose-built granulator to produce a fluid, pitch-able and continuous sound which retains the characteristic metallic harmonics of the source. Before someone enters the space, this is a monophonic drone. As the person moves through the space (i.e. every X times movement is detected) this differentiates to produce microtonal harmonies which change and build as they move down the corridor. These were implemented in a similar way to that described in sound installation AdSyMII, i.e. outputs are mapped pitch deviations to create microtonal harmonies. The principal Grandmother network needed to be run at a highspeed in order to retain the required temporal sensitivity, so a second network was implemented which was updated intermittently according to how fast the person moved.

Finally, a little 'hidden' generative sequence was designed which only appears if the visitor stands completely still for more than a fixed number of seconds (typically around 30). This is a slightly lighter series of melodic chimes, similarly produced by granulating a bowed cymbal sample. The period for this happening is set by hand, but the pitches of the chimes are controlled by the same homeostat which sets the pitches of the drones.

---

[3] Even if someone has been in the space a long time, but has not moved very much, it is generally fair to assume that they are somewhere near the start of the corridor. The only time that this is upset is when someone stands stationary part way down the corridor and moves theirs arms or head about a lot. In this case a false measurement is given resulting in an increase in volume of many parts. But at least they were enjoying themselves!

Figure 7.5: A screen shot from Jitter showing the visual effect of the frame differencing algorithm which is used for motion detection.

7.2.3 Technical Details

The installation was run inside Max/MSP/Jitter on two G4 PowerBooks linked via Open Sound Control (OSC) on a local area network. A diagram of the set up is given in Appendix B, Figure B.1. OSC is a communication protocol which enables high speed data sharing between networked comptuers. It is comparable to MIDI, but has much lower lag, allows the specification of data types and formats and can be sent via UDP or TCP/IP.

Video Analysis

Video analysis was performed inside Jitter. Jitter is an extension of Max/MSP which supports realtime manipulation of video, 3D graphics and other data sets within a unified processing architecture. This makes it relatively straight forward to grab and analyse a live feed from a digital camera connected via firewire. Motion detection is carried out with a simple frame-differencing algorithm. A video feed from a digital video camera in manual mode is grabbed at a resolution of 320 * 240 at 30 frames per second. To ensure no movement outside the installation area was detected, an adjustable mask was made, screening areas outside the corridor.

The frame differencing algorithm first calculates the difference in every pixel value between successive frames. The visual result of this is shown in Figure 7.5: nothing moves in the background areas, so the difference between frames for these pixels is [0 0 0 0] i.e. black; movement is greatest at the edges of the figure, or a limb which is moved suddenly (as visible on the left arm in the middle image) producing values approaching [1.0 1.0 1.0] i.e. white. A global measure for the whole frame is then calculated by taking the average difference and normalising. Although very simple, this was effective and could be tuned to suit environments with different lighting levels

Sensors and Switches

One of the installation issues was to ensure that the system was reset after each use, and to be able to ascertain when someone had entered the installation area. There are many ways this could be achieved. Here a combination of simple sensors and an engineering of people's movements with the physical space proved successful. The layout of the installation space made it possible to place the headphones just outside of the area under surveillance, this meaning that you could be fairly certain that they would engage with the piece before the sensor had detected their presence.

A modified passive infrared (PIR) sensor was used to detect when someone has entered the space. PIRs work by detecting changes in infrared radiation which is given off by all objects above absolute zero. These are used commonly in domestic burglar alarms. This was focused using the very low-tech but proven technique of sticking parallel strips of black electrical tape on the front of it, reducing its area of visibility to a narrow beam across the entrance of the space.

At the other end, a push-button switch was placed in a position prominent enough for most people to realise its function. If people did not use it, a fall-back procedure was implemented which reset all software for the next person.

7.3 Reactions and Discussion

Audience activity suggested that the piece achieved its basic function as a game of control: despite no instruction or explanation all users (bar a couple of ardent noise-core fans) quickly understood the game and submitted to the stealth-like movements necessary to play. In many cases visitors returned a second time. In the first run they seemed to quickly grasp the nature of the game and experiment with the range of allowed movement. On the second run they would enter very tentatively and try and traverse the whole corridor without triggering the screaach. The basic game is very simple, but there seemed to be a basic sense of 'achievement' with successful completion - and corresponding irritation by those unable to move sufficiently slowly !

It could be argued that simply playing a highly aversive sound is a fairly cheap way of controlling people's movement, but observing visitor's movements and speaking to them afterwards suggests that the experience was a little richer than just a game of control.

The ferocity of the first screeach inevitably made people stop in their tracks, so opening the space of minimal movements in which the snail-shells operated. Without these continuous feedbacks, people would perhaps would have given up and just stepped back out of the space. The snail-shell feedback was surprisingly effective in not only providing some form of entertainment, but apparently actively slowing down people's movements. In tests carried out in the lab, it was found that you could almost directly control the average speed of movement by adjusting the base-rate of the playback speed. The auditory feedback seemed to act as a positive feedback loop, with slow sounds making slow movements slower, and fast sounds making fast movements faster. The sound seemed to not only heighten awareness of movement, but create an illusion. Many visitors said that it felt like they were walking through gravel, or slurry (depending on the base-rate of playback speed). These two levels of feedback then were mutually supportive in that the screeach set an initial precedent, forcing people onto a slower pattern of movement, and the snail-shell reinforced their slow movements, keeping them away from the zone of detection.

One of the major issues was in setting the gain on the motion detection reading which was fed into the homeostat. Using the frame differencing method for motion detection was perhaps too simplistic as it does not take into account the distance of the person from the camera. As people approach the camera, becoming larger in the frame, it becomes effectively more sensitive. In practice, this wasn't too much of a problem as people became quite stealthy by the end of the corridor. If installed again it would be fairly straightforward to scale the output according to the percentage of frame filled by a moving object.

Whilst the homeostatic network was conceptually attractive, in practice a similar, if not more succesful effect could have perhaps been achieved with a simple switch. Self-modification is an conceptually attractive characteristic, but in practice, in this implementation, it just meant that users were given inconsistent feedback due to the changing threshold.

The balance of positive and negative feedback loops on a variety of timescales seemed to bolster engagement with the system. The organic, and slightly threatening nature of the sounds also supported the sense of an artificial reality which by all accounts was created by the experience. Whilst the use of such distributed interfaces and feedbacks is not perhaps suitable for on-stage performances, such environments provide fertile ground for exploring modes of interaction and engagement.

Chapter 8

Self-karaoke Machines: Collaborative Man-Machine Improvisations

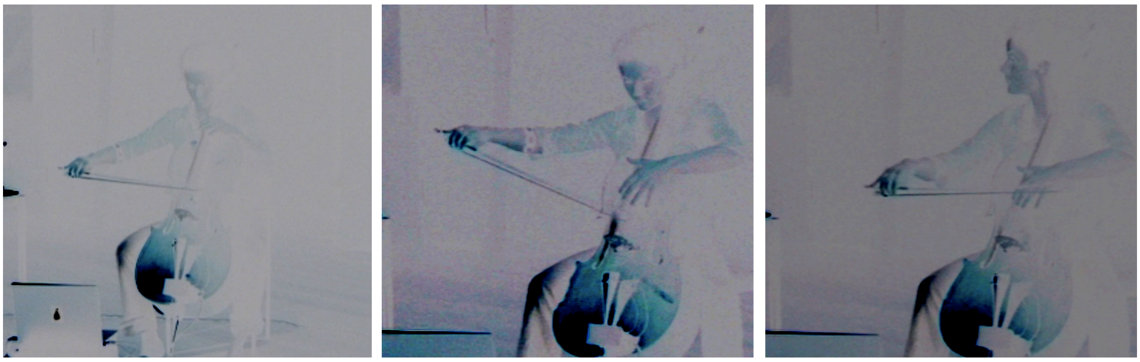


Figure 8.1: Fond Punctions performance at Third Iteration, Melbourne (2005).

This chapter brings us to the final project in this body of work which realises the end goal of inviting adaptive systems onto the stage. The *Self-karaoke Machine* is a generative system for collaborative improvisation which has been used for both live performances (with cello) and as an installation for public consumption. The project is placed squarely at the intersection of the generative arts and improvised interactive computer music and aims to show how aspects of the two practices can be mutually complimentary. As a performance system, the investigative aim was to explore whether the simple adaptive systems described in Chapter 5 can stand up in a live performance situation: whether despite their complete lack of musical 'knowledge', their formal behaviours can be implemented in such a way as to provide inspiration to the performer and engage the audience with a convincing man-machine collaboration. The artistic aim was to unite the artificial and the acoustic and explore the meeting point of digital generative practice and instrumental improvisation.

Section 8.2 provides a description of the system design and discusses its attributes from a performer's perspective as well as relating feedback from audiences at concerts. Performances have been very well received by audiences with a very wide range of musical taste, but as a performance it is of course difficult to pick apart the contributions of the system itself from those of the performer: all we can know is that the two worked well together. Whilst this is of central concern, in order to examine how much work the system itself was doing, a modified version was installed as an interactive installation at

The Big Blip 05, a week long festival of digital art. Section 8.2.4 describes the modifications made to the system to make it suitable for public use and discusses feedback from visitors to the exhibition.

Section 8.3 takes a step back and considers the impact of bringing a live performer into the generative loop.

- Documentation of the Self-karaoke Machine is provided on the accompanying DVD. A film of a Fond Punctions performance given the Third Iteration concert (2005) is given, and additional recordings are on tracks 21-24. Track 21 is an edit of this concert and another given in Sydney the same month. Tracks 22-24 were produced in the studio. They are similarly the result of improvisations with the system, but exclude any dry cello or vocal samples. An example of someone playing in the Self-karaoke Pond is given on track 25.

8.1 Generative Processes in Live Improvisation

Musical improvisation is a very natural setting for generative processes. Winkler (2001) suggested that Free improvisation represents the greatest challenge to the designers of digital music systems, but an improvisation framework also offers the greatest freedoms for generative art in enabling the ancient tradition and intuitive activity of human improvisation and the nascent vagaries of digital generative practice to come face to face on their own terms.

The intersection of cutting-edge digital arts and traditional art practices arguably offers the most fertile ground for sustainable cultural evolution. There is a slight proclivity in the new media arts toward techno-fetishism: employing techniques or tools, just because they are there rather than for any particular purpose, and losing touch with a wider arts context. This was expressed recently by a post on the generatorX forum:

> “. . . there’s a tendency towards being so immersed in the technology used that you forget to consider your work in a broader perspective. Once you start working on images with an artistic content, you not only have to relate to discourses of generative or new media art, but also start relating to the tradition of visual arts in a much broader sense. In addition to genre-specific discourses you have to start addressing issues of form, material, color, content, context, history, etc. in a much wider sense. Maybe the new media art scene sometimes should put a little less emphasis on “new” and “media” and more on “art” ?” [1]

One of the issues of course is that the widespread employment of interactive and generative digital processes has radically altered the forms and conceptual basis of many arts practices, particularly in the visual domain. This is evidenced by members of the community working hard to form new critical frameworks that can deal with the peculiarities of generative art (Woolf (2004), Whitelaw (2005)). Such texts deploy discourses, some aspects of which have little correspondence with critical approaches in the wider visual arts world.

In the musical domain and in improvisation in particular, it can be argued that generative and interactive processes fit very comfortably and extend an established tradition, rather than turning it on its head. This means we can build on discourses and practices of the past, rather than having to start over with a whole new set of practical and critical approaches. Collins (2003a) for example has noted that live computer music forms the

---

[1] Posted by Trond on October 6th 2005 at `http://generatorx.no` in response to Golan Levin’s ‘Three questions for generative artists’. Presented at GeneratorX 2005 and posted on Oct. 5th2005

perfect material for generative processes in terms of accommodating non-linear structures. Improvisation itself can perhaps be seen as a generative process, the product of the intuitions of the performer unfolding in the context of an environment formed by other's musical suggestions. Lewis presents his understanding of improvisation which could equally be taken to describe the realisation of a live interactive generative system:

> "In the general, everyday-life sense, the activity of improvisation can be viewed as a domain-specific, structure-generating interaction within a particular environment complex. In the musical domain, improvisation is neither a style of music, nor a body of musical techniques. Musical improvisation is one domain among the various possible domains of improvisation – an interaction within a multi-dimensional environment, where structure and meaning arise from the analysis, generation, manipulation and transformation of sonic symbols." - Lewis (1999), p.101

If we confer with this understanding of musical improvisation, then digital generative process can easily be accepted as 'just another musician in the band', and their differences welcomed, explored and exploited within the established traditions of musical improvisation.

8.2 Fond Punctions

Fond Punctions is a performance which uses the Self-karaoke Machine. The performance aimed to present a sense of collaboration between me, the cellist, and the digital system. The program was designed to explore the potential of simple adaptive systems in live performance and by extension to examine what forms of interaction are engendered. The desire to be able to perform solo electro-acoustic gigs (i.e. with no-one at the helm of the laptop) laid down a number of additional practical constraints which influenced the system design.

As I play the cello, the software needed to be able to run with no intervention. When sitting or standing behind a cello, bass, or any instrument with both hands fully deployed, it is physically awkward and invariably musically disruptive to turn to the track pad and keyboard of a laptop, so the system needed to be robust and rich enough to run unmanned.

Figure 8.2: Setting up for a Fond Punctions performance at Artpool. Budapest

The importance of engendering a pay-off between adaptability and dependability was discussed in a general context in Chapter 3, Section 3.3 where it was suggested that this balance is desirable on at least two levels. Firstly at the behavioural level, especially for live extemporisation, the system needs to be flexible enough to accommodate the intrinsic unknowns of improvisation, but reliable enough for live performance. This, it is suggested is one of the fortes of simple adaptive dynamic systems in exhibiting an unpredictable range of responsive dynamics within a circumscribed behavioural field.

Secondly as a composition tool, it was suggested that systems with a small number of parameters which influenced the global state of the system was desirable. This provides a global-control in performance situations which pushes toward a more collaborative model than the 'auto-pilot' approach proposed by Collins (2003a).

In a live situation where there is no one to twiddle knobs, some other solution for controlling these parameters is necessary. In this system, the modular approach adopted in the generative installation systems is developed to include two conceptually distinct but interacting dynamical systems which co-determine both the sonic output and form the basis of a visual projection.

Figure 8.3: Performance of Fond Punctions in the Friends Meeting House, University of Sussex.

As discussed in Chapter 2, it can be useful in designing interactive systems to extend the frame of consideration beyond the analysis and composition modules themselves and consider the performer and software as two interacting components in a larger performance network. The design of this system adopts this more systemic perspective, the implications of which are discussed in Section 8.3, and incorporates a visual element which plays a fundamental role in the overall performance network.

System design also sticks firmly to the minimal approach adopted throughout and investigates the slightly contentious effects of removing the numerical input that typically drives the computer system according to analyses of the player's output. Rather than analysing what the performer plays, the approach taken here is to take samples of the actual sound material which is then manipulated by the generative engine. This closes the loop via a sonic rather than a digital information circuit. The performance then becomes a collaborative effort with the player deciding what to 'feed' the system, and the system deciding what it will do with it – which in turn influences the course of the player's improvisation.

8.2.1 System Overview

Algorithmically the system is based on two distinct but interacting systems: an Ashbian homeostat as described in Chapter 5, Section 5.1.1 and a simple physics simulator which describes the motion and collision of floating particles. Both of these models act to parameterise a granular synthesis engine which operates on samples taken by the player during a performance and determine the movement of objects in the video display (a sample screen shot of which is shown in Figure 8.6). Structurally the systems function at different levels: the homeostat operates at a rhythmic and phrasal level, the physics simulator determines longer term structure. Finally a broad performance structure is implemented by specifying a set of rules in the form of conditions such that generated events in the system come to a natural close.

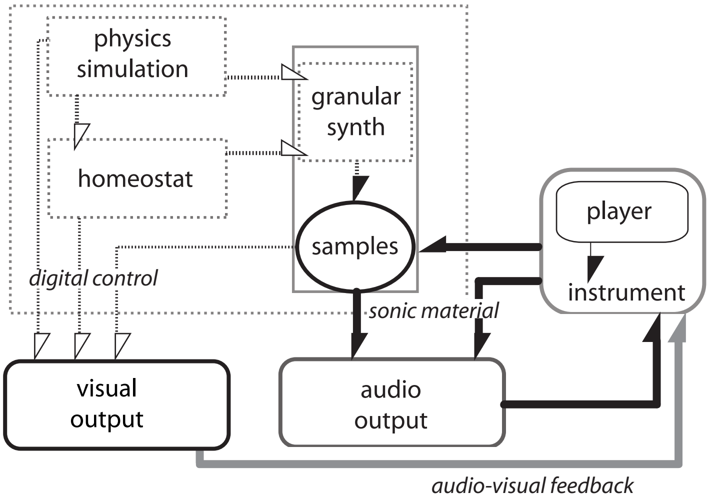


Figure 8.4: Schematic showing the network of influences between components in the whole performance system.

The homeostat acts as a responsive pattern generator, creating re-compositions of the musician's acoustic improvisation. Multiple audio samples are taken during the performance, and the output values of individual units in the homeostatic network are used to control *when* sound grains are triggered and from *where* in the sample they are taken. Different grain sizes and densities vary the acoustic/electronic or melodic/rhythmic feel, creating the impression of digital re-interpretations or timbral reflections of the performers improvisations. Details are given in Section 8.2.2.

The performer controls only when to take the samples and of course what to play, which as an improvisation is directly influenced by the sonic output of the system. In system terms, this closes the feedback loop on a macro scale; in performance terms this throws back fresh musical ideas which push the improviser in new directions.

8.2.2 Component Details

Homeostat control of the Granular Synthesis Engine

The main sound engine employs granulation techniques to recompose the samples taken by the performer. The granulation engine was implemented in Max/MSP using Nathan

Wolek's *gran* object[2].

During a performance up to eight different samples are held at any one time and are overwritten a number of times throughout. Typically these are between five and twenty seconds long although nothing prevents times outside this range. The length of each sample is stored and the current output range of the homeostat is scaled and mapped dynamically to the individual samples. The outputs of each of the eight units in homeostat are used to determine from where in the sample grains are taken. The granular synthesis engine allows manipulation of the size and amplitude and pitch (i.e. playback speed of the original sample) of each grain.

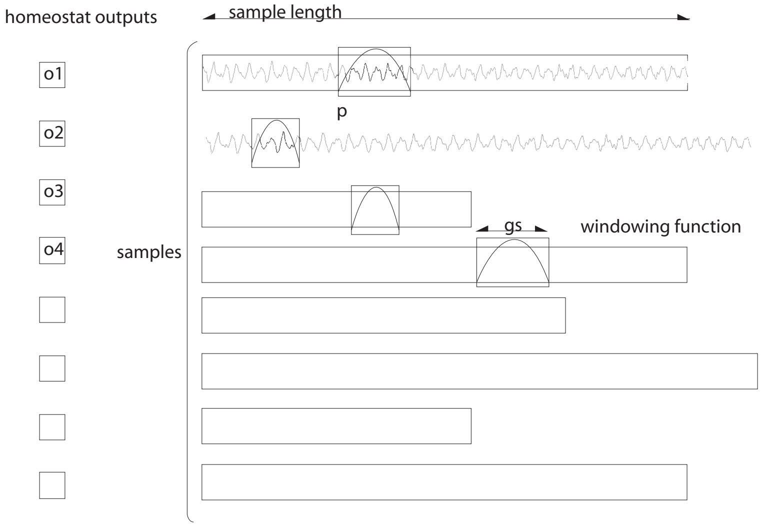


Figure 8.5: The outputs of each unit in the homeostat (O_n) are used to determine the point p in the sample that the grain is taken from.

The stored samples are divided into two halves. Samples 1-4 are read by eight gran objects with grain sizes between 400ms and 2000 ms. These are voiced at the original pitch. This preserves the pitch and timbral characteristics of the original sample and for higher values even melodic/ rhythmic fragments can be recognised. The position of the grain in each file is determined by the output value of individual homeostat units. Samples 5-8 are read by eight gran objects grain sizes between 90ms and 300ms and at higher speeds (typically between 8 and 32 times normal pitch, although some great effects can be made using higher values). This produces the pops and clicks characteristic of sparse granular streams.

For a certain range of viscosity values, the homeostat outputs tend to frequently oscillate at the viscosity value itself. This is probably a side-effect of the way in which viscosity is implemented[3]. This is exploited as a means of introducing variation. Grains from samples 1-4 are triggered whenever the output of their associated unit *is not* equal to the current viscosity value. The shorter grains from samples 4-8 are triggered when-

---

[2] This is an object in his Granular Toolkit package available at: `http://www.nathanwolek.com/software.html`

[3] Recall that the damping effects of viscosity are implemented by constraining the amount by which any one output can vary in any one time step. This means that if any unit output is near zero, and it receives a very large positive or negative net input, it will swing high or low and be constrained to the viscosity value. When this occurs for two or more units, this seems to set up an oscillation where they each get locked onto an oscillation between positive and negative values of the viscosity value, forming a stable attractor in which the system stays until weights or viscosity change. These effects have not been rigourously investigated, but proves to be robust enough for reliable use in this context.

ever outputs of the corresponding homeostat unit *are* equal to the current viscosity value. This typically occurs for mid-range viscosity values.

This second set of samples produce a more rhythmic texture which is elaborated with a simple probabilistic procedure to avoid very repetitive rhythms. Essentially, a filtering process is implemented that only lets every N^{th} trigger pass, where N is reset each time a collision occurs in the visual display. A complimentary process treats these rhythmic outputs: delay lines are set on half the triggers so that when a trigger does arrive, it is duplicated at varying fractions of the regular beat. This creates a break-beat effect by removing some of the regular beats, and fracturing those that are passed.

In all cases, panning is implemented simply by passing the grain to whichever channel (left or right) is free. If neither channel is free, the grain doesn't get voiced. This creates a self-selection process which automatically adjusts the density according to grain length. Equally, longer grains tend to get panned more evenly, often bouncing back and forth, whilst the shorter rhythmic samples tend to occupy one channel, getting thrown across only when densities are high.

An example of the system running with some long grain sizes is given on track 22 where whole fragments of the original samples can be clearly heard. Track 23 is similarly derived from cello samples, but shows the effect of using short grain sizes and high pitch multipliers. In both these examples, the outputs are triggered purely by the homeostat outputs, and not passed through the stochastic rhythmic process. The effects of this process can be clearly heard on track 24. Initially N is high, creating very sparse rhythmic textures. As the bass grain enters, N is reduced to create a rhythmically dense texture. These are all vocal samples.

Motion Simulation and Video Projection

The motion-collision equations in the physics simulation describe the movements of various objects in the video projections. One set of equations describes the trajectories of the three white bubbles which can be seen in Figure 8.6. These trace fixed paths described by simple functions (sine, quadratic etc.) and control the playback of the very first sample taken during the performance. This sample remains fixed throughout (see below). As each of these collide with the left and right boundaries of the space the initial sample is triggered (forward or reversed accordingly) at a speed determined by the length of the trajectory. This creates a polyphonic drone which shifts throughout the performance as the path lengths are incommensurate.

Another set of motion equations describe the movements of two cellular aggregations which move around a finite space, rebounding off the perimeters, and colliding with the bubbles. The cross-hatches which can be seen in Figure 8.6 mark the centre of each of these aggregations. Collisions between the bubbles and the cells perturb the homeostat, forcing it into new trajectories. Visually this is signified by a white flash. Acoustically weight changes invariably push the homeostat into a new field, meaning that the pattern of values across its output change, creating a sudden change in the parts of samples which get voiced and so a sudden change in the material heard. Each ring represents a sample: each appears as a new sample is taken, and the size is proportional to the length of the sample. Each coloured dot inside the ring represents the point in the sample at which the gran object is currently reading. When no grain is voiced, the corresponding dot is unfilled.

As the aggregations rise and fall, their vertical position controls the viscosity of the homeostat, as well as the grain amplitudes and lengths. This allows for a certain level of engineering of the overall shape of the performance in terms of dynamic range etc. Because these move along continuous paths, imminent collisions can be anticipated, and accounted for by the performer in their improvisations.

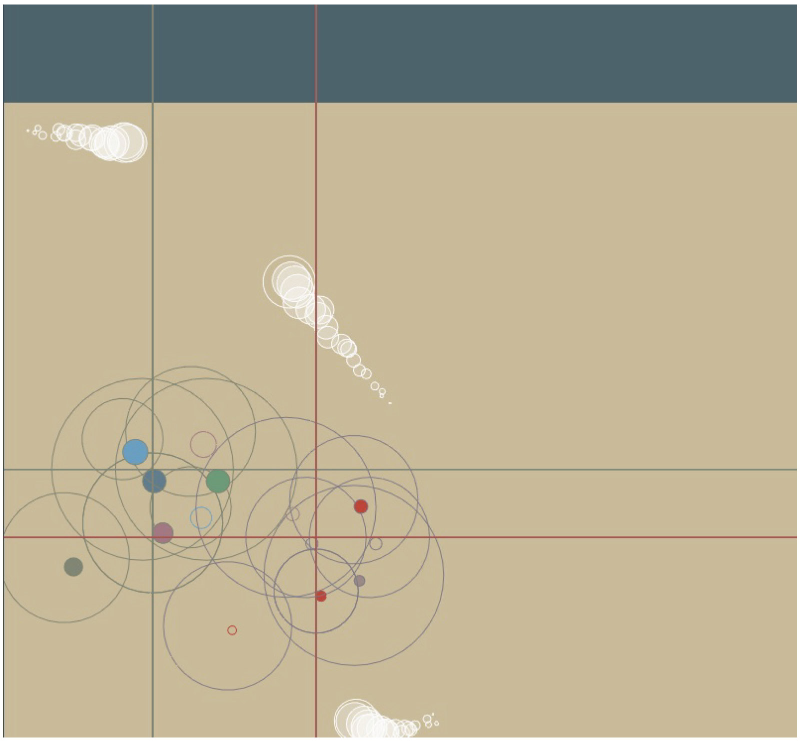


Figure 8.6: Screen shot of the video projections in the Self-karaoke Machine.

Implementation details

The system runs on a single G4 power book, running Max/MSP and processing. The homeostat is implemented as a Max object and all sampling and granular synthesis takes place in a Max/MSP performance patch. The graphics are implemented in Processing. The two systems communicate via OSC internally.

A foot pedal was made by hacking into a USB games controller, interrupting one of its switch circuits with a heavy-duty push-on push-off foot switch. This is one of the simplest ways to condition a voltage difference such that it can be read as a USB input device as the games controller contains a circuit that performs the same function for its own switches. The USB port can then be read using the *hi* object in Max which reports streams of date coming in from any controllers attached via USB.

Performance specific settings

Within human-human improvisations, performers often sketch out a rough structure. Within traditional forms this might be a set of chord progressions and agreement over the order of improvised solos as in traditional Jazz. Even in free improvisation players often formulate some form of game plan. This can be helpful for shaping the performance, but more pragmatically in a gig situation, ensures that the set accords with the time constraints which are set by most public concerts. Similar posts can be set when working algorithmically. For performance purposes a rough architecture is pinned by specifying a few conditional rules. These can be thought of as demarcating stages in the performance.

In performances made to date, the system starts 'empty'. No samples are included, and there is nothing on the screen. The first sample taken triggers the first of the white bubbles to be released which also acts to playback the sample as described above. The remaining bubbles appear consecutively as each traverses the bounds of the graphical space. The player is of course free to play as much or as little as they want between taking samples. The next sample taken causes the first cell to appear, and is duly voiced by its associated homeostatic unit. This continues until four samples have been taken, after which samples are overwritten. As each set of four is taken, the space within which the cells can float is increased, until they hit the ceiling. At this point the second set of four samples can be taken - the first set remaining fixed for the rest of the performance - and the same procedure followed. When both sets hit the top, the variable in the physics engine simulating buoyancy is reversed, such that the cells tend to float rather than rise. The performance ends when all the cells have sunk from view. As vertical height controls the amplitude of the grains, this coincides with the sound dying out.

Less structured alternatives have been explored, such as creating an energy variable, which is increased in accordance with a measure of the sum of average amplitude of current samples, and decreases as sound is emitted. This forms an interesting relationship whereby the performer has to 'keep the system alive'. For public performances however, it is a little too unpredictable.

8.2.3 Personal Reflections and Audience Reactions

Performing with a completely automated algorithmic system can be seen as a form of generative masochism: even an unmanned laptop running a fixed accompaniment represents a death wish with which few performers will willingly dally. However handing over such complete control to a generative system in public concerts provides the ultimate litmus test for the system's ability to make convincing musical contributions.

As a player the system demands an interesting balance of completely open intuition and careful strategic planning. The homeostat exhibits many of the features of a small child: sometimes throwing out inappropriate contributions at the most embarrassing moments, other times astounding you with naive yet perfectly formed insightful suggestions. Great consideration must be paid, particularly harmonically, to the selection of samples, as any element of any sample you take can get be thrown back in any combination. This has pros and cons as you may end up with an overly-sickly consonance or vile clashes. With careful planning however, interesting modulations can be achieved, overcoming one of the major drawbacks of loop samplers in which performers tend to stick in the same key for half and hour.

As a deterministic but unpredictable system the behaviours of the homeostat can only really be understood experimentally. The fact that the final output is a product of the state dynamics and structures of the samples that you take adds another layer of non-linearity which defies any forms of logical analysis and can only be approached on a very intuitive performative level. This is true not only of taking single samples, but in learning how best to supply the rolling bank of samples which the system holds. You can try and repeat the same effects, starting with the same seed, playing the same material and taking what you think are the same samples, but the sensitivites to various aspects of the environments which impact on the final outcome are such that something new and unexpected emerges each time. The mode of interaction therefore perhaps differs from both the instrument model, and the conversation model. There is certainly a level of mutual influence, but this is perhaps best described as a collaborative interaction. Overall the system provides a strangely comfortable cyber extension to improvising, transforming improvise fragments into something new and surprising which push your improvisations into new directions.

Performances to date have been very well received in a number of very different venues amongst ardent generative art fans and practitioners, the general public and musicians outside generative practice. Testimony to the universal appeal perhaps were the enthusiasms expressed at one gig by both a contemporary classical composer and an up and coming noise-core laptopist. The former exclaimed that he found the collaboration 'Awe inspiring', the latter proclaimed it: 'properly wicked'.

As an overall performance there is undoubtedly some degree of fascination with the combination of cello, unmanned laptop and visuals which perhaps woos people and distracts from any consideration of the actual musical content. Certainly people seem to enjoy this combination. The very physical aspects of cello playing are undoubtedly welcomed amongst electronic music communities dominated by laptop performers. Many people expressed an appreciation of the audio-visual relations as a successful augmentation of these gestural-sonic contingencies. The relationship between visuals and sound is perhaps more complex than in typical VJ performances, or even than situations where the visuals react to sonic output. When a sample is taken, there is a direct correlation between a performance gesture, its appearance in acoustic and digital sound and the appearance of a new object in the visual display, creating correspondences between what is seen and what is heard in both the synthetic and acoustic/ physical worlds. In addition, there is the reversed connections as events in the visual display - collisions and vertical movements - influence sonic events.

Whilst there are enough direct correspondences between what is seen and heard to reveal insights into the processes underlying the music, these cross-causalities evoke an element of detective curiosity in the audience. There seems to be an important balance in which just enough is revealed so that contingencies are perceived, but enough is held back so that people are engaged, almost analytically, in understanding the process. This may be seen by some as a distraction from a purer musical appreciation, but it is also a central aesthetic in interactive and generative arts which adds another dimension to play with as a composer.

It is encouraging too that an interest has been shown in the recorded outputs of the system, suggesting there is some value in the system musically, rather than just as a curious 'show'. Many electronic music producers and enthusiasts have been excited by the freshness and liveliness of the tracks produced through improvising with the system, suggesting that the basic behaviours of the homeostat accord with current yearnings of the computer music community. Personally I was pleased with the balance of the artificial and the real achieved, both sonically and structurally. The combination of the acoustic cello and its granulated samples are complimented by the balance between physical gestures of performance and the lively complexities created by the homeostat.

The Fond Punctions performances have been very well received, and as a performer and musician it feels personally that the system is doing some work in terms of extending both compositional and performance possibilities. However in wanting to put forward the use of simple adaptive systems per say, and even this specific implementation, it seems important to examine how other people interact with them. In considering some of the desired characteristics of creative digital tools, Golan Levin (1994) proposed a number of characteristics by which to judge the success of an instrument. In terms of professional musician's adoption of such systems, perhaps the most important are that the potential outcomes are "inexhaustable and extremely variable" (p.54) and in addition that it is "infinitely masterable" (p.56). We will return to a discussion of these issues in Section 8.3. These are considerations being examined for a broader range of models in a forthcoming workshop (see Chapter 9). In the first instance, however, the accessibility and flexibility of the Self-karaoke Machine in particular, was examined by installing it in a child-friendly week long exhibition.

8.2.4 Self-karaoke Pond in Installation

In an exhibition setting what must come before 'indefinitely masterable' is what Levin (1994) has described as "instantly knowable" (p. 56) i.e that the rules of operation are obvious and immediately available. If someone cannot work out how to interact with an installation in less than about 25 seconds, they will simply walk on to the next exhibit.

An open exhibition provides a fantastic opportunity to test the accessibility *and* depth of digital interactive works, as there will invariably be both four year old children present, wanting instant gratification, but also interactive art buffs wanting something fresh and engaging.

- An example of someone playing in the Self-karaoke Pond on hooter and blues harp can be found on track 25.

Figure 8.7: Visitors to the The Big Blip 05 playing in the Self-karaoke Pond.

In Chapter 4, the potential for over-theorising compositional schemes was noted, along with the attendant danger that the resulting output may be utterly incomprehensible to the listening audience. In the same way it is seems possible that what may feel like an intuitive and flexible interactive generative tool to its designer, may be similarly incomprehensible to a member of the public. It seemed important therefore to establish how naive visitors, and particularly children engaged with this system. In addition, on a software engineering level, there is no better test for the stability of a system than leaving it unattended for a week open to abuse from renegade children.

Installation specific set-up

Several adaptations to both the physical interface and software of the system were made for the purposes of the installation. The interface was adapted to take a microphone input and be operated with a games controller joystick. Instructions were given in the simple form of a diagram showing what the joystick controlled (shown in Figure 8.8). Physically the Self-karaoke Pond was installed in a small space that provided some privacy so that people were not afraid to make noises. In the space there was an arm chair and a coffee table which offered some toy instruments for the vocally-shy. Behind this was a back-projection of the visuals. The set up is shown schematically in Appendix A, Figure B.3 and can be appreciated from the image of the small boy playing in Figure 8.9 (left). The audio was delivered over loud speakers as it quickly became evident that people wanted to work in pairs and friends/mothers/children too shy to have a go themselves wanted to hear what was going on.

The software was essentially the same as that described above with a few surface modifications. The posts laid out for performances were removed such that everything floated freely in the space. A 'clear' button was added for obvious reasons which wiped all the stored samples and cleared all the images on the screen. When someone started afresh they could load up to 8 different samples as before, after which they started being overwritten in the order they were saved. As illustrated in Figure 8.8, the main thumb button of the joystick acted as the stop/start recording and the trigger acted to clear the memory. In addition several different settings options were offered which switched

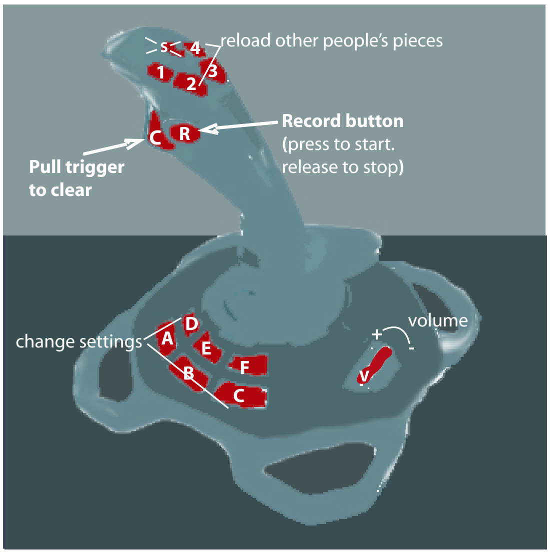


Figure 8.8: Diagramatic instructions for Self-karaoke Pond.

between a selection of preset grain sizes and pitches. This provided a range of output from sweet harmonic versions, where the original samples were easily recognisable, to more dissonant and electronic regurtations. Finally as means of encouraging interaction, there were a number of saved examples of other people's efforts that started playing if the system hadn't been used for more than 5 minutes. This served to to illustrate what was possible.

The installation was incredibly well received by many and left others somewhat nonplussed. Interestingly the enthusiasm levels conformed to a rather strict demographic trend: namely children and musicians loved it and other adults were either too embarrassed to make noises in public or seemed not to understand the appeal. One little girl stayed in there for 25 minutes making a 'halloween sound scape'. I found another mother on the verge of tears (of joy) at the 'beautiful music' her 18 month year old boy had made by babbling in time to it. What was even more encouraging is that whilst children of four years could make animal noises into it which got tangled into strange electronica, I found several accomplished musicians deeply engaged creating complexities I had never managed to achieve myself.

Although not particularly designed to support creativity or musicianship, the overwhelmingly common comment referred to people's surprise at their ability to make music. Some other comments from the exhibition include:

- Very nice you could lose yourself for hours
- This is amazing, I and my friend made beautiful music together.
- Very interactive lovely and fun art.
- My son said it was brilliant, I did no know I was so musically gifted.
- I like that thing love Charlie.
- Couldnt get my three kids off this
- This is sooooooo Goooooood!!!

- I cant believe that a human has made this program its so clever.
- Think its really good. Touch of genius.
- I like when it rkordid me [sic]

Figure 8.9: Some very small people playing in the Self-karaoke Pond at The Big Blip 05.

Feedback from the exhibition suggested that the system was 'instantly knowable' but also that the interaction had some depth, keeping some people there for up to half and hour, and making others return up to five times. Part of the interest perhaps was that it was sample based, delighting people merely in the sound of their distorted voice, but the many comments referring to suggest that the some aspect of the system, arguably the homeostat, is doing some work transforming their voices into something more than patterns of sound.

8.3 Live Improvisation in Generative Systems

The Self-karaoke Machine represents one way of integrating the exploratory potential of digital generative art within the traditions of live performance, adding a new twist to an ancient tradition. In this final section, the perspective is reversed, and the impact of introducing a live performer into the generative loop is examined. In Chapter 3, Section 3.3 the constraints of some ready-made generative composition tools were mentioned. It has been noted that whilst the generative process offers possibilities for exploring unchartered aesthetic territory and are seen to hold promise of exposing results "beyond our wildest imagination" (Rinaldo (1998), p.376), in practice many systems are constrained by the predilections of the programmer. In the case mentioned in Chapter 3, a member of the generative arts community commented that playing with Sseyo's KOAN system felt like remixing pieces pre-programmed by the development team rather than creating anything genuinely new. These sorts of conclusions are frequent within the generative arts community and several authors have outlined characteristics of generative systems which release these constraints. The next section reviews some previously proposed methods of over coming these restraints in the context of a generic scheme for the generative process.

8.3.1 Creative Constraints in Generative Systems

Constraints arising from inevitable decisions in the software development process are a problem not only for generative digital systems, but for software tools in general. Almost any tool or medium, physical or digital, leaves its characteristic mark on the artwork with which it is created. In many cases it is these very characteristics that inspire their use. For example all paintings made with oils will have something in common regardless of the style of painter or the subject matter, and these peculiarities will inspire the selection of oils as opposed to water colour or pastels for particular projects.

These characteristics can be seen as a constraint, but can be distinguished from *creative* constraints by considering factors such as those put forward by Levin (1994). Recall that he suggested that "a feature of a successful instrument is that its results are inexhaustible and extremely variable" (p.54) and in addition that it should be "instantly knowable, and indefinitely masterable" (p.56). This is achieved in physical tools like the humble pencil, a drum stick or even a piano: the smallest child can immediately pick any of these up and do *something* with them. Yet someone could also dedicate their life to practising and never exhaust the possibilities of further refinement. In addition competent drummer or illustrator will develop a personal style and be able to express themselves through the drum stick or pencil with their unique and personal voice. This flexibility is rare in any digital tool, and is a particular problem for generative systems which are presented as 'creative tools'. As Dorin (2001) notes, in many generative systems that are offered as creative tools, such as Latham's *Mutator*, it is impossible to express a personal voice, much less leave your characteristic mark. As he puts it: "none of the pixels voice the thoughts of the wanderer" (p.10). Whilst the interface may be instantly knowable it offers no scope for excellence: "there is no means for distinguishing a master from a relatively inexperienced user" (p.6).

In understanding the root of these constraints it is helpful to consider the components of the generative process in more detail. Dorin and McCormack (2001) have proposed a set of biological analogies which distinguish between different aspects of a generative work. Illustrated in Figure 8.10 we can conceive of these separate elements using the biological notions of genotype and phenotype as used in discussion of GAs within Alife research. The designer constructs a generative process (the genotype), and typically stands back to observe the phenotype unfolding in the hands of an automated procedure (the enaction of the specification). In many digital generative art systems, the genotype acts to structure a pre-specified medium, whether it be pixels (Todd and Latham (1991)), MIDI notes (Miranda (2003)), old washing machine parts (Berry (1986)), mould on photographic film (Montag (2000)), or the behavioural characteristics of robots (Rinaldo (2000)), creating the phenotypic realisation or artefact with which the audience engage. If the genotype specification includes mechanisms which are responsive to environmental feedback, the audience can also interact with the phenotype and potentially influence future outcomes of the system, as in the many implementations of aesthetic selection in an IGA or twiddling the parameter knobs in KOAN. In both these cases the artist/ programmer has designed an algorithmic engine (the genotype), a set of primitives (geometric forms or MIDI sequences) and a set of mappings which determine how these primitives are combined under the genotype. So not only the genotype, but also the material from which the phenotype is formed are designed within a digital system.

Several people have suggested particular properties of generative systems which potentially afford a greater freedom. Alan Dorin (2001) for example suggests that the designer's control may be relinquished by using aesthetic selection to steer the non-linear interactions of self-organising primitives in order to generate complex higher level emergent phenomena. The programmer would still specify the basic elements and how they interact, but the user could then enter an open-ended conceptual space, sculpting the system into a unique complex emergent structure un-envisaged by the author. This seems to open the space of artistic possibilities offered by other tools, i.e. to readdress the balance between the artistic skill of the tool's creator (e.g. Stradivarius) and its user (e.g. Menhuin). Within a generative art framework the thought of such control whisks us away

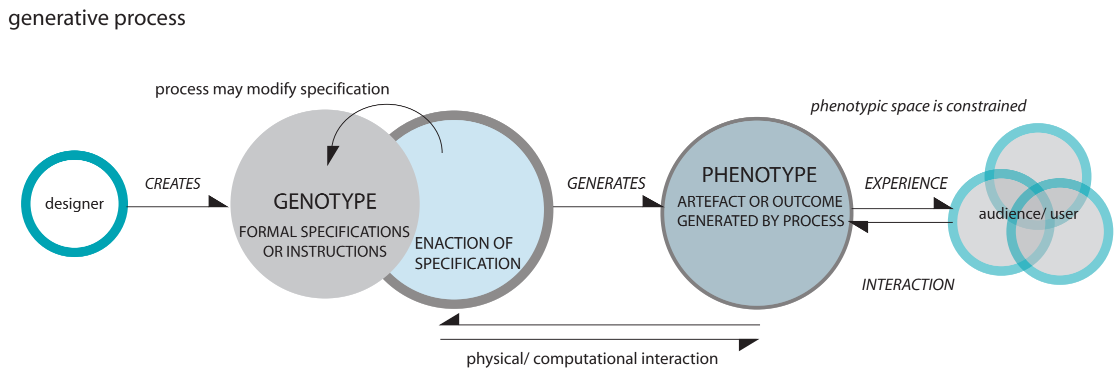


Figure 8.10: Overview of the interactions and influences in the generative process. (with kind permission from McCormack (2004)). The user's influence on the final outcome is constrained by the designer's decisions over the genotype, enaction process and material from which the phenotype is formed.

for a brief cyborg pas de deux around the grounds of the computational sublime. But the problem is, as has been noted elsewhere (Bird and Webster (2001)), that whilst certain types of emergent behaviour can be demonstrated in silico, there exists no un-contended digital system that exhibits truly open-ended dynamics (Smith and Bedau (2000)). The emergence of multi-level phenomenon is a deep open problem in biology (Bedau et al. (2000)), leaving intuition as the only guiding principle in the initial selection of suitable primitives. Finally as Bird and Webster (2001) suggest, the mapping of these (yet-to-be-digitally-attained) dynamics into a perceptual medium for artistic ends is non-trivial.

Another possibility which has been raised as a means of escaping the designer's control and broadening the scope of possible outcomes, is the creation and artistic application of 'creatively emergent' systems (Bird et al. (2002)). The concept of creative emergence is closely linked to Cariani's taxonomy of adaptive robotics (Cariani (1992)) which explicates how this can be achieved in organisms and robotic devices. Cariani outlines one way in which organisms and robots can be differentiated into three levels of adaptivity according to their component parts, or *primitives*. In a robot, these primitives refer to sensors (such as infrared), effectors (such as wheels) and *control mechanisms* which determine the behaviour of the robot by mapping between the two. We can think of these *primitives* as letters of an alphabet that can be combined in different ways to form different words, but cannot themselves be divided into constituent parts.

According to Cariani's taxonomy the simplest robots are described as *reactive*. All the primitives are fixed: control mechanisms are hard wired and sensors and effectors can never change. These are comparable to traditional acoustic or many electronic and digital instruments: the sensor (key, button, switch etc) is pressed and a fixed control system triggers a fixed response (a certain sound, pitch etc.). The simplest adaptive device is able to change the relationship between its sensors and effectors according to experience: it can't change its actual sensors or effectors, but the mapping between them can alter in response to feedback from the environment. The homeostat used in Ashby's Grandmothers Footsteps is arguably a very simple implementation of this sort of device. These Cariani calls *adaptive computational devices*. The most adaptive devices, he calls *structurally adaptive* to refer to the fact that they are capable of not only creating new mappings between a fixed set of primitives, but capable of creating *new* primitives. In the biological world there are many examples of this happening as in the evolution of colour vision, or flight or development of the cerebral cortex which have led to new sensory, effector and control mechanisms respectively (Bird et al. (2002)).

As Bird et al point out, Cariani's taxonomy is closely linked to different concepts of emergence. Adaptive computational devices achieve a *combinatorial emergence* as they can generate new combinations of existing primitives. Under the alphabet analogy, they can create new words. Structurally adaptive devices however, are capable of not only forming novel combinations of existing letters, but can create *new* letters: they are *creatively emergent*.

Bird et al describe two physical systems that are capable of such feats: Gordon Pask's Electrochemical Ear (Pask (1958)) and Paul Layzell's evolveable hardware (Layzell (2001)). Such systems undoubtedly broaden the space of possibilities beyond the confines of Latham's geometric primitives or KOANs pre-programmed musical fragments. Indeed as Bird et al. (2002) suggest, they may hold promise of satiating the Alife art desire for the generation of outcomes that "surpass our wildest imaginations". However this structural adaptivity could be a bit of a problem if we want to use such systems as creative tools. Firstly in terms of the aesthetic relevance of the outcomes, and secondly in terms of their usability. As I have suggested elsewhere (Eldridge (2005)), the incumbent epistemic autonomy in creatively emergent systems implies an aesthetic autonomy, i.e. it creates its own aesthetic norms. If we are concerned with creating artefacts for human consumption this may not be an attractive property. Woolf (2004) has made a similar point with respect to creative emergence in general, extending the alphabet metaphor, he questions how exciting it would actually be to be confronted with a novel written with a new alphabet ...

Secondly there is a problem if we want to use such a device as a 'creative tool'. As mentioned above, even the simple internal reconfigurations of the homeostat keep you on your toes as a performer: playing with the system under certain settings for a while, you can come to some form of performative understanding of its behaviours and thus learn to collaborate with it, but its unpredictabilities can never be fully fathomed. A structurally adaptive system would not only make slightly different responses to a certain stimuli (in this case a collision of objects in the visual display), it could at any moment respond to any other stimuli and offer an entirely new class of response. This would make working with the system quite difficult and render it 'unmasterable'. Finally, although these physical systems are arguably capable of exhibiting structural adaptation, it is a contentious and undecided issue whether a purely computational process can generate novel primitives (Boden (1996)).

Both Dorin's and Bird et al's suggestions address the problem of how the genotype and the enaction mechanism (shown in Figure 8.10) can be specified, yet unconstrained. Viewed within this framework, the simple move made in the Self-karaoke Machines in requiring the user to provide samples opens up this process. The genotype is still specified, but the enaction mechanism demands collaboration from the user, who doesn't just interact with the phenotype (the end product) as in the the vast majority of interactive art, but defines the very material from which the phenotype is formed. This simple move brings the human into the generative loop and immediately achieves a form of open-endedness which is unattainable in many purely digital, and even mechanical physical, systems.

If we return to Simon's parable of the ant on the beach mentioned in Chapter 2, we can understand Dorin's and Bird's concerns as addressing the problem of how to design an ant that can exhibit an unlimited range of behaviours as it walks across the same beach. Structurally adaptive systems engender creatively emergent ants. This is necessary if the beach is made up of digitally defined pebbles as in Latham's Mutator and the vast majority of digital generative art. But if we are concerned with the *behaviour* of the ant having unlimited potential the other alternative is to stick with a computationally adaptive, or even reactive, ant and let the user define aspects of the beach.

Chapter 9

Recapitulation and Future Development

This thesis aimed to demonstrate the appeal and application of simple adaptive systems as mechanisms for composing and improvising with computers. The work here is inspired by the principles and practices of Alife and autonomous robotics research and suggests a novel approach to the design and implementation of computer music software which complements mainstream AI-inspired approaches.

The choice of models stemmed from a consideration of the interactive and generative potential of existing creative applications of Alife models tempered by the requirements of the musician. The main projects centred around a model of an Ashbian homeostat. Aside from being somewhat of an Iconic device, the model implemented here is one of the simplest that can be seen to exhibit goal-directed and self-modulated adaptive dynamics, achieving what we might describe as minimally autonomous behaviour. The projects presented here in no way 'prove' this, but it is felt that these characteristics imbue a certain liveliness and coherence in both its dynamical output and its response to external events which come across when the streams of floating point numbers are used to specify sonic events. These attributes support an aesthetic which lies at the meeting point of digital and acoustic idioms to produce musical outputs which are "weird and surprising but strangely familiar".

Whether or not this claim is acceded to or this aesthetic is appealing to the reader, these types of adaptive dynamical systems seem to answer the callings of those wishing to indulge in live generative performance whilst retaining some modicum of respect, i.e. to achieve a balance of adaptability and dependability, of inspiration and obedience. The homeostat for example offers constant variation within a known behavioural field: it will invariably change dramatically when you poke it hard, and deviate only temporally when you prod it lightly. But what it will change to, or how long it will take to return are unknowns. Similarly you cannot control its behaviour directly, but the viscosity variable offers a form of global-control which allows you to reign it in or stoke it up as necessary. This is not meant to imply that the homeostat itself is the future of computer music, but it is a convenient device with which to illustrate a move toward a collaborative approach to algorithmic composition and performance.

The basic behaviours of the homeostat are indeed quite limited and there are many other possibilities yet to be explored. The relative simplicity of the device was another of its attractions as a starting point. Unlike many complex dynamical systems it can be readily parameterised by hand. One of the problems with the neural oscillator networks as implemented here was that the parameter space is so vast that whilst it is possible to discover some parameter sets which generate dynamics of musical use this was a slightly arduous process which relied largely on serendipity. The obvious next move in terms of systems design then is to explore more powerful systems and employ some form of artificial evolution to aid parameter search.

Hand designing complex systems in some ways goes against the principles of Alife. The observation that complex environments play a significant role in the creation of apparently complex behaviour leads to the use of artificial evolution for engineering the agent-side mechanisms necessary to deal with the environment. An agent can be given eyes and legs and told to follow walls, and artificial evolution will duly contrive the requisite sensori-motor connections. But as noted in Chapter 2, framing musical attributes in a formal fitness function is hard: 'play something interesting and relevant' is a decidedly ill-posed problem. The common alternative, interactive evolution, where a formal fitness function is replaced by a subjective choice on the part of the listener, is similarly fraught with difficulties, leading to a body of research on the 'fitness bottleneck'. Never-the-less, EC remains an attractive search method, particularly for setting parameter values for the sorts of complex systems explored here. The neural oscillator or broader CTRNNs models are full of musical promise, but exploring their parameter space by hand as done here, we can only get a peek at their full behavioural repertoire. An exciting possibility is that having engineered a minimally autonomous system, its response characteristics *could* be privy to formalisation. Rather than having to specify minutiae of musical details, evolution could be used to design fields of behaviour.

To develop these ideas further, it is also necessary to go beyond personal implementations and test the utility of these tools by getting other musicians to play with them. These needs are met in the form of a burgeoning collaborative project, *Behavioural Objects*. Working with computer scientist and musician, Ollie Bown and composer and improviser, Sebastien Lexer both of Goldsmiths College, UK, the project builds upon many of the ideas presented here and aims to present a compendium of adaptive dynamical objects for use in interactive computer music performance systems. Bown has recently built a CTRNN and has been using it in performance with trumpeter Tom Arthurs. Currently he has been using very simple fitness functions to ensure for example that outputs remain active. Working at this level, it seems entirely possible to be able to engineer simple relations with a performer. For example if pitch and/or onset analyses were fed into the network as inputs, it seems feasible that one could easily evolve a system to respond differentially to the performer using fitness functions such as 'play when I play', or 'remain active when I stop'. Such investigations are planned for the near future.

The project has also provided the opportunity to organise a focus group to explore other computer musician's and improviser's responses to these models. We plan to demonstrate ready made implementations of full interactive systems such as the basic Self-karaoke Machine, but more importantly to hand over the models in their raw form as Max/MSP objects and encourage other's experimentation. Feedback from this focus group will feed into planned future research which aims to move beyond the use of handcrafted behaviours and investigate methods for training and evolving behaviours.

At the start of this thesis were a list of reasons why it was an exciting time to be a musician. The last one of these is that it is now common practice to share ideas in the form of lines of code. At the end of the last chapter it was suggested that one answer to Dorin's concerns over the creative limitations of some generative systems for non-programmer users was to introduce the performer – and thus the vagaries of the real-world – into the enaction process, as in the Self-karaoke Machines. Programmes like Max/MSP and the social structures of its community of users mean that we no longer need to specify any form of enaction mechanism to share a generative idea. We can just post an ant and people can make their own eyes and legs and design their own beaches. Generative mechanisms are not just a new way of writing and performing music, but a new way of sharing and expressing musical ideas: a new musical currency.

Bibliography

Ames, C. and Domino, M. (1992). Cybernetic composer: an overview. In *Understanding Music with AI*, pages 186–205. AAAI Press.

Anderson, C. (2005). Dynamic networks of sonic interactions: an interview with Agostino Di Scipio. *Computer Music Journal*, 29(3):11–28.

Ariza, C. (2005). Navigating the landscape of computer-aided algorithmic composition systems: A definition, seven descriptors, and a lexicon of systems and research. In *Proceedings of the International Computer Music Conference*, pages 765–772, San Francisco. International Computer Music Association.

Arneodo, A., Coullet, P., and Tresser, C. (1980). Occurrence of strange attractors in three-dimensional Volterra equations. *Phys. Lett.*, 79A:259–263.

Ashby, W. R. (1952). *Design for a Brain: The Origin of Adaptive Behavior*. Chapman and Hall, London.

Bailey, K. (1983). Webern's opus 21: creativity in tradition. *Journal of Musicology*, 2:184–195.

Barandiaran, X. (2004). Behavioral adaptive autonomy. a milestone in the Alife route to AI ? In *Proceedings of the 9th International Conference on Artificial Life*, pages 514–521, Boston, Massachussets. MIT Press.

Barot, T. (1999). Songbirds forget their tunes in a cacophony of road noise. The Sunday Times, January 10th.

Barrass, S. and Kramer, G. (1999). Using sonification. *Multimedia Systems*, 7(1):23–31.

Bedau, M., McCaskill, J., Packard, N., Rasmussen, S., Adami, C., Green, D., Ikegami, T., Kaneko, K., and Ray, T. (2000). Open problems in artificial life. *Artificial Life*, 6:363–376.

Beer, R. (1995). A dynamical systems perspective on autonomous systems. *Artificial Intelligence*, 72:173–215.

Bentley, P. and Corne, D., editors (2001). *Creative Evolutionary Systems*. Morgan Kaufmann.

Berg, P. (1987). Foundations of computer music. In Roads, C. and Strawn, J., editors, *PILE – a language for sound synthesis.*, pages 64–108. The MIT Press., Cambridge, MA.

Berry, R. (1986). Percy's organ. `http://www.abc.net.au/arts/adlib/stories/s885079.htm15`.

Bidlack, R. (1992). Chaotic systems as simple (but complex) compositional algorithms. *Computer Music Journal*, 16(3):33–47.

Biles, A. (1994). GenJam: A genetic algorithm for generating jazz solos. In *Proceedings of the 1994 International Computer Music Festival*, pages 131–137, SanFransisco. ICMA.

Biles, A. (2002). GenJam in transition: from genetic jammer to generative jammer. In *Proceedings of Generative Art*.

Bilotta, E. and Pantano, P. (2002). Synthetic harmonies: An approach to musical semiosis by means of cellular automata. *Leonardo*, 35(2):153–159.

Bird, J., Layzell, P., Webster, A., and Husbands, P. (2002). Towards epistemically autonomous robots: Exploiting the artistic potential of physical systems. *Leonardo*, 36(2):109–114.

Bird, J. and Webster, A. (2001). The blurring of art and alife. In Dorin, A., editor, *Proceedings of Second Iteration*, pages 38 – 46, CEMA, Melbourne.

Birmbaumer, N., Lutzenberger, W., Rau, H., Braun, C., and Kress, G. M. (1996). Perception of music and dimensional complexity of brain activity. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, 6(2):267–278.

Blackwell, T. (2003). Swarm music: Improvised music with multi-swarms. In *Procedings of the AISB Symposium on Artificial Intelligence and Creativity in Arts and Science*, pages 41–49.

Blackwell, T., Y. M. (2004). Swarm granulator. In *EvoWorkshops Proceedings*, volume 3005 of *Lecture Notes in Computer Science*. Springer.

Blackwell, T., Y. M. (2006). Live algorithms for music manifesto. available from `http://www.timblackwell.com`.

Bly, S. A. (1982). Communicating with sound. In *Proceedings of the Conference on Human Factors in Computer Systems*, pages 371–375, New York. ACM Press/Addison-Wesley.

Boden, M. (1996). Introduction. In Boden, M., editor, *The Philosophy of Artificial Life*. Oxford University Press, Oxford, U.K.

Bongers, B. (2006). *Interactivation: Towards an e-cology of people, our technological environment, and the arts*. PhD thesis, Vrije Universiteit, Amsterdam, Holland.

Bown, O. and Wiggins, G. (2005). Modelling musical behaviour in a cultural-evolutionary system. In *Proceedings of the Computational Creativity Workshop, IJCAI*, Edinburgh.

Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. The MIT Press, Cambridge, MA.

Brewster, S. A. (2002). Overcoming the lack of screen space on mobile computers. Tech. Rep. TR-2001-87, Department of Computing Science, University of Glasgow.

Brooks, F. J., Hopkins, A. J., Neumann, P., and Wright, W. (1993). An experiment in musical composition. In Schwanauer, A. and Levitt, P., editors, *Machine Models of Music*, pages 23–40. MIT Press.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence Journal*, 47:139—160.

Brown, P. (2005). Creating mind. In *Proceedings Altered States Conference*, University of Plymouth, UK. Planetary Collegium.

Brown, S., Merker, B., and Wallin, N. (2000). An introduction to evolutionary musicology. In Wallin, N., Merker, B., and Brown, S., editors, *Origins of Music*, chapter 1, pages 3–24. MIT Press, 2 edition.

Burns, K. (1994). The history and development of algorithms in music composition. PhD thesis, School of Music, Ball State University, Indiana, USA.

Burns, K. H. (1997). Algorithmic composition, a definition. `http://music.dartmouth.edu/~wowem/hardware/algorithmdefinition.html`.

Burraston, D., Edmonds, E., Livingstone, D., and Miranda, E. (2004). Cellular automata in MIDI based computer music. In *Proceedings of the 2004 International Computer Music Conference*.

Burton, D. and Vladimirova, T. (1999). Generation of musical sequences with genetic technique. *Computer Music Journal*, 23(4):59–73.

Cariani, P. (1992). Some epistemological implications of devices which construct their own sensors and effectors. In Varela, F. and Bourgine, P., editors, *Towards a Practice of Autonomous Systems*. MIT Press, Cambridge, Mass.

Chennamangalam, J. (2003). The games of life an introduction to cellular automata. *Pristine Visions: The Mahatma Gandhi University College of Engineering Annual 2002-2003*.

Chowning, J. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, 21:526–534.

Clifton, T. (1970). An application of Goethe's concept of Steigerung to the morphology of diminution. *Journal of music Theory*, 14:165–189.

Collins, N. (2003a). Algorithmic composition methods for breakbeat science. *ARiADA*, 3.

Collins, N. (2003b). Generative music and laptop performance. *Contemporary Music Review*, 22(4):67–79.

Collins, N. and Cross, I. (2003). Beat tracking and reaction time. In *Rhythm Perception and Performance Workshop*.

Cook, N. (1987a). Musical form and the listener. *Journal of Aesthetics and Art Criticism*, 46:23–29.

Cook, N. (1987b). Perception of large scale tonal structure. *Music Peception*, 5:197–205.

Cook, N. (1990). *Music, Imagination and Culture*. Clarendon, Oxford.

Cook, N. (1998). *Music a Very Short Introduction*. Oxford University Press, New York.

Cope, D. (1991). *Computers and Musical Style*. Oxford University Press, Oxford.

Cope, D. (1992). Computer modeling of musical intelligence in emi. *Computer Music*, 16(2):69–83.

Cope, D. (1999). One approach to musical intelligence. *IEEE Intelligent Systems*, 14(3):21–25.

Cope, D. (2000). *The Algorithmic Composer*. A-R Editions Inc, Wisconsin.

Cope, D. (2005). *Computational Models of Creativity*. MIT Press, Cambridge MA.

Dahlstedt, P. (2001). A MutaSynth in parameter space: Interactive composition through evolution. *Organised Sound*, 6(2):121–124.

Dahlstedt, P. and Nordahl, M. G. (2001). Living melodies: Coevolution of sonic communication. *Leonardo*, 34(3):243–248.

Dale, K. (2000). *When Worlds Collide: some simple models of navigation using vision and a compass.* PhD thesis, School of Biological Sciences, University of Sussex, UK.

Dannenberg, R. (1989.). Current directions in computer music research,. In Matthews, M. and Pierce, J. R., editors, *Real-time scheduling and computer accompaniment*, pages 225–61. MIT Press, Cambridge, MA.

Davis, K. D. (1996). Dystopic toys. *World Art*, 1:30–33.

Dawkins, R. (1986). *The Blind Watchmaker*. W. W. Norton & Company, New York.

de Vaucanson, J. ((1742).). *An account of the mechanism of an automaton or image playing on the German-flute.* Buren, Netherlands. trans. J.T. Desaguliers.

Dean, R. (2003). *Hyperimprovisation: Computer-Interactive Sound Improvisation.* A-R Editions Inc., Middleton, Wisconsin.

Demuth, N. (2003). *Musical Forms and textures*. Rockcliff Publishing Co. Ltd., London.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(5):429–452.

Di Scipio, A. (1990). Composition by exploration of nonlinear dynamical systems. In *Proceedings of the 1990 International Computer Music Conference*, pages 324–327, San Francisco. International Computer Music Association.

Di Scipio, A. (1999). Synthesis of environmental sound textures by iterated nonlinear functions. In *Proceedings of the 2nd COST g-6, Workshop on Digital Audio Effects DAFX'99*, Trondheim, Norway,.

Di Scipio, A. (2003). Sound is the interface: from interactive to ecosystemic signal processing. *Organised Sound*, 8(3):269–277.

Dinkla, S. (1994). The history of interfaces in interactive art. *Proceedings of the 1994 International Symposium on the Electronic Arts*.

Dodge, C. and Jerse, T. A. (1997). *Computer Music; Synthesis, Composition, and Performance*. Wadsworth Publishing Company.

Dombois, F. (2001). Using audification in planetary seismology. In Hiipakka, J., Zacharov, N., and Takala, T., editors, *Proceedings of ICAD 7*, pages 227–228. Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Espoo, Finland.

Dorin, A. (2001). Aesthetic fitness and artificial evolution for the selection of imagery from the mythical infinite library. In Kelemen, J. and Sosik, P., editors, *Lecture Notes in Artificial Intelligence. Advances in Artificial Life, Proceedings of the 6th European Conference (ECAL)*, pages 659–668, London, UK. Springer Verlag.

Dorin, A. and McCormack, J. (2001). First iteration, generative systems (guest editors introduction). *Leonardo*, 34(2):335.

Drew, D. G. and Grimes, T. (1987). Audio-visual redundancy and TV news recall. *Communication Research*, 14:452–461.

Eacott, J. (2000). Form and transience - generative music composition in practice. In Soddhu, C., editor, *Proceedings of Generative Art*.

Ebcioglu, K. (1988). An expert system for harmonizing four-part chorales. *Computer Music Journal*, 12(1):43–51.

Eldridge, A. (2002). Adaptive systems music. Master's thesis, School of Cognitive and Computing Sciences, University of Sussex.

Eldridge, A. (2005). Cyborg dancing: generative systems for man-machine musical improvisation. In Innocent, T., editor, *Proceedings of Third Iteration.*, pages 129–142, Melbourne, Australia. Centre for Electronic Media Art (CEMA).

Eno, B. (1996). Generative music (talk delivered at the imagination conference, San Francisco, June 8th 1996). *Motion Magazine, July*.

Epstein, J. (1996). *Growing Artificial Societies: social science from the bottom up*. Brookings Institution, Washington.

Findahl, R. (1981). The effect of visual illustrations upon perception and retention of news programmes. ERIC documentation reproduction services. no ED 054 631.

Fitch, T. and Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system. In Kramer (1994a), pages 307–326.

Forrest, S. and Jones, T. (1994). Modelling complex adaptive systems with ECHO. In Stonier, R. and Yu, X., editors, *Complex systems: Mechanisms of Adaptation*. IOS Press, Amsterdam.

Futayama, D. and Slatkin, M. (1983). *Coevolution*. Sinauer, Sunderland.

Galanter, P. (2003). What is generative art ? complexity theory as a context for art theory. In Soddhu, C., editor, *Proceedings of Generative Art*.

Gardner, M. (1970). Mathematical games the fantastic combinations of John Conway's new solitaire game "Life". *Scientific American*, 223:120–123.

Gardner, M. R. and Ashby, W. R. (1970). Connectance of large dynamic (cybernetic) systems: Critical value for stability. *Nature*, 228(5273).

Garnett, G. (2001). The aesthetics of interactive computer music. *Computer Music Journal*, 25(1):21–33.

Gartland-Jones, A. and Copley, P. (2005). Musical form and algorithmic solutions. In *Proceedings of the Creativity and Cognition Conference*, pages 226–231, New York. ACM.

Gaver, W. W. (1989). The sonic Finder: An interface that uses auditory icons. *Human Computer Interaction*, 4(1):67–94.

Gerhard, D. and Hepting, D. (2004). Cross-modal parametric composition. In *Proceedings of the International Computer Music Conference*, pages 505–512, San Francisco. International Computer Music Association.

Gogins, M. (1991). Iterated functions systems music. *Computer Music Journal*, 15(1):40–48.

Grout, D. and Palisca, C. (1996). *A History of Western Music*. W.W. Norton and Company, New York, 5 edition.

Gurney, K. (1997). *An Introduction to Neural Networks*. Routledge, London, UK.

Häfner, R. and Krätz, O. (1978). Der makabre android ... sein spiegelbild in der literatur. *Kultur und Technik*, pages 22 – 27.

Handel, S. (1989). *Listening*. MIT Press, MA.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall.

Hayward, C. (1994). Listening to the earth sing. In Kramer (1994a), pages 369–404.

Hiller, L. (1981). Composing with computers: A progress report. *Computer Music Journal*, 5(4).

Hiller, L. and Isaacson, A. (1959). *Experimental Music,*. McGraw-Hill, New York.

Hillis, W. (1992). Co-evolving parasites improve simulated evolution as an optimisation procedure. In Langton, C., Taylor, C., Taylor, J., and Rasmussen, S., editors, *Artificial Life II*, pages 313–324. Addison-Wesley, Reading, MA.

Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.

Hornel, D. and Dagenhardt, P. (1997). A neural organist improvising baroque-style melodic variations. In *Proceedings of the International Computer Music Conference*, Thesoloniki, Greece.

Horner, A. and Goldberg, D. (1991). Genetic algorithms and computer-assisted music composition. In *Proceedings of the 1991 International Computer Music Conference*, pages 479–482, San Fransisco. International Computer Music Association.

Impett, J. (2000). Situating the invention in interactive music. *Organised Sound*, 5(1):27–34.

Impett, J. (2001). *Computational Models for Interactive Composition/Performance Systems*. PhD thesis, University of Cambridge.

Jacob, B. (1996). Algorithmic composition as a model of creativity. *Organised Sound*, 1(3):157–165.

Johnsson-Laird, P. (1991). Jazz improvisation: A theory at the computational level. In Howell, P. West, R. and Cross, I., editors, *Representing Musical Structure*. Academic Press.

Jones, K. (1981). Compositional applications of stochastic processes. *Computer Music Journal*, 5:41–61.

Jorda, S. (2002). Improvising with computers: A personal survey. *Journal of New Music Research*, 31(1):1–10.

Jost, E. (1994). *Free Jazz*. 1^{st} Da Capo Press edition. Da Capo Press, New York.

Kay, A. and Goldberg, A. (1977). Personal dynamic media. *IEEE Computer*, pages 31–41.

Kelso, J. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior.* The MIT Press/A Bradford Book.

Kennel, A. (1996). AudioGraph: A diagram reader for blind people. In *Proceedings of the Second Annual ACM Conference on Assistive Technologies*, pages 51–56. ACM Press, New York, NY.

Kitagawa, N. and Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, 416:172–174.

Kotosaka, S. and Schaal, S. (2001). Synchronized robot drumming by neural oscillator. *Journal of the Robotics Society of Japan*, 19(1):116–123.

Koza, J. (1993). *Genetic Programming.* MIT Press/ Bradford Books, Cambridge, MA.

Kramer, G., editor (1994a). *Auditory Display: Sonfication, Audification and Auditory Interfaces, Proceedings of the First International Conference on Auditory Display*. Addison-Wesley, Reading, MA.

Kramer, G. (1994b). Some organising principles for representing data with sound. In Kramer (1994a), pages 85–221.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., and Neuhoff, J. (1997). Sonification report: Status of the field and research agenda. Technical report, National Science Foundation.

Krause, B. (1993). The niche hypothesis, a hidden symphony of animal sounds, the origins of musical expression and the health of habitats. *The Explorers journal*, Winter:156–160.

Krueger, M. (1976). *Computer Controlled Responsive Environments.* PhD thesis, University of Wisconsin, Madison.

Krumhansl, C. (1990). Melodic structure: Theoretical and perceptual aspects. Paper presented at the International Wenner-Gren Symposium: Music, Language, Speech and Brain.

Lachman, R., Lachman, J., and Butterfield, E. (1979). *Cognitive psychology and information processing: An introduction.* Lawrence Erlbaum, New Jersey.

Langton, C. (1989). Artificial life. In Langton, C., editor, *Artificial Life. Santa Fe Institute Studies in the Sciences of Complexity*, pages 2–3. Addison-Wesley.

Layzell, P. (2001). *Hardware evolution: On the Nature of artificially evolved electronic circuits*. PhD thesis, School of Cognitive and Computing Sciences, University of Sussex, UK.

Leach, J. and Fitch, J. (1995). Nature, music, and algorithmic composition. *Computer Music Journal*, 19(2):23–33.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music.* MIT Press, Cambridge, MA.

Levin, G. (1994). *Painterly Interfaces for Audiovisual Performance*. PhD thesis, Massachusetts Institute of Technology.

Lewis, G. (1999). Interacting with latter-day musical automata. *Contemporary Music Review*, 18(3):99–112.

Lewis, G. (2006). Improvising with creative machines. Talk given at Improvising With Computers. pre-NIME workshops. IRCAM, Paris, France.

Lewis, M. and Granic, I., editors (2000). *Emotion, Development, and Self-Organization. Dynamic Systems Approaches to Emotional Development.* Cambridge University Press, Cambridge and New York.

Lotka, A. J. (1925). *Elements of physical biology.* Williams and Wilkins Co., Baltimore.

Loy, D. (1989). Composing with computers - a survey of some compositional formalisms and music programming languages. In Matthews, M. and Pierce, J., editors, *Current Directions in Computer Music Research*, pages 291–396. MIT Press, Cambridge, MA.

Loy, D. (1991). Connectionism and musiconomy. In Todd, P. and Loy, D., editors, *Music and Connectionism*, pages 20–36. MIT Press, Cambridge, MA.

Lunney, D. and Morrison, R. (1990). High technologies for visually handicapped chemistry students. *Journal of Chemistry Education*, 58:228.

Machover, T. (1992). *Cyberarts*, chapter Hyperinstruments: A Composer's Approach to the Evolution of Intelligent Musical Instruments, pages 67–76. William Freeman.

Machover, T. and Chung, J. (1994). Hyperinstruments: Musically intelligent and interactive performance and creativity systems. *Proceedings International Computer Music Conference*.

Maes, P., editor (1991). *Designing Autonomous Agents.* MIT Press. Boston, MA.

Manovich, L. (2002). *The Language of New Media*. MIT, London.

Matsuoko, K. (1985). Sustained oscillations generated by mutually inhibiting neurons with adaptation. *Biological Cybernetics*, 52:367:376.

Maurer, J. (1999). The influence of chaos on computer-generated music. `http://ccrma.stanford.edu/~blackrse/chaos.html`.

McCabe, K. and Rangwalla, A. (1994). Auditory display of computational fluid dynamics data. In Kramer (1994a), pages 321–340.

McCormack, J. (2001). Eden: an evolutionary sonic ecosystem. In Kelemen, J. and Sosik, P., editors, *Lecture Notes in Artificial Intelligence. Advances in Artificial Life, Proceedings of the 6th European Conference (ECAL)*, volume 2159, pages 133–142, Berlin, Germany. Springer-Verlag.

McCormack, J. (2004). *Impossible Nature: The Art of Jon McCormack.* ACMI.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588).

Melara, R. D. and Marks, L. E. (1990). Interaction among auditory dimensions: Timbre pitch and loudness. *Perception and Psychophysics*, 482:169–178.

Melo, A. (1988). A connectionist model of tension in chord progressions. Master's thesis, School of Artificial Intelligence, University of Edinburgh, Edinburgh.

Miranda, E. (1993). Cellular automata music: An interdisciplinary music project. *Interface*, 22(1):3–21.

Miranda, E. (1999). Towards an Artificial Life approach to the origins of music. In *Proceedings of the XIX Annual Congress of the Brazilian Computing Society / VI Brazilian Symposium on Computer Music*, Rio de Janeiro. Brazilian Computing Society.

Miranda, E. (2000a). *Composing Music With Computers*. Focal Press, Burlington.

Miranda, E. (2000b). On the music of emergent behaviour: what can evolutionary computation bring to the musician? In *Proceedings of Genetic and Evolutionary Computation Conference*, Las Vegas, USA. University of Central Florida.

Miranda, E. (2002). Emergent sound repertoires in virtual societies. *Computer music Journal*, 26(2):77–90.

Miranda, E. (2003). On making music with Artificial Life models. In *Consciousness Reframed 2003*.

Miranda, E. and Biles, A., editors (2007). *Evolutionary Computer Music*. Springer-Verlag, London.

Montag, D. (2000). *Bioglyphs: Generating images in collaboration with nature's events*. PhD thesis, University of Hertfordshire, UK.

Mozer, M. (1994). Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multi-scaling processing. *Connection Science*, 6(2-3):247–280.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395:123–124.

Neuhoff, J. G., Kramer, G., and Wayand, J. (2000). Sonification and the interaction of perceptual dimensions: Can the data get lost in the map? *Journal of Experimental Psychology: Applied.*, 8(1):17–25.

Neuhoff, J. G. and Wayand, J. (2002). Pitch change, sonification and musical expertise: which way is up? In Nakatsu, R. and Kawahara, H., editors, *Proceedings of ICAD 8*, pages 351–356. ATR, Kyoto, Japan.

Newell, A. and Simon, H. A. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126.

Nyman, M. (1999). *Experimental Music: Cage and Beyond (2nd Ed.)*. Cambridge University Press.

Orio, N., Leemouton, S., and Schwarz, D. (2003). Score following: State of the art and new developments. In *Proceeedings of NIME*, pages 36–41, Montreal, Canada.

Pachet, F. (2002). Interacting with a musical system: The continuator. In Anagnostopoulou, C. Ferrand, M. and Smaill, A., editors, *Music and Artificial Intelligence: Lecture Notes in Artificial Intelligence*, volume 2445, pages 119–132. Springer- Verlag,.

Paine, G. (2002). Interactivity, where to from here ? *Organised Sound*, 7(3):295–304.

Papadopolous, G. and Wiggins, G. (1998). A genetic algorithm for the generation of jazz melodies. In *STeP'98*, Jyvaskyla, Finland.

Papadopolous, G. and Wiggins, G. (1999). AI methods for algorithmic composition: a survey, a critical review and future prospects. In *Proceedings of the AISB '99 Symposium on Musical Creativity*, pages 110–117, Brighton, UK. SSAISB.

Paradiso, J. Hsiao, K. and Hu, E. (1999). Interactive music for instrumented dancing shoes. In *Proceedings of the International Computer Music Conference*, Bejing.

Pareles, J. (1990). The MIDI menace: Machine perfection is far from perfect. *New York Times*.

Pask, G. (1958). Physical analogues to the growth of a concept in mechanisation of thought processes. In *Proceedings of a Symposium held at the National Physcial Laboratory on 24-27 Nov*, London. H.M.S.O.

Pearce, M., Meredith, D., and Wiggins, G. (2002). Motivations and methodologies for automation of the compositional process. *Musicae Scientiae*.

Penny, S. (1995). In pursuit of living machines. *Scientific American*, page 216.

Pereverez, S. V., Loshak, A., Backhaus, S., Davis, J. C., and Packard, R. E. (1997). Quantum oscillations between two weakly coupled reservoirs of superfluid 3He. *Nature*, 388:449–451.

Petsche, H., Lindener, K., Rappelsberger, P., and Gruber, G. (1988). The EEg: An adequate method to concretise brain processes elicited by music. *Music Perception*, 6:133–160.

Phon-Amnuaisuk, S. Tuson, A. and Wiggins, G. (1999). Evolving musical harmonisation. In *Proceedings of ICANNGA*.

Pickering, A. (2002). Cybernetics and the mangle: Ashby, Beer and Pask. In Pestre, D. and Dahan, A., editors, *Social Studies of Science, La Science des Années 1950*, volume 32, pages 413–37. Presses de l'EHESS, Paris.

Poppel, K. (1994). Temporal mechanisms in perception. *International Review of Neurobiology*, 37:185–202.

Port, R. and Van Gelder, T., editors (1995). *Mind as Motion: Explorations in the Dynamics of Cognition.* The MIT Press/A Bradford Book., Cambridge, MA.

Pressing, J. (1988). Nonlinear maps as generators of musical design. *Computer Music Journal*, 12(2):35–45.

Prusinkiewicz, P. (1986). Score generation with L-systems. In *Proceedings of the 1986 International Computer Music Conference*.

Prusinkiewicz, P. and Lindenmayer, A. (1990). *The Algorithmic Beauty of Plants*. Springer:Verlag, New York.

Raphael, C. (1999). Automatic segmentation of acoustic musical signals using hidden markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 21(4):360–370.

Raphael, C. (2004). Music minus one - what is it ? `http://xavier.informatics.indiana.edu/~craphael/music_plus_one/mpo.html`.

Ray, T. S. (1991). An approach to the synthesis of life. In Langton, C., Taylor, C., Farmer, J. D., and Rasmussen, S., editors, *Artificial Life II, Santa Fe Institute Studies in the Sciences of Complexity*, volume XI, pages 371–408. Addison-Wesley,, Redwood City, CA.

Reynolds, C. W. (1987). Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics (SIGGRAPH '87 Conference Proceedings)*, 21(4):25–34.

Rinaldo, K. (2000). The flock. `http://accad.osu.edu/~rinaldo/works/flock/flock.html`.

Rinaldo, K. E. (1998). Technology recapitulates phylogeny: Artificial Life art. *Leonardo*, 31(5):371–376.

Roads, C. (1985). Research in music and artificial intelligence. *ACM Computing Surveys*, 17(2):163–190.

Roads, C. (1996). *The Computer Music Tutorial.* MIT Press, Cambridge, MA.

Roads, C. (2001). *Microsound*. MIT Press, Cambridge, MA.

Robertson, A. and Plumbley, M. D. (2006). Real-time interactive musical systems: An overview. In *Proceedings of the Digital Music Research Network Doctoral Researchers Conference*, London, UK.

Rokeby, D. (1990). A very nervous system. `http://homepage.mac.com/davidrokeby/vns.html`.

Rowe, R. (1992). *Interactive Music Systems: Machine Listening and Composing*. The MIT Press.

Rowe, R. (2001). *Machine Musicianship*. MIT Press, Cambridge, MA.

Ruiz-Mirazo, K. and Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life 10*, pages 235–259.

Schnell, N. and Battier, M. (2002). Introducing composed instruments, technical and musicological implications. In *Proceedings of NIME*.

Schoenberg, A Stein, L. T. (1954). *Structural Functions of Harmony*. Norton, New York.

Schutz, J. (1976). Fragments of the phenomenology of music. In Smith, J., editor, *In Search of Musical Method*, pages 23–71. Gordon and Breach Science Publishers, London, New York and Paris.

Sekuler, R., Sekuler, A. B., and Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385:308–308.

Severin, W. J. (1967). Another look at cue summation. *Communication Review*, 15:233–245.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature*, 408:788.

Shepard, R. (1982). Structural representations of musical pitch. In Deutsch, D., editor, *The Psychology of Music*, pages 343–390. Academic Press, New York.

Simon, H. (1969). *The Sciences of the Artificial*. MIT Press.

Sims, K. (1991). Artificial evolution for computer graphics. In *Computer Graphics (Siggraph '91 proceedings)*, pages 319–328.

Smith, R. and Bedau, M. A. (2000). Is Echo a complex adaptive system? *Evolutionary Computation*, 8:419–442.

Smithers, T. (1997). Autonomy in robots and other agents. *Brain and Cognition*, 34:88—106.

Sommerer, C. and Mignonneau, L. (1997). A-Volve - an evolutionary Artificial Life environment. *Artificial Life V*, pages 167–175.

Sommerer, C. and Mignonneau, L., editors (1998). *Art & Science.* Springer-Verlag., Wien.

Spector, L. and Alpern, A. (1995). Induction and recapitulation of deep musical structure. In *Working Notes of the IJCAI-95 Workshop on Artificial Intelligence and Music*, pages 41–48.

Spiegel, L. (1989). Distinguishing random, algorithmic, and intelligent music. Internet: `http://retiary.org/ls/writings/alg_comp_ltr_to_cem.html`.

Steedman, M. (1984). A generative grammar for jazz chord sequences. *Music Perception*, 2:55–77.

Steedman, M. (1989). *Lexical Representation and Process*, chapter Grammar, Interpretation and Processing from the Lexicon. MIT Press.

Stockhausen, K. (1964). *Texte zu eigenen Werken, zur Kunst Anderer*, chapter Elektronishce Studien I und II. DuMont Schauberg, Cologne.

Sundberg, J. and Lindblom, B. (1976). Generative theories in language and music descriptions. *Cognition*, 4:99–122.

Supper, M. (2001). A few remarks on algorithmic composition. *Computer Music Journal*, 25(1):48–53.

Taube, H. (2004). *Notes from the Metalevel: An Introduction to Computer Composition.* Swets Zeitlinger Publishing.

Thaut, M. (2003). Neural basis of rhythmic timing networks in the human brain. *Annals of the New York Academy of Sciences*, 999:364–373.

Thelen, E. and Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action.* The MIT Press., Cambridge, MA.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind.* Harvard University Press.

Toch, E. (1948). *The Shaping Forces of Music.* Dover Publications, New York.

Todd, P. (1900). A sequential network design for musical applications. In Touretsky, D., Hinton, G., and Sejnowski, T., editors, *Proceedings of the 1988 Connectionist Models Summer School*, pages 76–84, San Mateo. Morgan Kaufmann.

Todd, P. (1989). A connectionist approach to algorithmic composition. *Computer Music Journal*, 4(13):27–43.

Todd, P. and Loy, D. (1991). *Music and Connectionism.* MIT Press, Cambridge MA.

Todd, P. and Miller, G. (1991). On the sympatric origin of species: Mercurial mating in the quicksilver model. In Belew, R. and Booker, L., editors, *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 547–554, San Mateo, MA. Morgan Kaufmann.

Todd, P. and Werner, G. (1999). Frankensteinian methods for evolutionary music composition. In Griffith, N. and Todd, P., editors, *Musical Networks: Parallel Distributed Perception and Performance*, pages 173–190. MIT Press, Cambridge, MA.

Todd, S. and Latham, W. (1991). Mutator, a subjective human interface for evolution of computer sculptures. Technical report, IBM, United Kingdom Scientific Center Report.

Tovainen, P. (1995). Modeling the target-note technique style of bebop-style jazz improvisation: an artificial neural network approach. *Music Perception*, 4:399–413.

Truax, B. (1990). Chaotic non-linear systems and digital synthesis: an exploratory study. In *Proceedings of the International Computer Music Conference*, pages 100–103, Glasgow, Scotland. International Computer Music Association.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21:615–665.

Van Gelder, T. (1999a). Dynamic approaches to cognition. In Wilson, R. and Keil, F., editors, *The MIT Encyclopedia of Cognitive Sciences.*, pages 244–246. The MIT Press., Cambridge, MA.

Varela, F. J. (1979). *Principles of Biological Autonomy.* Elsevier North Holland, New York.

Volterra, V. (1926.). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. R. Accad. Naz. dei Lincei*, VI(2).

Walker, B., Kramer, G., and Lane, D. (2000). Psychophysical scaling of sonification mappings. In Cook, P., editor, *Proceedings of ICAD 6*, pages 99–104. International Community for Auditory Display.

Walker, B. N. and Kramer, G. (1996). Mappings and metaphor in auditory displays: An experimental assessment. In Frysinger, S. P., Kramer, G., Ahearn, D., Pettit, C., and St. Amant, K., editors, *Proceedings of ICAD 3*, pages 71–74. Santa Fe Institute, Santa Fe, New Mexico.

Walter, W. G. (1953). *The Living Brain.* Duckworth, London.

Waschka II, R. (2001). *Artificial Life Models for Musical Applications*, chapter Theories of Evolutionary Algorithms and a 'New Simplicity' Opera: Making Sappho's Breath, pages 79–86. Editoriale Bios, Cosenza, Italy.

Werner, G. and Todd, P. (1997). Too many love songs: Sexual selection and the evolution of communication. In Husbands, P. and Harvey, I., editors, *Fourth European Conference on Artificial Life*, pages 434–443. MIT Press/Bradford Books, Cambridge, MA:.

White, T. (1983). A man out of time beats the clock. *Musician magazine*, 60.

Whitelaw, M. (2004). *Metacreation: Art and Artificial Life.* MIT Press, Cambridge, MA.

Whitelaw, M. (2005). System stories and model worlds: A critical approach to generative art. In *Readme 100: Temporary Software Art Factory*, pages 135–154, Norderstedt. BoD.

Wiener, N. (1948). *Cybernetics or Control and Communication in the Animal and the Machine.* MIT Press, Cambridge: MA.

Wiener, N. (1967). *The Human Use of Human Beings: Cybernetics and Society*. Avon Books, 2nd Ed., New York.

Williamson, M. (2002). Oscillators and crank-turning: Exploiting natural dynamics with a humanoid robot arm. Technical report, Hewlett Packard, 202-132.

Winkler, T. (2001). *Composing Interactive Music: Techniques and Ideas Using Max.* The MIT Press.

Wishart, T. (1985). *On Sonic Art.* Imagineering Press, York.

Wolfram, S. (1982). Cellular automata as simple self-organising systems. Caltech preprint CALT - 68-938.

Woolf, S. (2004). *Interactive and Generative Processes in the New Media Arts*. PhD thesis, School of Informatics, University of Sussex, UK.

Wuensche, A. (1997). Attractor basins of discrete networks. Cognitive Science Research Paper CSRP461, University of Sussex, School of Cognitive and Computing Sciences, University of Sussex, Brighton.

Xenakis, I. (1971a). *Cybernetics, art and ideas*, chapter Free stochastic Music, pages 124 – 142. Studio Vista ltd, London.

Xenakis, I. (1971b). *Formalised Music*. Indiana University Press.

Yee-King, M. (2000). Audioserver - collaborative, interactive genetic algorithm for sound design. Master's thesis, School of Cognitive and Computing Sciences, University of Sussex.

Young, D. (2001). The hyperbow controller: Real-time dynamics measurement of violin performance. In *Proceedings of NIME*.

Zicarelli, D. (1987). M and jam factory. *Computer Music Journal*, 11(4):13.

Appendix A

Supplementary Results for Chapter 5

A.1 Neural Oscillator

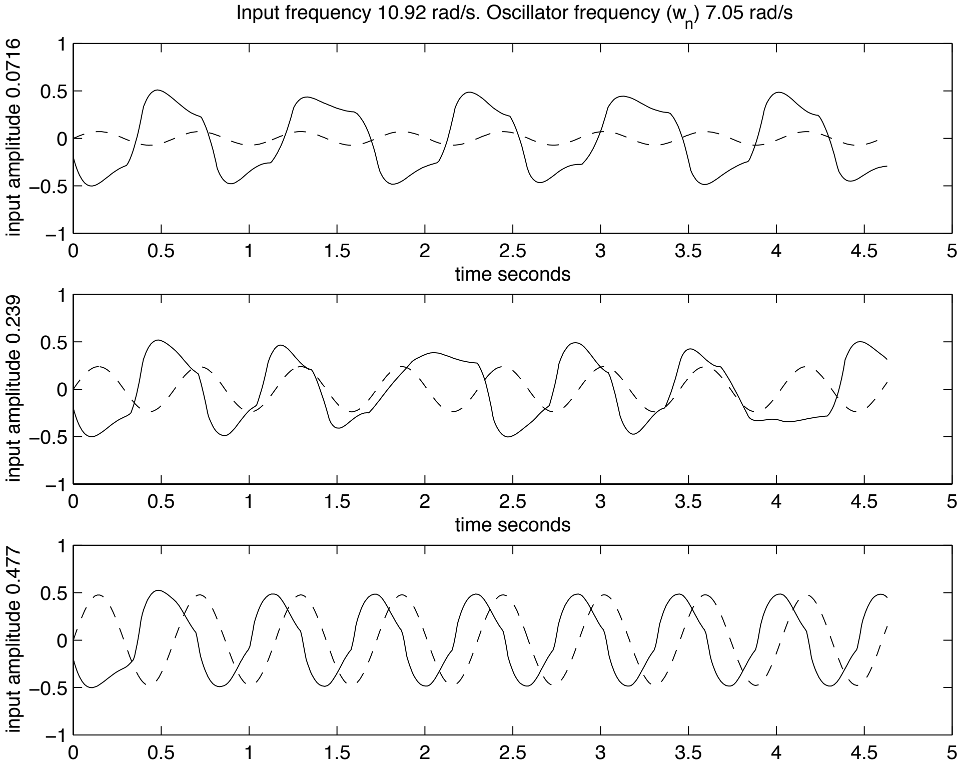


Figure A.1: Plots showing the output of a neural oscillator pair and sinusoidal input signal for varying amplitudes of input. (from Williamson(2002))

When the amplitude of the input signal is small (top), the oscillator is not entrained but oscillates at its endogenous frequency. As the input is increased the oscillator is almost entrained but slips every few cycles. For larger input amplitudes, the oscillator locks to the input frequency (bottom).

A.2 Self-regulating homeostat

The following results suggest that the modifications made to the basic homeostat did not disrput its essential characteristics. As shown in Figure A.2, the system returns to the same stable state after minor perturbation (marked **a** at iterations 250 and 525), and re-stabilises following critical perturbation (arked **A** at iteration 350). As we would expect, the increase in size means that the network takes longer to stabilise.

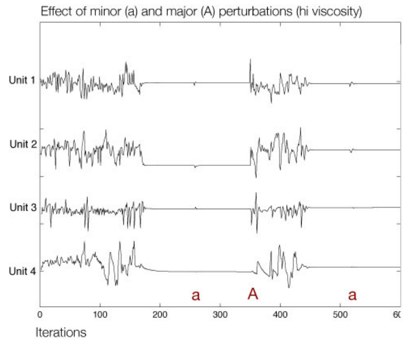


Figure A.2: Outputs of a ten-unit self-regulated homeostatic network demonstrating stability to minor perturbation and re-stability after critical interference.

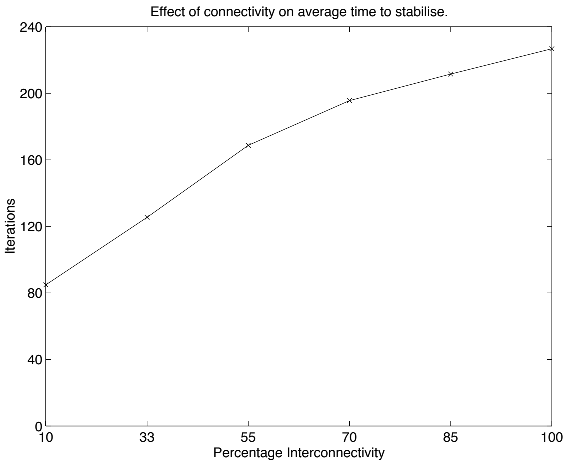


Figure A.3: Change in stability as a function of interconnectivity for an auto-regulated network

Figure A.3 shows the increase in the average time to stabilise with an increase in percentage connectivity for an auto-regulated network. The inverse relationship between connectivity and stability observed in the standard homeostat is preserved. Here connectivity refers to the degree of *interconnectivity*, each module being fully *intra*connected.

Appendix B

Technical set up for installations and performances

B.1 Set up for Ashby's Grandmother's Footsteps

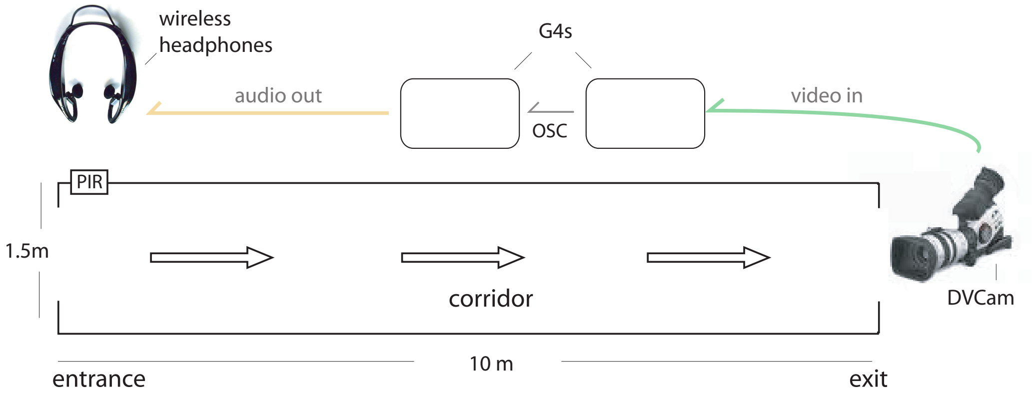


Figure B.1: Set up for Ashby's Grandmothers Footsteps installation

B.2 Technical Set up for Self-karaoke in installation and performance

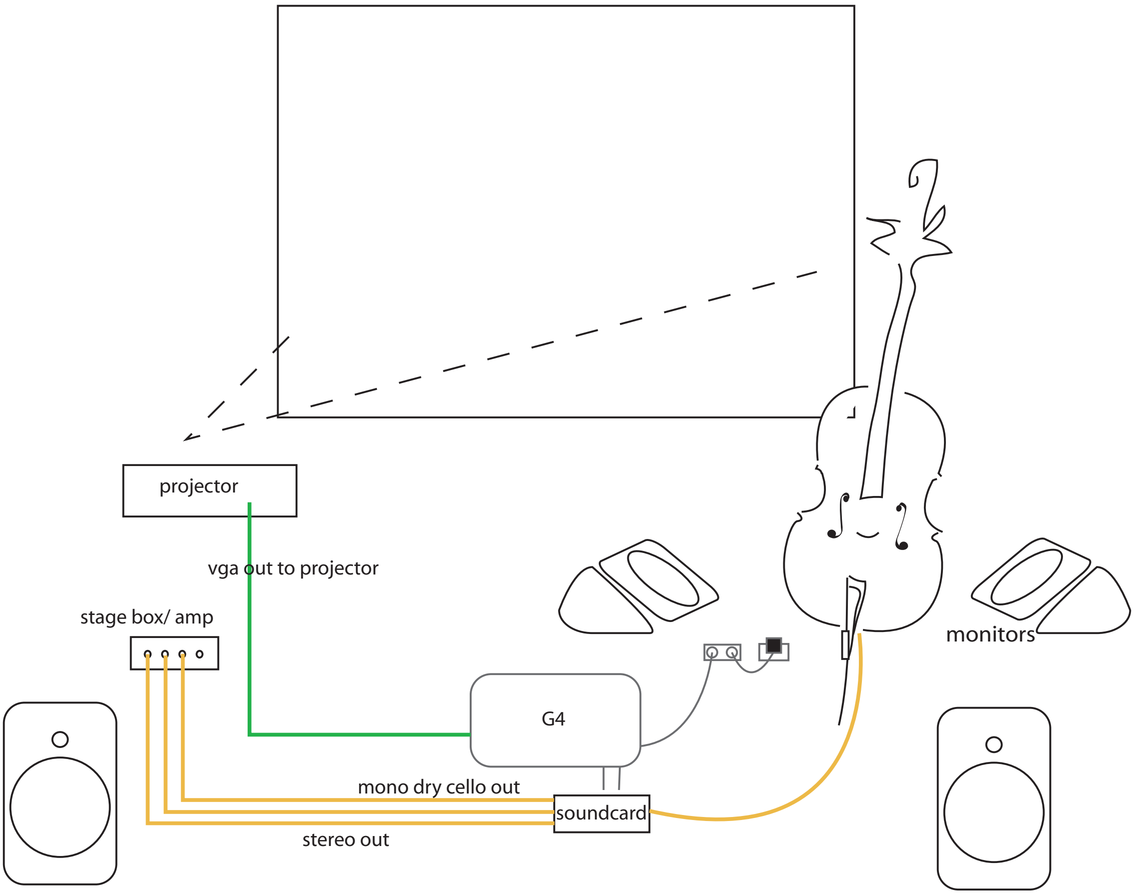


Figure B.2: Set up for Fond Punctions Performances

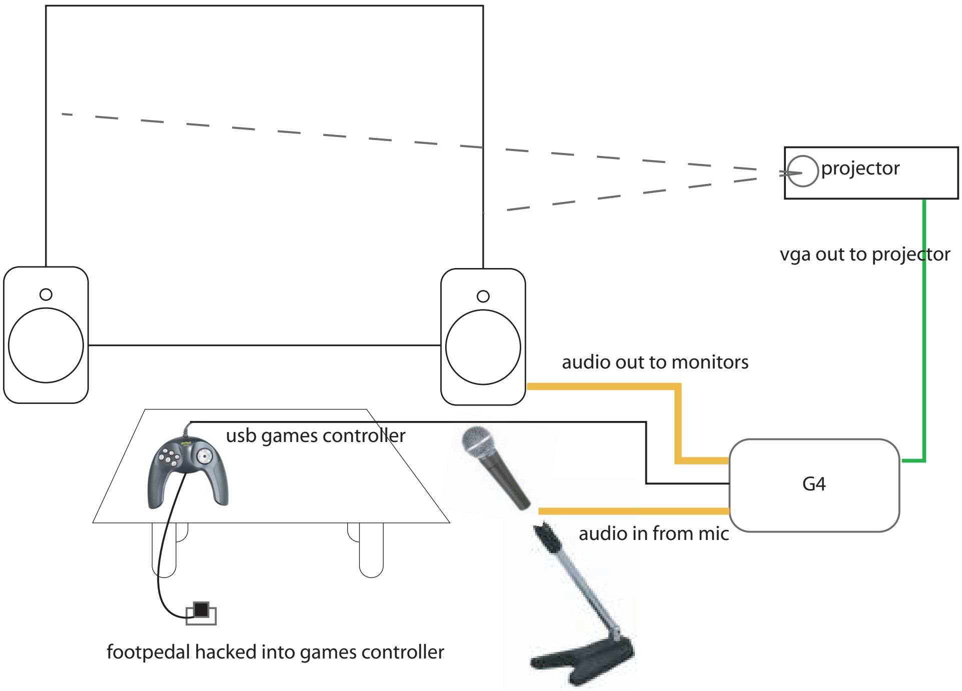


Figure B.3: Set up for Self-karaoke Pond in installation