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`

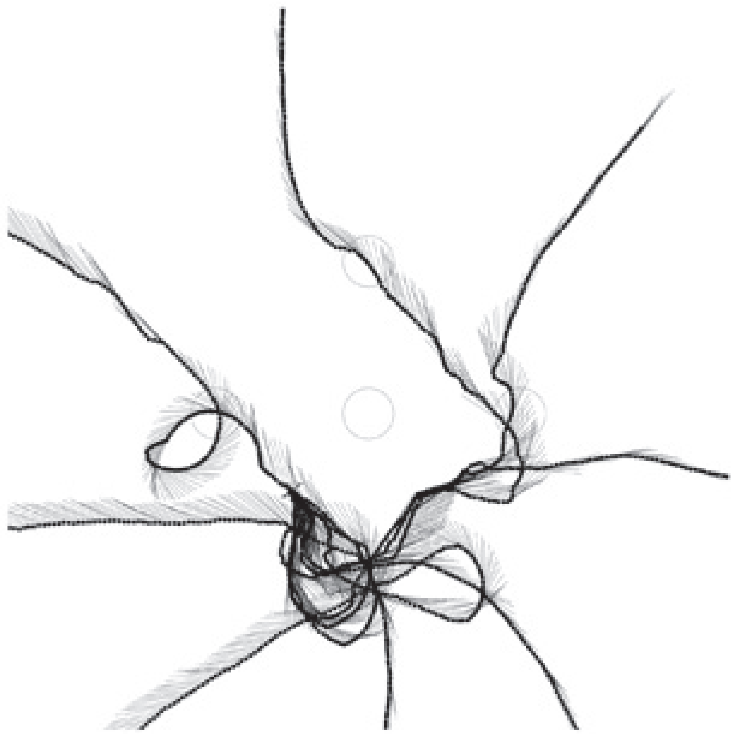

Figure 4.1: Visualisation of simulated foraging behaviour. Dale (2000).

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.

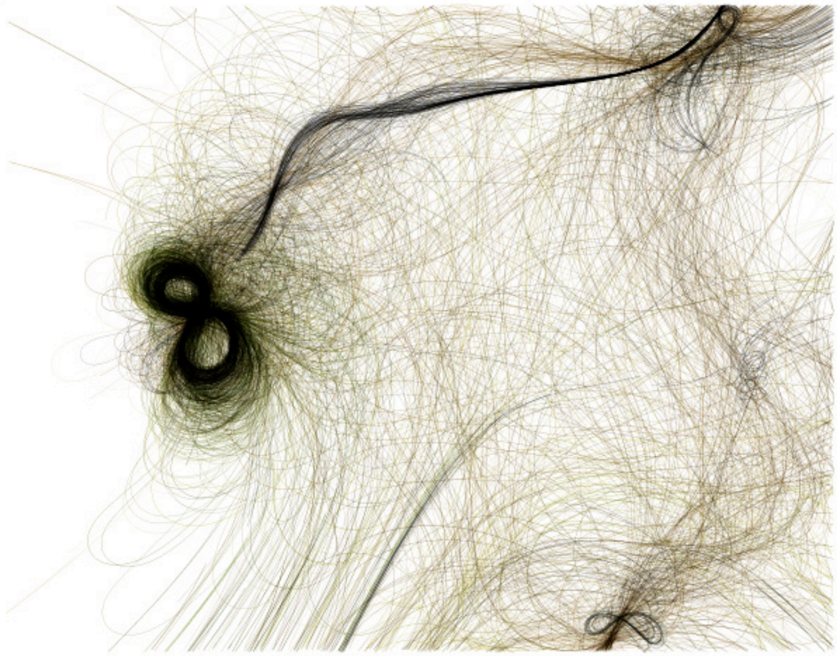

Figure 4.2: One of a set of 28 prints made with the *Tissue* software. Casey Reas 2002.

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).

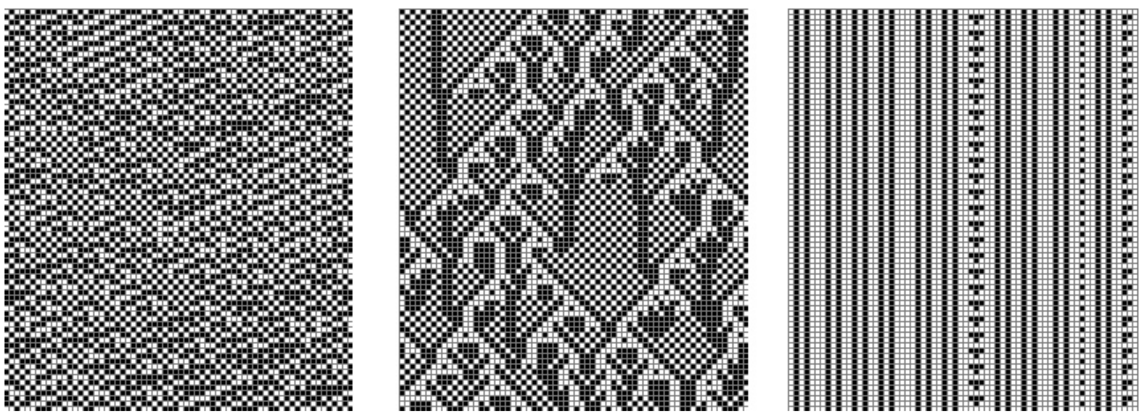

Figure 4.3: Examples of visual stimuli for chaotic (left) complex (middle) and ordered (right) rule sets.

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').

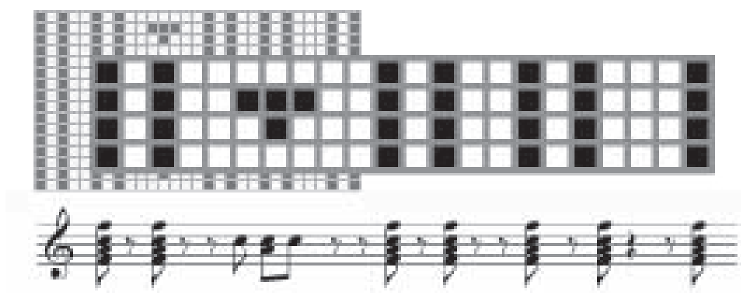

Figure 4.4: Rhythmic mapping: cell states are transformed to musical events: 1 = play, 0 = rest. Four lines are voiced simultaneously

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).

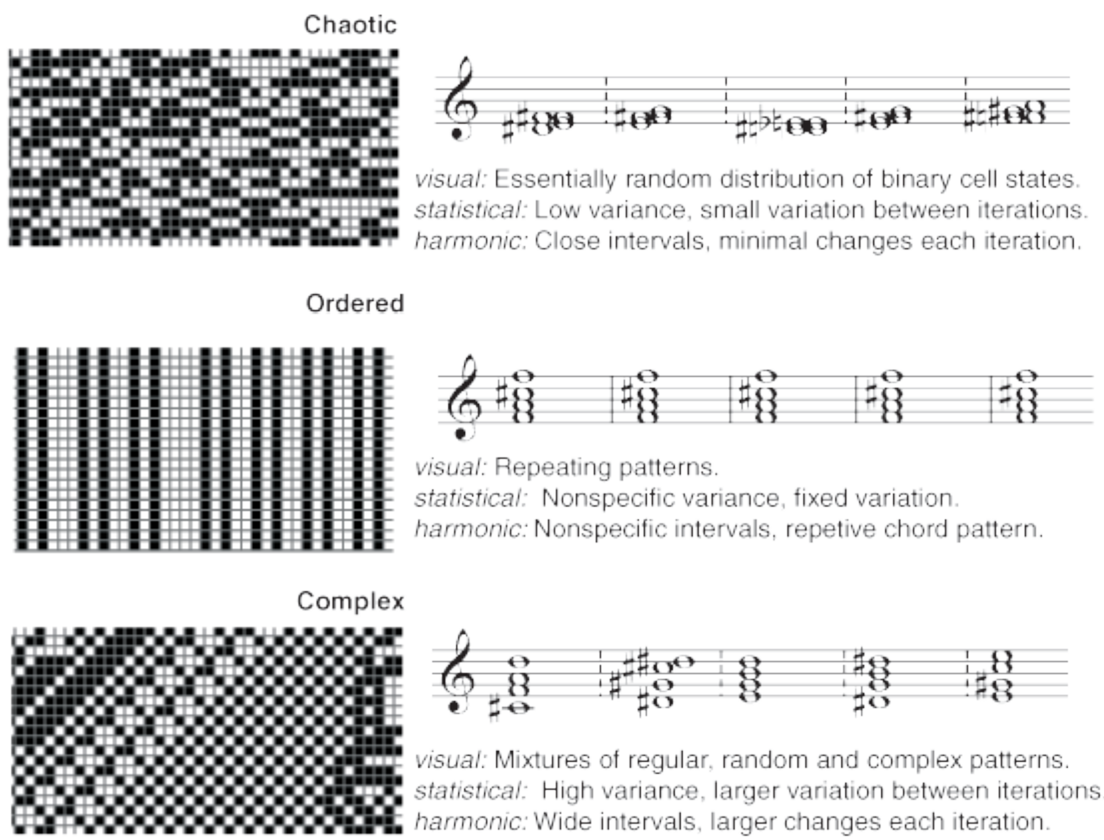

Figure 4.5: Harmonic mapping for the CA.

[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.

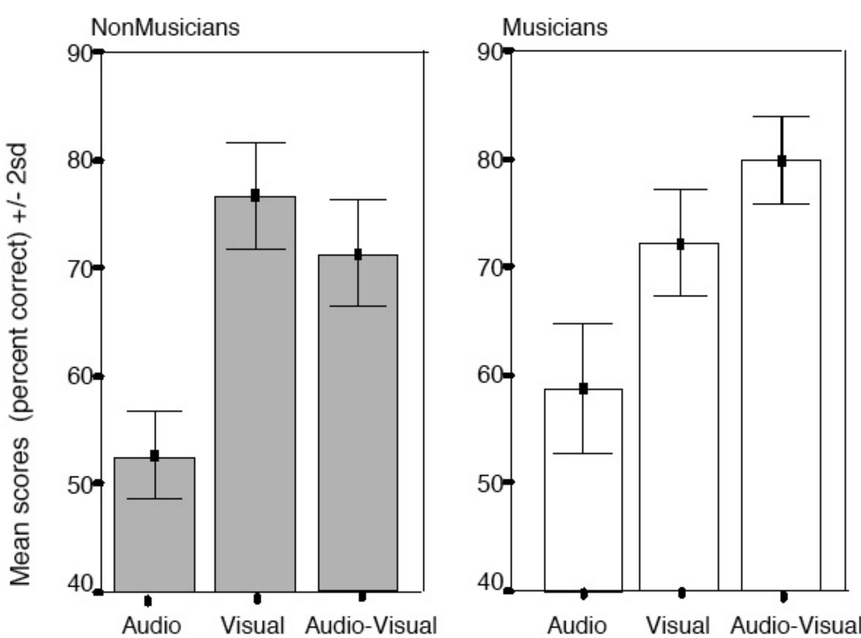

Figure 4.6: Mean scores and standard deviations for each group across all conditions

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

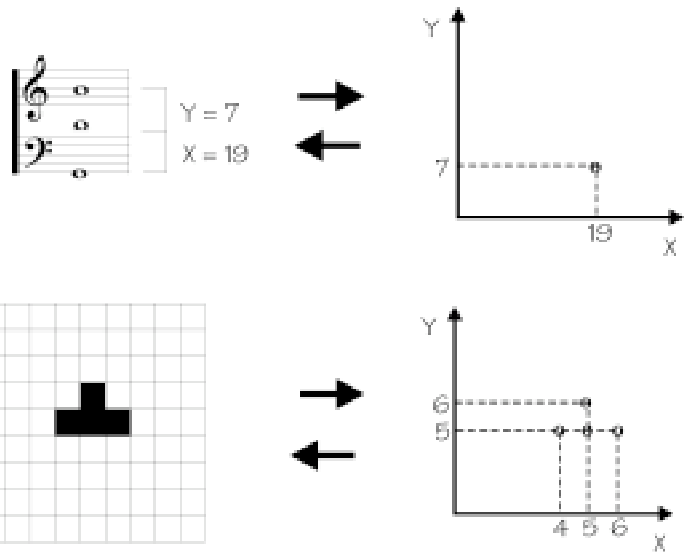

Figure 4.7: Mappings used in CAMus. The cartesian coordinates of a live cell are mapped to a triple (top). Each iteration of the rule set produces a number of such chords (bottom).

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.

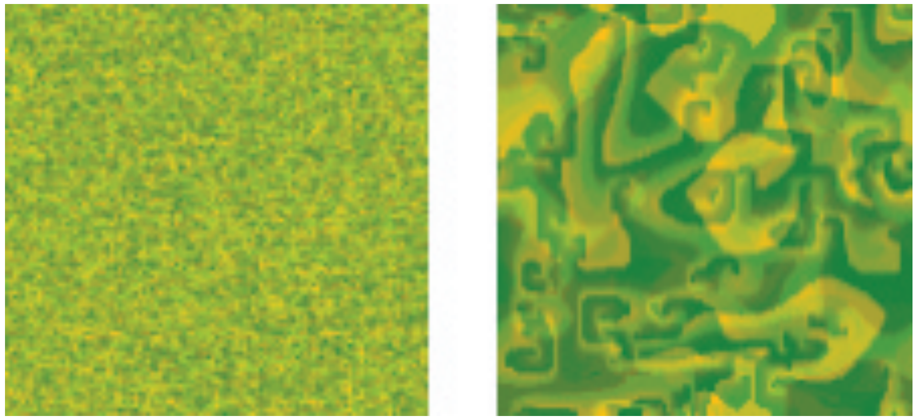

Figure 4.8: Evolution of CA used in Chaosynth: initial random distribution of cells (right) evolves to an oscillatory pattern (left).

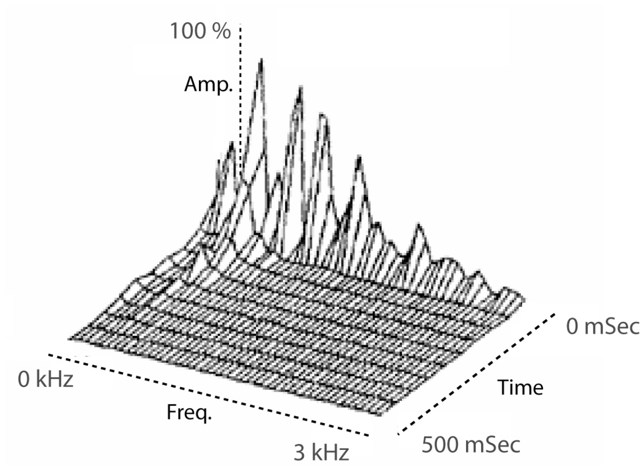

Figure 4.9: 3D wire sonogram showing evolution of spectra from initial white noise to 2nd, 3rd and 4th harmonic for a Mridangam (east Indian drum) stroke.

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.