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.

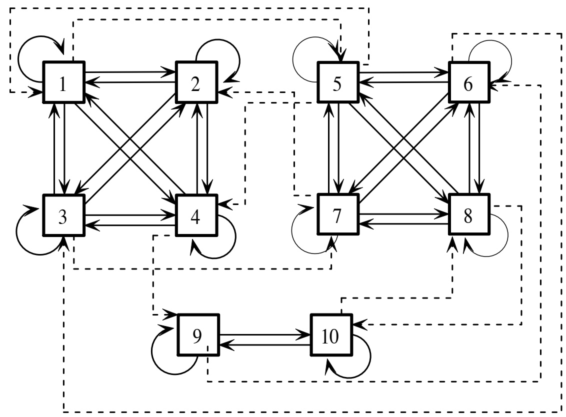


Figure 6.1: Diagram of homeostat network showing full intraconnections and 10% interconnectivity

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.

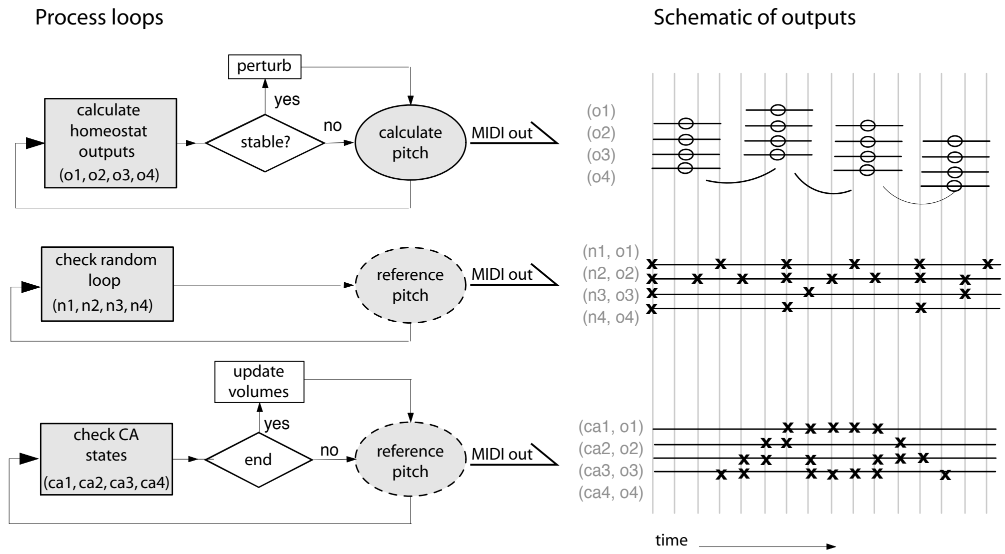


Figure 6.2: Schematic of the three process loops in AdSyMII and outputs produced.

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

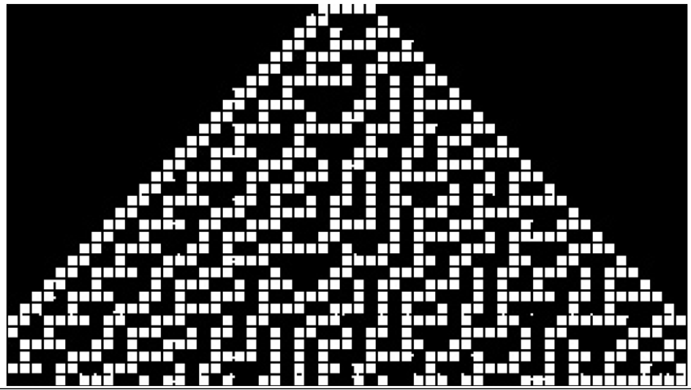
A simple stochastic method was used to determine the timings of the 'melody' line. N numbers $n_0, n_1 ... n_n$ are randomly selected. N defines the 'density', and can be any number, but somewhere between 3 and 10 works well. The value of n determines the frequency of each note i.e. it will be played 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 180*ms*. This reset to its original value of 300*ms* 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.

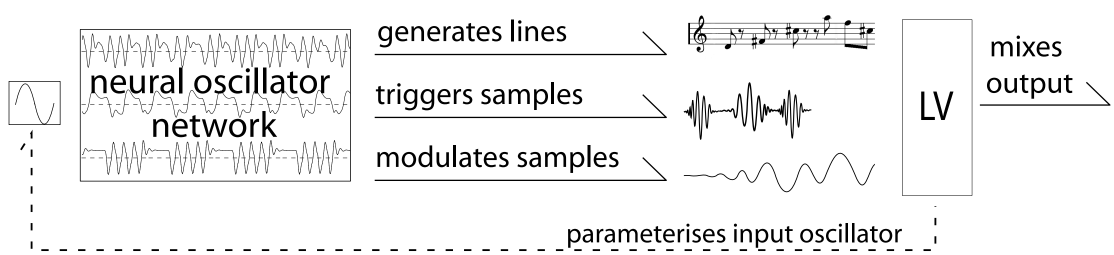


Figure 6.6: Schematic of the organisation of models and mappings in Organised Entry.

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.

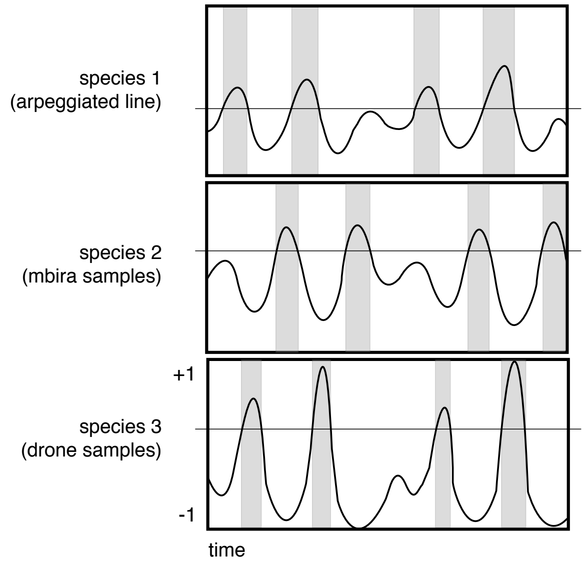


Figure 6.7: Three-species Lotk-Volterra equation as mixer. The horizontal lines represent the hand-set cut-off values. As the population density of each species exceeds the cut-off, the species-density is used to specify the amplitude of the associated part. Shaded areas signify the times when each part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