Chapter 8

Self-karaoke Machines: Collaborative Man-Machine Improvisations

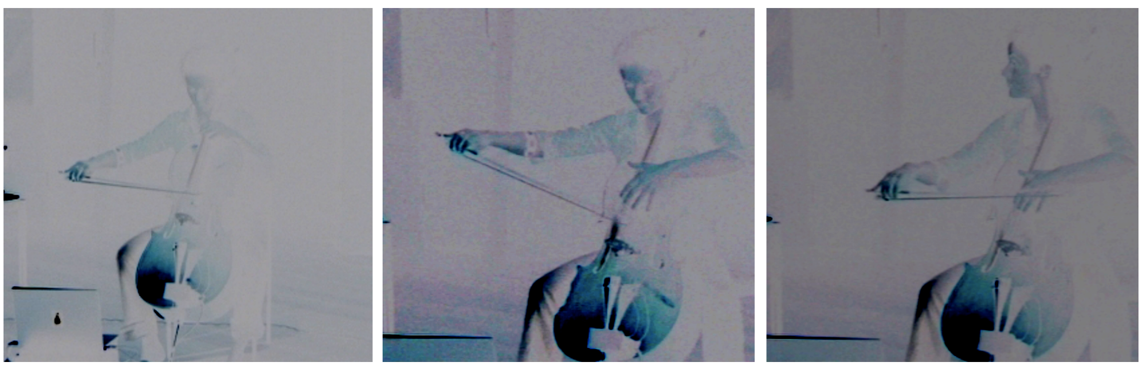

Figure 8.1: Fond Punctions performance at Third Iteration, Melbourne (2005).

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 6^{th} 2005 at `http://generatorx.no` in response to Golan Levin's 'Three questions for generative artists'. Presented at GeneratorX 2005 and posted on Oct. 5^{th}2005

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.

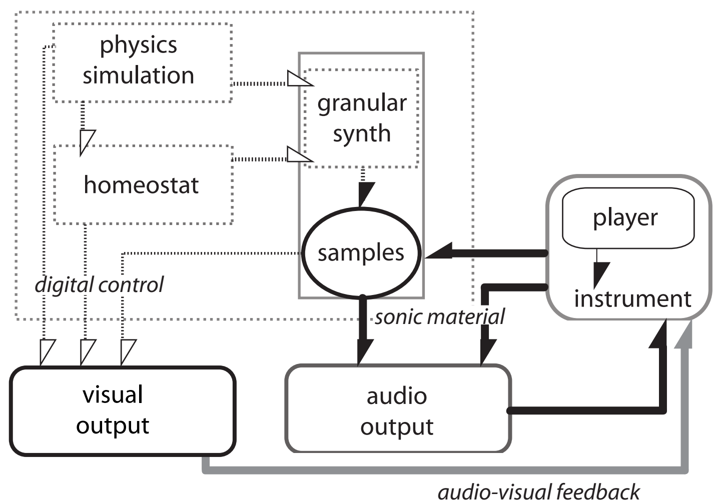

Figure 8.4: Schematic showing the network of influences between components in the whole performance system.

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.

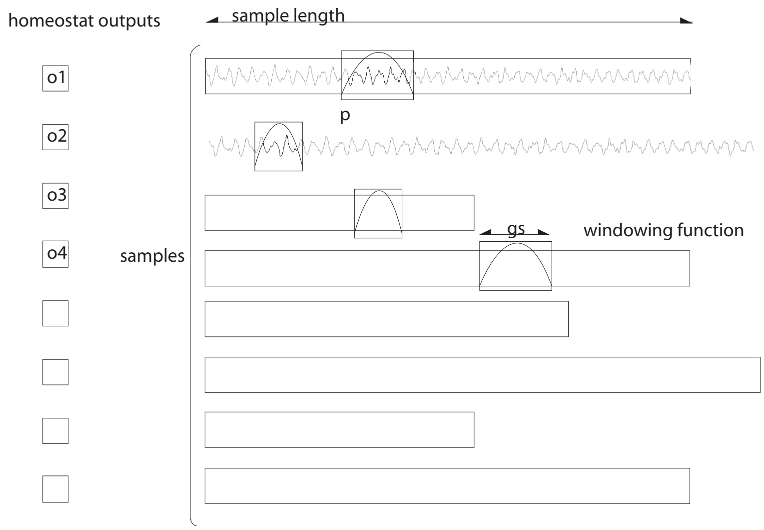

Figure 8.5: The outputs of each unit in the homeostat (O_n) are used to determine the point p in the sample that the grain is taken from.

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.

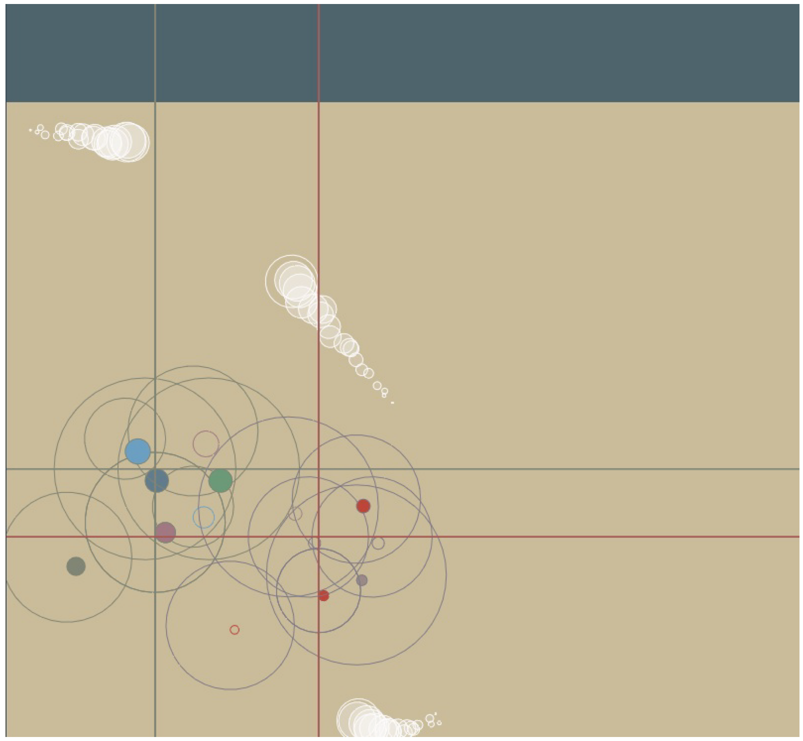

Figure 8.6: Screen shot of the video projections in the Self-karaoke Machine.

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

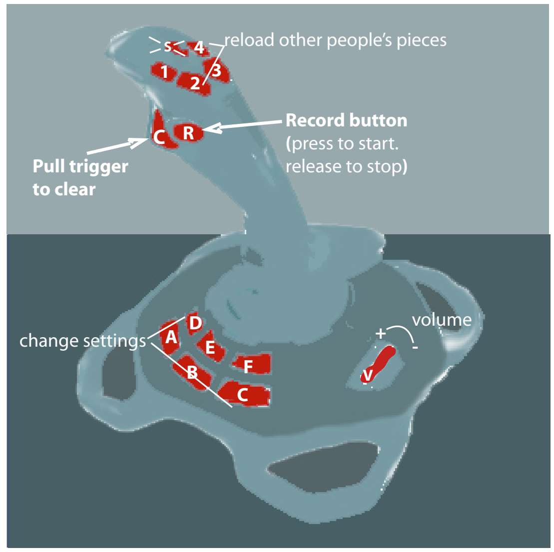

Figure 8.8: Diagramatic instructions for Self-karaoke Pond.

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

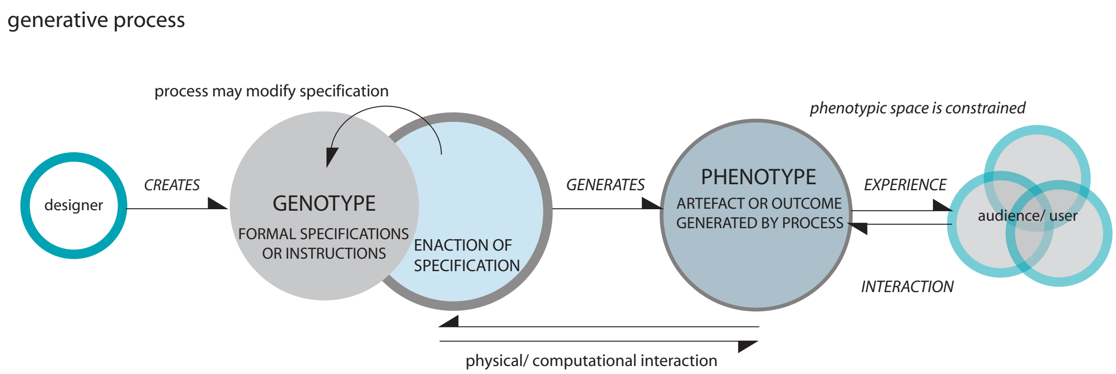

Figure 8.10: Overview of the interactions and influences in the generative process. (with kind permission from McCormack (2004)). The user's influence on the final outcome is constrained by the designer's decisions over the genotype, enaction process and material from which the phenotype is formed.

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.