Interactivity, where to from here?

GARTH PAINE

De Montfort University, Music, Technology and Innovation Research Group, Room 2.07, Clephan, Leicester LE1 9BH, UK
E-mail: Gpaine@dmu.ac.uk
URL: http://www.mti.dmu.ac.uk and http://www.activatedspace.com.au

This article is intended to raise some points of interest and mark out some pointers for alternative approaches to the design and execution of interactive music systems and artworks which pursue interaction that:

- **does not include any pre-defined pathways,**
- **takes dynamic morphology as its foundation, and**
- **implements dynamic software infrastructures, built on the object-oriented model, providing dynamic instrument instantiation, orchestration and timbral control.**

It is intended that such design would be a precursor to a new approach to interactivity that responds more directly and uniquely to those who engage in the work, and in so doing rewards them more richly for their time, energy and enthusiasm.

1. INTRODUCTION

Over the last six or so years, I have created a number of large-scale audio-visual environments that have been controlled, ‘performed’, or engaged with, by people through the dynamic of their movement and the patterning of their behaviour. This installation practice has fuelled a consideration of the nature of interactivity. The terms interactive and interactivity are broadly applied in the new media arts; however, the diversity of application has lead to a lack of focus. It will be argued that the term interactivity is therefore widely abused and, in line with Bongers (2000), that most systems are not interactive, but simply reactive or responsive because they lack a level of cognition.

In order to strip away the influences of current usage and artistic practice, dictionary definitions of the term interaction are introduced as a starting point in an attempt to establish a solid understanding of the semantics of the term, and subsequently, explore how such a definition might influence system design, artistic applications and a working perception of what an interactive system might be.

It will also be argued that a streamed approach to data representation and interpretations for digital (interactive) instruments and responsive (interactive) sound installations is preferential to event-based systems, because it more accurately reflects an ongoing interrelationship between the user and the system, and as such provides a more flexible, responsive and artistically rewarding outcome than a system based on the triggering of pre-made finite content (audio samples, for example).

Dynamic morphology (Wishart 1996) is explored as a conceptual framework for dealing with streamed data that facilitates an exploration of dynamic timbre in interactive, responsive music systems, and more broadly as a conceptual framework for the design of truly interactive systems, covering human–computer interface and sound synthesis applications. Smalley’s spectro-morphology (Smalley 1986) follows a similar line of enquiry, but from an analytical perspective. Both theories deal with the evolution of sound over time and provide a model for dealing with all sounds as compositional material. A detailed treatment of the nature of the relationship between the excitation source and the evolution of the sound is absolutely critical if those engaged in an interactive, responsive sound system are to perceive a correlation between the quality of their input and the quality of the resulting sound(s).

2. INTERACTION

In the technology age, interactivity has become a major consideration in the development of a contemporary artistic practice that engages with the proliferation of computer-based technologies. This proliferation has seen a revolution in the fields of animation and image generation as well as sound art and music composition. It is the harbinger of an entirely new artistic praxis, giving rise to time-based genres that apply everything from ‘artificial intelligence’ to tactile engagement as a means to actively engage the ‘spectator’/audience, and in so doing, make the work vary dynamically. Such a work is defined in the moment, by the nature of the direct or indirect influence(s) brought to bear. The audio-visual output is made up of either a collection of pre-defined material, triggered on the basis of distinct conditions, or generated algorithmically in real time. Much of this work is called interactive. However, what do we mean by interactive? So many things are said to be interactive that the common usage of the term is suffering from a lack of focus.

Before detailing the current approaches to interactive musical systems, the semantics of the term interactive

Organised Sound **7**(3): 295–304 © 2002 Cambridge University Press. Printed in the United Kingdom. DOI:10.1017/S1355771802003096

must be set on a firm foundation, and so I turn to the following dictionary definitions. The *Oxford English Dictionary* (2000) defines interaction as follows:

> the prefix **inter**-, [meaning] between, among, mutually, reciprocally.
> **interact**, [meaning to] act reciprocally or on each other.
> **interaction**, a noun, [meaning to] blend with each other.

The *Collins English Dictionary* (1992) contains the following definitions:

> **interact** is a verb, [meaning] to act on or in close relation with each other.
> **interaction** is a noun, [meaning] 1. a mutual or reciprocal action or influence. 2. physics, the transfer of energy between elementary particles, between a particle and a field or between fields.
> **interactive** is an adjective, [meaning] 1. allowing or relating to continuous two-way transfer of information between a user and the central point of a communication system, such as a computer or television. 2. (of two or more persons, forces, etc.) acting upon or in close relation with each other; i.e. interacting.

The prefix of the word interaction, inter-, is defined as something between, among, mutual, reciprocal. This definition implies that the two parties act upon each other; that the parties exchange something, they act upon each other in a way that is reciprocal. The *Collins English Dictionary* definition of interaction outlines an action that involves reciprocal influence. In the field of physics, it leads us to understand that an exchange of energy takes place.

How then do these definitions translate into the area of interactive music systems, and the new media arts? Does an exchange of energy occur when one is viewing a CD-ROM? That is, an exchange of energy between the viewer and the product? An exchange of information certainly occurs, but an exchange of energy? Probably not. The user requests a piece of information, and the computer, through the programming of the CD-ROM, delivers that information to a screen in such a way that the user can comprehend it.

One could argue that a transfer of energy takes place when someone is playing a computer game that requires a racing car driving wheel, a gear changer, brake and acceleration pedals to be used (many of these can be seen in amusement arcades). In this case the user is directly transferring energy through the interface by turning the steering wheel, changing gears and possibly operating the accelerator and brake pedals. This energy is certainly transferred to the interface – we know this from the 'OUT OF ORDER' signs that regularly grace the amusement arcades. The variation in condition of the interface is transferred as data to the computer program. This data is used by the computer to construct a scenario that is then drawn on the screen, to which the user responds. There is clearly a causal loop here. A causal loop being a scenario in which all parties require the others for their survival, and where the interaction of all parties maintains a balanced system. However:

- Does the racetrack alter in response to the behaviour of the driver? and,
- Is there actually a reciprocal energy transfer taking place?

The answer to both these questions is no, the user is simply attempting to maintain a state that is acceptable to the criteria of the game, i.e. to keep the car moving forward and on the track. The computer program defines the conditions to be met, and these conditions do not change as a result of users' input. They are the same every time the game is played.

3. MUSIC AND INTERACTION

In *Interactive Music Systems* (1994), Robert Rowe defines an interactive computer music system as 'one whose behavior changes in response to musical inputs'. In this case, we can assume the term 'musical' implies a common understanding. Winkler (1998) expresses this 'musical' understanding as being made up of a 'huge number of shared assumptions and implied rules based on years of collective experience'. Rowe continues by discussing the interpretation of low-level musical signals into structured high-level representations. Rowe suggests that interactive music systems:

> . . . interpret the input by evaluating human musical understanding (Rowe 1994: 3) . . . In their interpretation of musical input, interactive systems implement some collection of concepts, often related to the structure musicians commonly assume. Each interactive system also includes methods for constructing responses, to be generated when particular input constructs are found. As methods of interpretation approach the successful representation of human musical concepts, and as response algorithms move towards an emulation of human performance practices, programs come increasingly close to making sense of and accomplishing an instruction such as broaden the end of the phrase. (*ibid*.: 4)

Clearly Rowe's position only remains true while the input is of an instrumental nature. If the system input is a human gesture, be it a dance troupe or a solo music performer, or for that matter, a member of the public (as is true of most interactive, responsive sound installation works), the definition becomes problematic.

Here we face the first challenge in searching for a definition of what interactivity means when applied to digital music systems. The Rowe definition is founded on pre-existing musical practice, i.e. it takes chromatic music practice, focusing on notes, time signatures, rhythms and the like as its foundation; it does not derive from the inherent qualities of the nature of engagement such an 'interactive' system may offer.

Todd Winkler (1998) expands the Rowe definition with four distinctions. He suggests that interactive systems present differing levels of interaction, and specifies the following models:

- *The Conductor Model*, exemplified by a symphony orchestra, where everyone is controlled from a single source. The musical interpretation and execution is dictated by a master controller, the conductor.
- *The Chamber Music Model*, exemplified by the string quartet, where control may be passed from player to player at different moments within a performance. Winkler says, 'Intonation, phrasing and tempo are constantly in flux, with control often passed around to the musician with the most prominent musical material' (Winkler 1998: 25).
- *The Improvisation Model*, exemplified by the jazz combo, where the musicians operate within a defined framework, frequently pass control, and vary the musical score with improvised interjections, interplay and improvised solo passages. Winkler comments, 'Musicians trade off taking control of the music, fashioning their solos into spontaneous personal statements that alter and influence the surrounding accompaniment' (*ibid*.: 26). The outcome is perceived as musical because it has a shared form and concurs with Winkler's definition of musical understanding above. This behaviour establishes a kind of musical intelligence.

An extension of the improvisation model above is free improvisation, where a broad range of often chaotic interchanges, governed by a common musical understanding (the musical intelligence just summarised), encourages a homogenous musical outcome from a vast amalgam of inputs. No formal structure is agreed in advance, and so the musicians continuously respond to each other ad infinitum.

Whilst the Winkler categories form a useful platform for discussion, it should be noted that the notion of interaction is not well defined in musical theory. For instance, Bongers (2000) points out in *Physical Interfaces in the Electronic Arts, Interaction Theory and Interfacing Techniques for Real-time Performance* that many interactive music systems are in fact reactive systems, due to the absence of cognition. These systems represent Winkler's conductor model, failing to exhibit the dynamic intelligence of the other models.

As was true of Rowe's definition of an interactive music system, discussed above, these definitions are all coaxed in terms of existing musical practice. Interactivity may offer an entirely new approach to music-making, and so in order to avoid getting stuck in the current musical paradigms, we should question not only the nature of the system input (such as musical notes, tempi, rhythms, or human gestures, dance movement, or a conductor's gestures), but we should pay equal attention to the output of the system, and the qualitative relationship between the two.

In Interactive Music Systems, Robert Rowe outlines three stages of an interactive music system:

(1) *Sensing*, which includes pitch and rhythmic pattern detectors,
(2) *Processing*, which includes the scheduling of tasks that create musical events in response to the sensed inputs, and
(3) *Response*, where the constructed audible response is delivered back to the interactive agent.

In a situation where the system is designed to accompany or improvise with a musician, the construction of the responses within an agreed musical aesthetic makes sense; however, this approach does nothing to further our exploration of the inherent qualities of an interactive music system, it simply squeezes interaction into a known template.

I would like to proffer one further model based on the process of human conversation. Winkler also raises this conceptual model, commenting that human conversation, like any good interaction, is a 'two-way street . . . two people sharing words and thoughts, both parties engaged. Ideas seem to fly. One thought spontaneously affects the next' (*ibid*.: 3). The human conversation model represents a number of other characteristics as well: the conversation is

- unique and personal to those individuals,
- unique to that moment of interaction, varying in accordance with the unfolding dialogue, but is
- maintained within a common understood paradigm (both parties speak the same language, and address the same topic).

Within such an interaction the starting point is known by one of the parties, and whilst in some discussions there is a pre-existing agenda, in general the terrain of the conversation is not know in advance. It is a process of exchange, of the sharing of ideas. It is a cybernetic-like product of the whole, a product that is unique to the journey they share during their discourse.

This process of interaction is extremely dynamic, with each of the parties constantly monitoring the responses of the other and using their interpretation of the other parties' input to make alterations to their own response strategy, picking up points of personal interest, expanding points of common interest, and negating points of contention. This kind of interchange is much more in line with the kind of public/artist/system relationship that becomes apparent when an interactive system is exhibited in a public space as an interactive, responsive sound installation or immersive environment. In this scenario, the system's creator(s) cannot expect those engaging with the system to have knowledge of formal musical paradigms. Furthermore, when the input to the interactive system is a human gesture, it is questionable whether a musical construct, constrained by the precedents of historical musical practice (chromatic music, for instance), is an appropriate response. The appropriateness of response, i.e. a perceivable relationship between the gestural input and the system output, is a central issue in the design of interactive systems. This relationship should express the unique characteristics of each person's engagement and commitment to the instrument/art work/system and is, I suggest, best represented by a system based on streamed data techniques rather than triggered, pre-defined events. The mapping of sensed input data to processing algorithms is the most complex and subjective aspect of system design.

The mappings must be such that there is extensive scope for exploration and the discovery of new outcomes, but where the outcomes prove repeatable to the extent that they confirm the cognitive map that the interactor is developing as their relationship with the interactive system deepens. The mapping should reflect the nuance of the individual's engagement with the piece, and must therefore be made in real time, an issue discussed in more detail later in this paper.

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

In this situation, the designer must consider why, for instance, would one wish to convert the movement of an arm to a musical chord, which in turn embeds itself within a chromatic music structure, but does not necessarily address the weight and subtleties of the interactive gesture? What role or relevance does a sense of tonality have to the experience of interaction? Equally one has to contemplate the role of other historically ingrained aspects of musical composition, rhythm, melody, musical form and structure, which so clearly constrain the Rowe and Winkler definitions.

In order to address these issues, the interactive input must closely represent natural physical activities. Video sensing allows continuous tracking of human movement, and provides a continuous stream of data that represents a qualitative indication of the movement currently taking place. The following qualities can be derived from that data:

- direction of movement,
- speed of movement,
- size of a moving object,
- proximity of movement to other moving objects,
- inertia,
- consistency of movement, and the
- Cartesian co-ordinates for both the momentary position, i.e. per video frame, and the position of rest.

Whilst direction and speed of movement may be applied to existing musical parameters, they make much more sense when equated to the characteristics of naturally occurring sounds. A consideration of the concepts of dynamic morphology (a conceptual model developed by Trevor Wishart (1996) for describing all sounds, through the change of timbre, pitch and time in a 3D space) and spectral morphology (an analytical model developed by Dennis Smalley (1986) for analysing how spectra influence timbre) as an alternative approach to system output becomes relevant.

If interactivity is predicated on the ability of both parties to change in a way that reflects the developing relationship or discourse between them (as discussed above by way of the racing car computer game), we have to accept that multimedia systems that do not evolve their behaviour in relation to accumulated patterns of input (as described in the human conversation model above) are therefore not interactive, but simply responsive, a pattern of engagement that has been prevalent since the invention of the Theremin.

In order for the system to represent an interaction, it must be capable of changing and evolving. The process of evolution must promise continually new outcomes that are based upon the nature of a response–response relationship where the responses alter in a manner that reflects the cumulative experience of interrelationship. For this to be upheld, each exchange must be personal. That is to say, it must reflect the unique qualities of each particular dialogue. It must do this in a way that reflects the overall quality of the developing relationship and the transient moments of engagement.

We experience this kind of interaction every day, as outlined in the human conversation model earlier. If this kind of interaction were applied to the design of interactive instruments, they would need to adopt a structure that recreates, or evolves the system structure, mappings and generation algorithms dynamically in response to interactive input. The very structure of the system would become dynamic, and subsequently, the overall aesthetic potential of the system would evolve as a result of the momentary and historical interaction. How then do we contemplate the implementation of such approaches in interactive music systems?

4. SYSTEM DESIGN

As I mentioned earlier, interactive music research has largely focused on the study and creation of systems that perform tasks appropriate to note-based, chromatic music, where pitch is paramount, harmonic structures are integral, and a regular pulse forms a persistent rhythmic definition. This approach takes the vertical and horizontal approaches of existing musical praxis as its foundation, thereby firmly rooting all conceptual development within the framework of lattice-based music where pitch is one axis and time is the other axis. This approach is very well illustrated by the work of the Hyper Instruments Group at the Massachusetts Institute of Technology (MIT), who have focused on the adaptation of lattice-based musical approaches to electronic interfaces in the pursuit of interactive instruments (see their Virtual Violin and Beat Bugs, amongst other examples).

I suggest that the imposition of musical characteristics on interactive music systems only makes sense when designing a system as an accompaniment or improvisatory application where the other elements of the ensemble are adhering to chromatic music performance practice. Accepting a pre-existing musical framework as the criterion for the design of an entirely new musical paradigm clearly restricts the selection of development criteria. If one accepts that the aim is to create new means for musical expression, to generate new or innovative music, then the acceptance of historical criteria associated with chromatic music (i.e. pitch, pitch structures that define tonalities, the recognition of chords, the recognition of tempi and time signatures, etc., as outlined by Rowe (1993) as the characteristics of an interactive music system) is inappropriate. The field of philosophy defines this kind of criteria confusion as a Category Mistake (Ryle 1995), i.e. the application of characteristics, descriptive of one phenomenon onto another where they are not necessarily characteristic.

Perhaps it is time to develop a new conceptual framework for interactive music systems.

5. DYNAMIC MORPHOLOGY

The contemplation of existing musical paradigms suggests the creation and collection of individual events that equate to single points of action and only become part of a continuum of movement when contextualised within known musical forms. Dynamic morphology (Wishart 1996) is a conceptual model that can be applied to both sound generation and sensing systems, and suggests a continuously evolving stream of sound events (audible or silent).

The morphology of a sound is expressed in a ‘three-dimensional pitch-duration-timbre space’ (*ibid*.: 94), indicating a simultaneous and continuous multi-dimensional deviation through the characteristics of pitch, timbre and time. It is important to note here that Wishart is referencing timbre as one of the inherent musical characteristics of a sound. With few exceptions, existing musical composition practice investigates timbre as a function of orchestration rather than a product of a single source over time. A naturally occurring sound has just such a morphology, fashioned by the many facets of its initial sounding source and the architectural and acoustic space(s) it inhabits. Rodet, Potard and Barrière (1984) point out that our recognition of certain consonant sounds in speech is heavily dependent on the nature of this morphology, so it is vital to our understanding of the world, and our everyday experience of it. Wishart defines two primary areas of morphology:

(1) Gesture, which is ‘the articulation of the continuum by the agent which instigates the event’.
(2) The classification of morphology used ‘in relation to perceived natural phenomena’.

The morphology is only apparent if the sound-object is perceivable as a whole. Wishart comments:

> We may expect, however, that the category of gesture is in some ways more restricted than that of natural phenomena structures . . . It is in other respects more extensive than the category of natural phenomena – higher organisms are capable of very subtle articulations of the continuum, which we should only expect to find by chance in the structures of a natural phenomena. The interface of these two types of description may be seen in the relationship between vocal and instrumental music. A musical instrument is a device used to stabilise, through its resonance structures, the pitch and dimensions of a sound-event. The morphological structure of the sound-event is thus dominated by the characteristics of the natural phenomena of resonance. (Wishart 1996: 102)

A human movement is characterised by a smooth, continuously changing relationship of the limb to the body, each movement being made up of many infinitesimally small variations and adjustments. The overall movement defines a gesture, and the nature of the way in which that gesture is enacted is described as the weight of the gesture. It is not possible to extricate from a gesture the individual moments of movement and adjustment that make up the overall gesture. In this sense, human movement reflects Wishart’s statement that ‘sound objects with a dynamic morphology can only be comprehended in their totality’ (*ibid*.: 94).

Computer music is also characterised by continuance of variation and adjustment. Similarly, a single sound is perceived as a whole, not as a collection of the myriad samples that create it in a digital playback system, but as a stream of contiguous sonic information. Whilst it is possible with a computer to extract a single sound sample, it is meaningless to the human ear. Wishart points out that:

> In general, sound objects with a dynamic morphology can only be comprehended in their totality and the qualities of the process of change will predominate in our perception over the nature of individual properties. (Wishart 1996: 94)

When considering interactive, responsive installations that use sound as their principle medium, dynamic morphology can be applied to both the sound production process and the analysis and management of the relationship between the person(s) interacting, and the system response(s). For the installation to respond dynamically, there must be dynamic variation in timbre, pitch, amplitude and associated envelopes. The spectromorphology model (Smalley 1986) is also a useful tool for the analysis of the input gestures. As an analytical tool, it allows the differing weights of gesture to be categorised as distinct. If the gestural input data is stored in audio files (Wessel and Wright 2002) then this technique becomes particularly applicable. A real-time interactive system is dependent on the resolution of qualitative characterisation of the input data, and in the same way as a complex audio signal can be broken down into primary frequency partials, so too is it important to examine the individual movements that make up a gesture as a whole. Complex data analysis of this kind can provide a large number of control variables that make for subtle and unique synthesis responses.

If dynamic morphology is applied to the design of responsive and interactive instruments and installations, it becomes clear that the system design itself must be dynamic, and that during an interaction, an instrument must be able to change in fundamental ways to produce timbres that were impossible at its inception. In other words, it must be possible for it (in accordance with the nature of interaction) to evolve into a new instrument altogether. Dynamic software architectures are a hallmark of object-oriented software design.

6. OBJECT-ORIENTED DESIGN

The model for dynamic morphology is embedded within the concepts of object-oriented (OO) programming. An OO program is essentially a structure that facilitates the creation of variations on the defined classes (the term class carries a specific meaning in OO programming, explained below). This process allows the dynamic creation of as many versions of the class as are necessary. For instance, in a student marks register program, it would be the number of students on the course. This process thereby allows the dynamic creation of an infrastructure of functioning units from a set of blue prints (the classes). It allows this infrastructure to be created at the time of need, and to be resized as necessary (add a student, or lose a student).

The building block of all OO software design/application is the Class, a blue print for a functional unit that has a range of behaviours. In sound synthesis terms, the class may look for a range of variables and perform some operation on an audio input stream, or may generate an audio stream which is fed into a software-based mixer within the program (patch/instrument).

The class will contain some parameters in its description; these parameters must be provided by the process that creates the instance of the class (each instance of a class is called an object) in order for the class to become extant. For instance, the parameter(s) may be the frequency of the oscillator, or the variable that will be mapped to frequency, or the mixer channel or bus input into which the generated audio is to be fed. The functionality of the class is the domain of the code written within it, and it is essentially a blueprint for an object. Once extant, it becomes an object that performs a function, as described in the class design, until it is no longer required. At this stage the object may be disposed of, and a garbage collection mechanism removes the object from the computer's memory, thereby providing memory allocation for new object creation.

As mentioned earlier, the beauty of the OO approach is that units of functionality can be dynamically created, plugged together, detached and disposed of in such a way as to fulfil the momentary requirements of the global program. An analogy may be a symphony orchestra whereby, aurally, instruments are added when required and removed when no longer required. A composer will orchestrate a composition to use particular instruments for a desired timbral quality at points throughout the work; however, the capacity of the orchestra is essentially fixed. If four flutes are required, then the flutes are seated in the orchestra at the beginning of the work and remain until the end. What would happen if a work were being orchestrated in accordance with a responsive/interactive schema, predicated on dynamic morphology? Such a situation may require twenty flutes at one point, and no violins, and within minutes require sixty violins and no flutes. A software infrastructure can allow for such occurrences within the limits of the processing power of the host CPU (central processing unit), and the limitations of the software design.

If designing an OO interactive music system, dynamic morphology would extend beyond the scope of varying a fixed audio stream through the use of equalisation or other filtering, or an otherwise variable synthesised output from a collection of algorithms (which no matter how the algorithm is designed will have a finite range of aesthetic and timbral variation), to a dynamically forming orchestration. In such a dynamic orchestration, a new sound object would be created when the morphological scope (by which I mean the aesthetic range of expression) of the current algorithm is reaching its limits. The new algorithm may exist only as long as it is required, and may be augmented by other dynamically created instruments, before being disposed of. Something very similar may be generated again based on interactive input, but it will exhibit variation in accordance with the varied collection of the conditions of creation. In other words, if this instrument were a flute, the type of flute would be defined at the point of creation, by the process that creates it; hence, the object could be either a piccolo, alto flute, bass flute, wooden flute as the moment requires and in accordance with the scope of the parameter variables of the class.

Any synthesis approach that does not allow for the real-time addition of new algorithms is inherently limited. For instance, it is not possible to add a band to a filter in real time, an entirely new filter object must be created and the previous object disposed of, which creates a disruption in the audio signal. Lexically separate synthesis processes are required for true dynamic morphology of synthesis processes.

Existing sound synthesis software tools allow some individual instruments to be created and disposed of dynamically, so long as their algorithms are defined at the instantiation of the main program; the mixing infrastructure (routing, EQ, etc.) must also be created at the point of instantiation of the program so all the algorithms run at once, from the time the piece/system is started. The synthesis process must be connected to the base synthesizer (the foundation synthesis engine and audio output structure) at the time the program is instantiated. An entirely new synthesis process cannot be added after this point. This includes algorithms developed in response to changing conditions of interaction. The sounds may be audible or not as the input demands, but the processes are all constantly running. Clearly, this is not an efficient use of resources, and is not a good way to encourage dynamic morphology beyond the initial capabilities of the instrument(s).

Drawing on the previous orchestral analogy, the above situation would equate to the orchestra playing a wind quintet, but requiring all the other members of the orchestra to remain on stage, and worse, to require the entire orchestra to attend all the quintet rehearsals and sit silently in their chairs during the rehearsals and performances.

The predominant paradigm of music composition is assumed, whereby the composer/programmer creates the resources they expect to need for the entire composition at the beginning of the work. These resources contain a set group of instruments, with an inherently limited aesthetic and morphological scope, a limitation that in my view has no place in interactive electronic music performance, because:

- it does not address the human context, by which I mean it has evolved to cater for existing musical practice and does not addresses the potential flexibility of computer music systems or the developing range of approaches to interfacing with interactive music systems;
- it is not driven by artistic values, it is predicated by programming limitations;
- it is aesthetically limited;
- it does not allow for the evolution of a musical work over an extended time frame, where the context for the work may also change; and
- it caters to a paradigm based on a pre-determined musical work being performed by an expert performer, and as such does not cater to the indeterminate form or resource requirements of an interactive musical installation being ‘performed’ by inexpert agents (the general public).

The public exhibition of interactive musical environments is a good forum for the exploration of interaction in three-dimensional spaces, such as art galleries and public spaces. This method of working frees the artist/musician/composer from the predetermined and often unconscious patterning of existing musical practice, and by making the works truly three dimensional, engages with space as a compositional medium in much the same way as is true for dance. As mentioned earlier, this process allows for the application of the concept of dynamic morphology and spectromorphology to the movement and behaviour patterns of people within the installation, the weight of gesture, the qualitative relationship between the weight of gesture, the aesthetic of the sound outcome, and the way in which the space is utilised.

In a sense, the public exhibition of interactive musical environments creates an ecosystem formed by:

- the human presence and nature of behaviour,
- the response of the technology (the aural or visual response as experienced by the inhabitant of the installation), and
- the space itself.

The process of understanding this dynamic relationship between the human condition and the physical space is supported by the study of cybernetics, and in particular the closed causal loop, a phenomenon which forms the third and final pillar in this suggested approach to the design and exploration of interactive music systems. However, before moving to consider the closed causal loop, a further extension of software design approaches needs to be considered.

Throughout the above consideration of design concepts for interactive, responsive musical environments, I have talked of the environment, or instrument evolving over time in response to learned patterns of input. Such a system would require a level of cognition, i.e. a software infrastructure that could establish the patterns of interaction based on historical knowledge, and act accordingly. Neural networks are an approach to computing, developed for pattern recognition tasks, and capable of being trained with an initial set of sensitivities, and learning to evolve those sensitivities in response to varied input over time. The inclusion of a neural network in the kind of software design discussed above would allow the system to be trained:

- To recognise individuals from their gesture patterns and movement characteristics (a useful feature for training an interactive environment to respond independently to different people, and also useful for interactive dance performances where it could be compositionally valuable to attribute different response patterns to different dancers).
- To make subjective, qualitative judgements about

the observed movement/gesture patterns, so that the system could determine the intent of the movement or gesture. Qualitative data of this kind would greatly extend the scope of current systems, that can respond, as described earlier in this document, to speed of movement, position of movement, relationship between two bodies, etc. The accumulation of subjective, qualitative data would make additional layers of intention-based aesthetic responses available through the structured control of more sophisticated synthesis algorithms. These additional layers would in turn make for much more sophisticated, unique and individualised interaction.

- To control vast numbers of synthesis variables in a structured manner, directly related to the subjective, qualitative data output of the neural network, and in turn control much richer aesthetic outcomes.
- To analyse the aesthetic output of the interactive system, and generate new algorithms that would extend or fine tune the aesthetic scope of the output of the system. This may see the output algorithms of an interactive, responsive environment evolving over time, so that the response patterns of the installation adapt to an accumulated knowledge of how people interact with it, and in so doing may totally discard the algorithms the artist(s) established for the piece. This kind of development is currently occurring in artificial life animated worlds both in exhibition environments and online (see Sommerer and Mignonneau 1998).

Neural network development would directly support the ideas of dynamic orchestration, as outlined above through the use of object-oriented programming. For synthesis output to reflect the minutiae of individual gestures, the synthesis algorithms must become much more sophisticated, which in turn requires more variables, more than would be easily controllable in a non-structured manner. Neural networks provide a mechanism to control vast numbers of synthesis algorithms in a controlled and subjective, qualitative manner, and as such are a valuable addition to the interactive systems development tool kit.

The examination of cybernetics moves the design approach away from the foundations of existing musical or computer sciences practice into an area of contemporary exploration of the phenomena of the natural world and the human condition, areas that in my view are vital considerations in the development of a new artistic paradigm.

7. CYBERNETICS – THE CAUSAL LOOP

The relationship between the physical exhibition space, the technology used to execute the work, and the human movement and behaviour patterns that form the basis of the engagement, is a critical consideration in the development of interactive, responsive environment installations. One of the principal concepts of cybernetics is the causal loop. A closed causal loop is one in which each of the elements contained in the loop act upon the others in a constant and varying fashion to maintain equilibrium. The only influences on a closed causal loop are the elements it contains. As I shall outline below, an interactive, responsive environment installation exhibits the qualities of a closed causal loop.

Within an interactive, responsive environment installation, human movement and behaviour patterns act upon the technology, the sensing system collects information about the nature of the human movement, the weight of the gesture, the speed and direction of movement, and feeds the data to audio and video algorithms that respond in whatever fashion the artist has designed. The response of the system is presented in the physical space, and takes the form of changing sound patterns and variations in video or animation projections. In this way the technology acts upon the space, altering the architectural and energetic nature of the exhibition space, and these changes in the physical space cause an alteration of behaviour by those that inhabit the exhibition. This alteration of behaviour, be it one of excitation or placation, will be driven by an intention to bring the system to equilibrium, or drive it into an unsteady or chaotic state.

The human response to the alterations in the environment forces the closed causal loop into a further iteration. The input to the technology, the human movement, will probably vary, the output of the technology will vary and the physical space will in turn be changed, generating a new and distinct response from those within it, and so the loop begins again. Interdependent relationships are formed between:

- the technology that mediates the installation response,

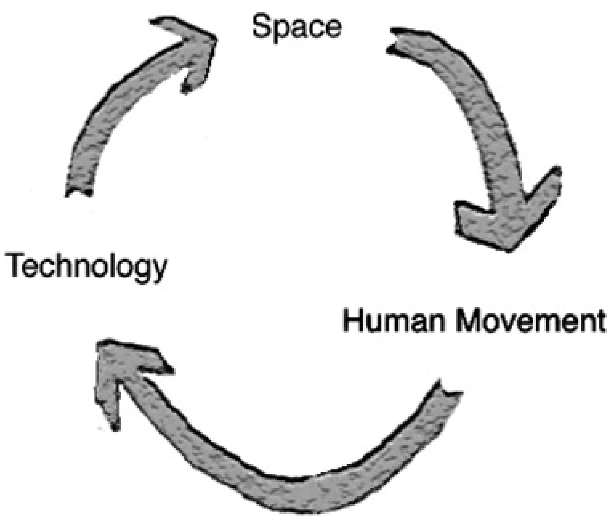

Figure 1. Closed causal loop.

- the human condition, the behaviour, emotions and relationships that are exhibited in the exhibition space, and
- the definition and experience of the physical space.

I have taken particular inspiration from Norbert Wiener, a key proponent of Cybernetics. He conducted a great deal of research into the application of cybernetic principles to the organisation of social systems. In 1948 he wrote:

> It is certainly true that the social system is an organisation like the individual, that is bound together by a system of communication, and that has a dynamic in which circular processors of a feedback nature play an important role. (Wiener 1948: 24)

Wiener's comments can be applied to interactive, responsive environment installations, and in so doing, imply that a well-designed interactive, responsive environment may represent patterns of social interaction, and subsequently provide a basis for the consideration of aspects of the human condition.

Fritjof Capra (1996) also focused on this aspect of the closed causal loop when he wrote:

> . . . the discovery of feedback as the pattern of life, applicable to organisms and social systems . . . (helped) . . . social scientists observed many examples of circular causality implicit in social phenomena, . . . the dynamics of these phenomena were made explicit in a coherent underlying pattern. (Capra 1996: 62)

So too are the patterns of relationship in an interactive, responsive environment made explicit and coherent through many iterations of the closed causal loop discussed above, each one rendering with greater detail the nature of the relationship. The user/inhabitant of the interactive, responsive environment installation develops a cognitive map of the responses of the installation, tests this map through iterative exploration, confirming prior experience, and actively engages in the evolution of the ecosystem in which they find themselves.

In 1998, the virtual reality and interactive installation artists Christa Sommerer and Laurent Mignonneau expressed similar thoughts when discussing the development of the interactive digital arts:

> . . . the art work . . . is no longer a static object or a predefined multiple choice interaction but has become a process-like living system. (Sommere and Mignonneau 1998: 158)

Sommerer and Mignonneau (1998) also comment that:

> From the insight that interaction itself and the interrelation between entities are the driving forces behind the structures of life, . . . artists investigate interaction and the creative process itself. Creation is no longer understood as expression of the artists inner creativity or 'ingenium' (according to Hegel) but becomes itself an intrinsically dynamic process that represents the interaction between the human observer, his/her consciousness and the evolutionary dynamic and complex image process of the work ('Art as a Living System'). Sommerer and Mignonneau assume, similar to Gregory Bateson (1982), that the patterns of mind (consciousness) and the patterns of matter are reflections of one another and part of an unbroken dynamic whole. (Sommere and Mignonneau 1998: 148)

One of the pioneers of interactive arts, the American video, and responsive environment artist Myron Krueger expresses a similar sentiment when discussing his early interactive video works:

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

Here, Krueger, as an artist, draws the same parallels expressed by Wiener, Capra, Sommerer and Mignonneau; he indicates that the experience of engaging in a responsive environment involves an active engagement with each moment, and that each moment of engagement contributes to the creation of the art work. The participant does not have the option of taking the stance of a detached spectator; they are inherently part of the process, part of the artwork/instrument itself. Indeed, the area of animation and digital imagery (the focus of Krueger, Sommerer and Mignonneau's work) is leading the way in creating interactive, responsive environments that evolve as a result of changing input conditions, a lead from which those involved in the development of interactive music systems could learn a lot.

8. SUMMARY

This paper has presented a range of approaches to interactive music system design, and in so doing draws some distinctions between the current note-based approaches, and a more sound-based approach that embraces dynamic morphology as a foundation for evolving aesthetic and musical outcomes. For interactivity to develop into an interesting and truly engaging art form, it must move away from the concept of interaction as defined in current interactive musical performance systems (embedded as they are in existing musical practice), and in CD-ROM and virtual reality systems where the user experiences a mediated journey along predefined paths, during which they are presented with predefined audio and possibly visual stimulation. We must seek to develop experiences that are not predetermined, and that reflect each individual's nuance of input in a unique and fulfilling manner. In working towards this point, interactive instrument and installation systems must be developed which change in response to accumulated user input in the ways discussed above. If they do not, as is the case for nearly all current interactive, responsive environment installations, they are, as outlined earlier, essentially responsive, and not interactive.

One important focus that assists this change of direction is the creation of sensing systems that explore streams of input data rather than individual triggered events. A data stream can provide ongoing qualitative data, referencing momentary changes in behaviour and movement patterns. An event-based system provides no information about the state of its input in-between sensed points of collision or contact with the 'hot spots', which produce the event trigger(s).

If the concept of streamed data is adopted as the basis for an interactive system, the software infrastructure must in turn be able to respond to dynamic changes of states. The combination of the concepts of dynamic morphology, spectromorphology and object-oriented software design have the potential to provide a conceptual and pragmatic framework for interactive music systems that allow variation beyond the scope of the original design, allowing the system to evolve over time to provide a range of sonic outcomes that truly enunciate the experienced input dynamics. Such a system would be supported by a neural network that conditions the inputs and evolves in accordance with perceived weightings of behavioural patterns over time, and may also be trained to recognise movement patterns, and make subjective judgements about the quality of sensed movement. Subjective outcomes would allow a substantial qualitative evolution of system response, and would also require that the synthesis infrastructure have the ability to dynamically create instruments that cater for the extended and evolving timbral requirements of the interactive process.

Cybernetic research has illustrated the application of systems thinking to social systems and all living systems. These concepts extend the dynamic morphology approach within a broader context – the human condition, the source and focus of much artistic endeavour. They provide a basis for the consideration of alternative approaches to the design of interactive systems, i.e. ones in which the total is greater than the sum of the parts, and in which the relationships of the individual elements, and their momentary responses define the characteristics of the entire system.

REFERENCES

Bateson, G. 1982. *Mind and Nature*. New York: Dutton.

Bongers, B. 2000. Trends in gestural control of music. In *Physical Interfaces in the Electronic Arts. Interaction Theory and Interfacing Techniques for Real-time Performance*. IRCAM – Centre Pompidou.

Collins English Dictionary, 3rd edn. 1992. Sydney, Australia: HarperCollins Publishers.

Krueger, M. 1976. *Computer Controlled Responsive Environments*. Doctoral Thesis, University of Wisconsin, Madison.

Oxford English Dictionary. 2000. Oxford: Oxford University Press.

Roads, C. 1966. *The Computer Music Tutorial*, 2nd edn. Massachusetts: The MIT Press.

Rodet, X., Potard, Y., and Barrière, J.-B.. Fall 1984. The CHANT Project: from the synthesis of the singing voice to synthesis in general. *Computer Music Journal*. The MIT Press.

Rowe, R. 1993. *Interactive Music Systems*. Massachusetts: The MIT Press.

Rowe, R. 2001. *Machine Musicianship*. Massachusetts: The MIT Press.

Ryle, G. 1995. *The Concept of Mind, The Oxford Companion to Philosophy*. Oxford University Press.

Salzman, E. 1967. *Twentieth-Century Music: an Introduction*, 2nd edn. New Jersey: Prentice-Hall inc.

Smalley, D. 1986. Spectromorphology and structuring processes. In S. Emmerson (ed.) *The Language of electroacoustic Music*. Basingstoke: McMillan.

Smalley, D. 1997. Spectromorphology: explaining sound-shapes. *Organised Sound* **2**(2). Cambridge, UK: Cambridge University Press.

Sommerer, C., and Mignonneau, L. (eds.) 1998. *Art & Science*. Wien: Springer-Verlag.

Wessel, D., and Wright, M. 2002. Problems and prospects for intimate musical control of computers. *Computer Music Journal* **26**(3): 11–22 Massachusetts: The MIT Press.

Winkler, T. 1998. *Composing Interactive Music, Techniques and Ideas Using Max*. Massachusetts: The MIT Press.

Wishart, T. 1996. *On Sonic Art*. Amsterdam: Overseas Publishers Association B.V.

Wiener, N. 1948. *Cybernetics*. Massachusetts: The MIT Press; reprinted 1961.