The perception of potential: interference, dimensionality and knowledge

Jonathan Impett[1] and Leonardo Impett[2]

[1]School of Media and Performing Arts, Middlesex University, London
[2]St John's College, Cambridge
j.impett@mdx.ac.uk

Abstract

This paper presents a system that investigates the sonification of wave interaction in a performance space and its interaction with a live performer – the illumination of sonic activity within a real space, in contrast to conventional ALife algorithmic, event- or material-based approaches. The model maintains three parallel representations of the entire live/virtual system: wavespace, symbol space and performance space. The cross-modal analysis and representation of behavior is important to the evolution of the system, which displays emergence on multiple levels of structure. Micro-evolution takes place within the population of wave-emitting and –listening agents. A higher level of structure emerges from their aggregate in interaction with the live performer, and a formal level as symbol space learns from the performer. Cross-modal representation is seen as a significant factor in the evolution of Western art music, in the development of multi-leveled structure and of work that affords many dimensions of engagement. We discuss the nature of knowledge produced through working with such systems and the role of the subject in ALife-generated knowledge. New models of simulation-derived knowledge are seen as important to cultural understanding.

Introduction

ALife and creative music systems

The relevance of the ALife paradigm to music is well established. Dynamic, evolving populations of behaviors interacting within a constrained space, the emergence of structure unforeseeable from initial conditions, the role of self-simulation and learning and the negotiation of otherwise intractable networks of relationships are system properties that resonate with musical minds.

ALife approaches are particularly attractive to artists working in real-time or interactive environments (see, for example, Miranda 2011, Miranda and Biles 2007). Stricter algorithmic systems are locked within their fixed aesthetic and behavioral dimensionality. Alife architectures appear to be a way of distributing creativity (you can't make 44,100 good decisions a second); composer, performer or environment effectively become co-agents or super-agents. At the same time, there is often a suggestion that such systems in turn might reveal something of the mysteries of music, or of human aesthetic responses, or of sound-based evolution. In the general case, we are dealing with a constructive metaphor; any verisimilitude is aesthetic. This is not to discount the knowledge inherent in such experience. Indeed, we shall discuss below how the example of an ALife music system can usefully raise questions about the relationship of the subject with such models and the nature of knowledge thus produced.

Interactive music applications have been the most fertile area for development. Technology affords exploration of the space between conventional notions of improvisation and composition, a space non-navigable by traditional means. Concepts such as real-time composition present non-trivial questions as to the relationship between apparently opposed activities of instantaneous decision–making and reflective architectural planning – apparently, because the instant (James's *specious present*) is crucially informed by intention, expectation and habit, and formal architecture is invariably modulated by the sequential, situated development of initial conditions.

The present system

ALife-based music systems tend to work on an event basis. That is, they deal with notes of sound objects or MIDI data. Such decisions are pragmatic (this devolves much of the processing), perceptual (we tend to think we listen to music in terms of notes) and cultural (the study of Western music remains largely text- or symbol-based). Artists such as Ryoji Ikeda and Carsten Nicolai work with the evolution and interference of wave-forms, but as entities abstracted from their environment (Ikeda and Nicolai, 2011). Feedback-based work exploits sonic characteristics of particular spaces and technologies (Alvin Lucier's *I am sitting in a room* and Nicolas Collins' *Pea Soup* are canonical examples). This system discussed here attempts to address the apparent dichotomy between "compositional" approaches based on the behavior of pre-designed materials and "environmental" approaches that explore a context. These views have distinct discourses; in the current UK Research Excellence Framework "Sound Art" falls within the purview of fine art, "Composition" under performing arts.

The present system sets out from a different perspective. We consider the real space of performance as a unitary volume within which micro-sonic wave activity is ubiquitous. The apparently silent blank canvas of music is illusory, such an anechoic reality even distressing. Such a view echoes Ingold's observations on soundscape:

> … neither sound nor light, strictly speaking, can be an object of our perception. Sound is not what we hear, any more than light is what we see. The scaping of things - that is, their surface conformation - is revealed to us thanks to their illumination. When we look around on a fine day, we see a landscape bathed in sunlight, not a lightscape. Likewise, listening to our surroundings, we do not hear a soundscape. For sound, I would argue, is not the object but the medium of our perception. It is what we hear in. (Ingold 2007, 11)

In his early work on sound, Bill Viola proposed a taxonomy of sonic behaviors, likewise developed from an analogy with light:

> A partial list of some of the most basic physical phenomena studied by the acousticians reads like a set of mystical visions of nature:
>
> Refraction …
> Diffraction …
> Reflection …
> Interference …
> Resonance …
> Sympathetic Vibration …
>
> Each of these phenomena evokes wonder, even after their scientific representations have been rationally understood. […] The processes of contemporary media systems are latent in the laws of nature – they have existed in various forms since the beginning of history. (Viola 2013, 41-2).

Figure 1: interference phenomena exhibited by the present wavespace system

The work presented here was developed for one of a sequence of compositions exploring wave phenomena in situated sound. It investigates the use of interference phenomena (fig. 1); others explore refraction and diffraction. Reflection plays a role in all three, as a key component in defining the relationship between space and activity.

We take a wave-based approach to modeling activity within the space. The potential complexity of even aurally trivial musical constructs presents a challenge. Instead of attempting to consider an intractable number of relationships we approach this complexity by looking at the behavior of the gradients and interference patterns generated by the musical material-generating agents that inhabit the space. In this respect, meteorology perhaps provides an analogy; a tornado is not an external event projected into a neutral space, but a product of dynamics within the environment itself. Nevertheless, in the domain of human activity it is best understood and responded to as an autonomous entity with its own behavior. We might think of the process presented here as second-order sonification.

We suggest that an important component of musical composition is the remapping of materials and phenomena between different modes of representation – from audial (the sonic imagination) to symbolic (notation) to physical (working at the piano), for example. The present system maintains multiple views of its space: representations of its "real", virtual/mathematical and symbolic or eventual state.

Technical Description

Implementation

In this implementation the performer uses a metatrumpet – a microcontroller-extended instrument that communicates its internal and external soundworlds and all physical activity to the computer via Bluetooth and radio. The system software is written in C++, using Max/MSP as interface and sound engine; they communicate using the UDP-based OSC protocol.

Wave space and performance space – a double bind

The system architecture incorporates three parallel spaces: a virtual wavespace based on a graphical model (maintained in the C++ programme), a symbolic space using CMMR-like representations (quantized to semitones and in time) and the performance space where sounds are played and heard in the physical world (both handled in Max/MSP) (fig. 2).

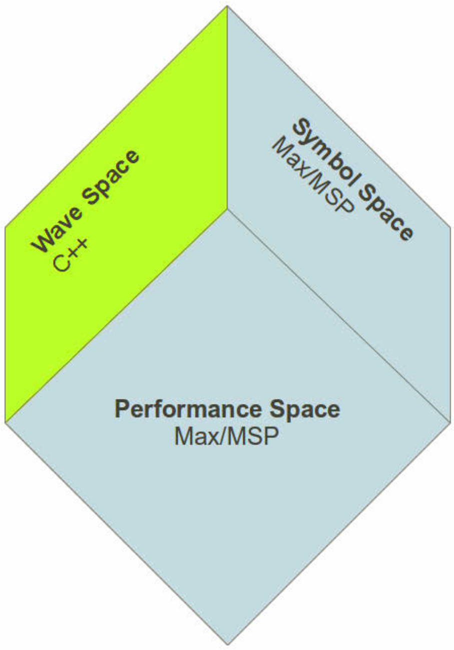

Figure 2: The three parallel spaces can be metaphorised geometrically as three orthogonal planes; navigation in one plane does not affect navigation in the other two, but all three are affected by the dynamics of the system

The performer initiates wave-generating agents. These are singularities in the wave-field - effectively beacons radiating sine waves. The wave behavior of each agent is initialized with values of initial position, frequency, speed of sound and spatial attenuation. Associating speed and attenuation with each individual agent allows for distortions of the space; different agents might find themselves in media through which sound propagates at different speeds. This also permits the effective representation of nonlinear spaces much larger than the actual performance space – large but bounded acoustic spaces within which interference artifacts fall within the range of human hearing. The agent's position is recalculated at every time step, and can be fixed, periodical, rule-based or learned (fig. 3). A repulsive force of proximity prevents terminal convergence. Once initialized, an agent generates one or more sine waves for the duration of its existence. In the initial model, agents are active or not. For reasons of practicality, the model is a steady state approximation. The field is conservative; it has no memory. If an agent is destroyed, all its effects immediately disappear from the space.

```
struct wave{
   float amplitude;   //The starting amplitude of the emitted wave
   float xpos;        //The x,y positions of the agent
   float ypos;        //...
   float pointvelocity;  //The agent's velocity through space
   float omega;       //The emitted wave's angular velocity
   float beta;        //The emitted wave's spacial attenuation
   float learn;       //The agent's learning agressiveness
   float contr;       //The agent's estimated contribution...
   float controld;       //...and in the previous timestep, for comparison
   float angle;       //The agent's spatial direction
   bool dominant;        //The agent's state: is it in the most harmonic pair?
};
```

Figure 3: C++ data struct for each wave-emitting/-listening agent

At each system time step we know the phase dependent contribution of each agent to the resultant signal at each point in space. In the initial model, system speed is 1 Khz, and the dimensions of the wavespace 256 * 256 points. Interference patterns within a space are perceived clearly when the dimensions of that space are an order of magnitude above the wavelengths in question. If we assume the scale of a typical performance space to be of the order of 10m, this grid allows for the representation of signals within the range of human hearing – wavelengths of the order of 1m. A balance of precision and computability is particularly critical in a real-time system servicing streams of audio output.

Sound production

The total wavespace field is calculated by summing the contribution from each agent at each point. Identifying areas of interest – loci of active interference and potentially emergent behavior – requires deconstructing intuitive understandings. The wavespace itself is indifferent. Initial strategies included looking for peaks in the image derivative – effectively high pass filtering – and edge detection. Both produced indiscriminate amounts of data and proved computationally expensive. The strategy adopted is to identify points of maximum interference by looking at the difference between the scalar sum and the vector sum of the contributing waveforms.

Figure 4 shows the gradients of interference produced by three agents in such a space, and the point of perceived maximum interference calculated as described above.

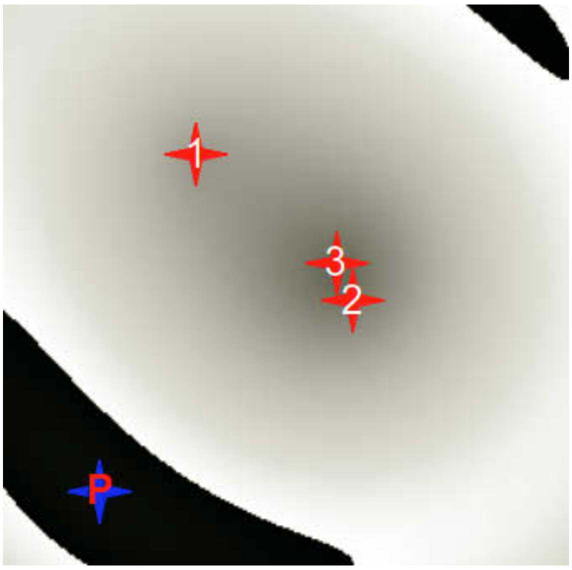

Figure 4: Three agents in wavespace (1-3) and the point of maximum perceived interference (P)

Interference in this sense is result of phase differences between sinusoids. Another less formal understanding of interference might entail the production of additional noise or artifacts. This is an embodied phenomenon. In the case of hearing, for example, modulation artifacts are a function of the response of the basilar membrane (Moore 2002).

Having localized the point of maximum interference, a hypothetical sound is produced by taking the weighting the various agent audio signals by values stored in a matrix that represent their estimation of their contribution at that point. This data is valid for one wavespace timestep (1ms). The audio stream is subsequently filtered at timestep frequency to mitigate windowing artifacts.

We consider this signal as being emitted radially by the maximum as a point source in wavespace. This sound is never heard directly. In keeping with our embodied understanding of interference, the beacon agents also become listeners and performers in performance space We then hear the combination of their listenings to the combined reflected signal, which now themselves exhibit interference due to their phase displacements. The agents are dynamically located in physical space using IRCAM's *Spat* spatializing application within Max/MSP (http://support.ircam.fr/docs/spat/3.0/spat-3-intro/co/overview.html).

This information is sent via OSC to the agent's Max/MSP avatar. The audio stream of each agent is constructed by additive synthesis in Max/MSP from the discrete frequency spectrum. It is then buffered to maintain a stable audio stream. It is delayed and attenuated according to its position in wavespace and its local speed of sound. In addition to the frequency/amplitude/phase data representing the point source signal, the performance space is informed of the positions of the other agents and of the maximum. Before synthesis, the agent can make certain musical decisions. In particular, frequency addition artifacts above the range of human hearing (c. 20 KHz) are transposed down to become audible. This initiates a recursive exchange between wavespace and performance space through which increasingly complex sounds can accumulate.

Accumulating agency

In an early iteration of the system, agent position is determined dynamically by a fixed velocity and a direction of motivation towards the point of maximum interference. A dance develops as agents track the maximum point-source, the location of which is in turn determined by the agents' position. Quasi-oscillation, temporary attractors, reflections, line-following, convergence/divergence,and leaps through space all occur. The latter is not just a function of the bounded wavespace. The real performance space it reflects is also bounded, and such events are therefore considered to be of potential musical-structural significance. Figure 5 shows the movements of three agents and the point of maximum perceived interference for the 56 timesteps leading to the positions shown in Fig. 4.

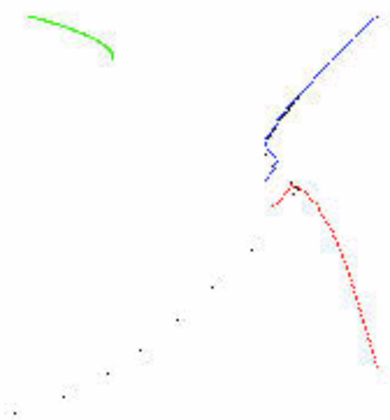

Figure 5: movement of three agents and point of maximum perceived interference through 56 timesteps preceeding fig. 4

Microevolution. A more advanced system is developed where a dynamic number of agents N navigate the space, and the system judges harmonic relationships between pairs of agents in an NxN matrix. The position of the smallest element in this matrix indicates the pair of agents with the strongest harmonic relationship. In a reflection of the bio-evolutionary paradigm, the agents which are not part of this pair reproduce imperfectly and are destroyed, creating imperfect re-initializations of themselves on destruction. Consistent with the spatial model, the degree of mutation is dependent on the proximity of the agent to the point of maximum interference, representing information loss. After successive generations, a different set of two agents will exhibit a stronger relationship, and the set of agents which reproduce will change accordingly. Each agent continues to have a behavior dictated by both the initializing conditions outlined above, and learned behavior (controlling agents' own frequency and direction of spatial movement) maximizing their contribution to the point of maximum interference. The aggressiveness of the learning algorithm is itself an initialized parameter. Agents within the harmonic relationship are understood as changing their behavior whilst agents outside the harmonic relationship change their nature.

Polyphonic rhythm. The performed sounds of each agent are subject to a further thresholding process. Agents calculate their own contribution to the sound perceived at the maximum point source by comparing the interference intensity at that point with and without their contribution.

They become active as performers only when this contribution exceeds a threshold – that is, the sound of each agent is turned on and off with a rhythm of variable linearity with windows of periodicity. This creates a polyrhythmic fabric, each strand of which is related to the virtual maximum. The texture can be inverted – transformed to its negative by reversing the threshold test. A challenge in music systems with any degree of automation is to generate the multiple levels of structure found in art music. The system again analyses the relationships between individual agents as pairs, but this time looking for convergence in audial performance space. When both audio streams of a pair of agents converge, giving out a 'drone', a timer is initialized; after a fixed time period, both agents are silenced in performance space. However, they continue to operate in both the wave and symbol spaces, which under the micro- and macro-evolutionary paradigms outlined eventually cause the pair of agents to exhibit different audial-spatial behavior, and their inaudible audio streams diverge. Under these conditions, the audio streams of the two agents are restored to performance space, again becoming audible to the performer (fig. 6).

Figure 6: Thresholding behavior of three agents over time

Macroevolution – symbolic representation

A third parallel space maintains symbolic representations of performance activity. Streams of sound output, synthesized and live, are analyzed in terms of pitch and rhythm. In both domains this constitutes a process of re-representation and categorization roughly analogous to human cognitive behavior; the point here is not anthro-verisimilitude but process architecture. Perceived pitch is quantized to semitones, rhythm to a 50 ms. grid. Inevitably these processes lead to loss of precision, of information, but they allow for the apprehension of relationships on other levels. The aggregate wavespace output is then analyzed in terms of harmonicity, density of behavior (a measure of number of events perceived over time), “noisiness” (signal-derived value) and stability of amplitude. The rhythmic profile of individual streams are compared by searching for a lowest common denominator – an additive pulse – and looking for matches or complementarity.

Initial threshold values are set for these parameters together with a global coefficient of surprise. Principal Component Analysis is used to produce values representing the balance of factors in the current state. Symbol space directs change in wavespace under two kinds of condition:

- when a subset of parameter thresholds is exceeded
- when the performer triggers a change. In this case, symbol space learns the current PCA state. When the same PCA state recurs (within a given tolerance), symbol space will subsequently trigger the same changes autonomously. PCA thus represents a higher-level representation, less dependent on specific detail of behavior. In this way a repertoire of formally salient states evolves through interaction with the performer.

The changes sent from symbol space to wavespace are distributed among agents by the main loop – effectively a super-agent. They relate to the following properties:

- number of agents: the population may be forced to multiply or divide
- frequency bands: the potential frequency space inhabited by each agent (initially 20-20,000 Hz) may be divided into separate bands, such that the agents are distributed among bands of varying width and distance.
- range and distribution of speeds of sound
- triggering threshold
- repertoire of wavetables (wavetables written from live sound and from wavespace output are added to the initial sine tables)

The system thus evolves over longer time-scales, approximating to formal time in compositional terms, by learning a repertoire of transition states through interaction with the performer.

The performed space

The live performer is incorporated through performance space. FFT data is passed to wavespace where the performer appears as an agent among others, contributing to the aggregate interference pattern. This is captured by a microphone at the position of the performer, such that the performed sound of the agents as heard by the performer also figures in this stream. Clearly the performer is not as free to move in wavespace (and doesn’t have the mental capacity to calculate interference maxima). Instead, motion data from the metatrumpet directs both movement in virtual space and the positioning of sound in performance space. The performer controls the range of frequency components to be passed. In a four-band implementation, for example, one might select components 9-12 to avoid the perceived pitch and its low integer multiples, focusing the system’s attention on the activity of higher partials. To facilitate calculation and avoid additional artifacts (from the arbitrary relationship of table length to frequency), a cycle at the lowest perceived frequency is written into a wavetable.

The performer can work with current tendencies, attempting to counter or encourage them, trigger new agents directly, or mark a current state as a moment of transition in symbol space.

Further Developments

Our intention is to enhance the relationship between virtual wavespace and actual performance space by incorporating an acoustic space-tracing algorithm (Dokmaniç et al, 2013). This impulse response-based technique can be incorporated both in musical terms, by allowing the perceived interference patterns to evolve from a single impulse, and technologically, in that the necessary microphones already form part of the system. Wavespace precision will be unchanged, but the representation can be mapped nonlinearly onto the physical space.

Agents will be empowered to search their own perceived spectrum, to select and transpose series of higher partials that contribute a greater range of color to the whole.

The performer will be able to mark a state as stable as well as transitional. If the potential for any of the evolving repertoire of stable states is perceived in symbol space, wavespace will be encouraged in this direction.

A predictive function will be incorporated in symbol space, in the hope of enhancing the measure of novelty. It will learn by correlation with the performer’s real-time decisions.

Conclusion

The re-representing of behavior across modal boundaries is characteristic of Western art music of many kinds (Impett, 2009). Indeed, we could understand its historical development as being the incremental and technological exploration of this possibility. Polyphony – a complex system which cannot be fully resolved or reduced (unlike, say a simple melody-and-bass structure) – is the most characteristic property of the last thousand years of art music. The kind of cross-modal interaction explored here allows for the emergence of potential structure on multiple levels and their interaction to generate a complex, compound architecture in which detail on one level relates to form on another. We attempt to create a dimensionality of engagement (of expectation, of analysis, of interpretation) higher than that of purely algorithmic or performer-driven systems. Looking at interference and convergence provides a way of grasping the state of a complex system and deriving understanding of its potential architecture.

This system can be tuned to the tastes of the musician; it is replicable but not falsifiable. Rather than invalidating aesthetic models as knowledge-producing environments, this obliges us to confront the role of the subject in knowledge production. The apprehending subject is part of the knowledge producing system; production and apprehension become inseparable. ALife models are used to model societies, cultures, populations and complex systems, but rarely to understand the individual. This may be partly historical, in contradistinction to GOFAI for which the implicit paradigm was perhaps a cartoon of human intelligence. We cannot avoid the role of the subject, however, particularly as such open systems present behavior that cannot be represented in its entirety. In exploring the nature of this new kind of knowledge, Delanda describes their behaviors as singularities in a *space of potential* (Delanda 2011, 18).

This work is part of a wider theoretical project, the *work-without-content*. It takes its cue from Giorgio Agamben's observation that in a culture of infinite difference, of the right to expression in unique individual languages, the common experience on which cultural exchange is possible becomes the confronting of the blank page. He describes the instantiation of such work as giving form to potentiality (Agamben 1999). Here the material derives entirely from the relationship of performer and space.

What is the relationship between the subject as pilot of the system and subject as co-agent? The situation can be characterized in terms of power and rules; it is effectively a legal question. We can look to Agamben once more. He considers the role of the sovereign, the individual in relation to which laws obtain, and yet who may stand outside them. He can be assassinated, but not condemned to death. Likewise his *homo sacer*, the outlawed subject whom the state will not kill but another individual may choose to (Agamben 1998). Sovereign power or homo sacer, the rules and behavior of a system such as this produces knowledge – if the production of truths is how we understand the function of art (Badiou 2001, 41-57) – only by incorporating such a subject.

There is another important aspect of emergent knowledge-production - an idea that science, in turn, might take from critical theory. Such knowledge is experiential and situated, it denies commodification and resists attempts at repetition and falsification. In this respect, music might be seen to provide a model for the new knowledge. Meanwhile the commodification of culture proceeds unabated. Science has built for itself a position of immense power and responsibility – effectively a sovereign power, in fact, with which it can validate and rewrite the conditions for knowledge. It is important that the sciences of the artificial look at the nature of the new kinds of knowledge they generate and reflect them back to culture.

References

Agamben, G. (1998) *Homo Sacer: Sovereign Power and Bare Life* (trans. D. Heller-Roazen). Stanford University Press, Stanford.

Agamben G. (1999) *The Man without Content* (trans. G.Albert). Stanford University Press, Stanford.

Badiou A. (2001) *Ethics: An Essay on the Understanding of Evil* (trans. P.Hallward). Verso, London.

Delanda, M. (2011). *Philosophy and Simulation: The Emergence of Synthetic Reason*. Continuum, London.

Dokmaniç, I., Parhizkar, R., Walther, A., Lu, Y. M. and Vetterli, M (2013). Acoustic echoes reveal room shape. *Proc. National Academy of Sciences* June 17, 2013, doi:10.1073/pnas.1221464110.

Impett, J. (2009). Making a mark: the psychology of composition. In Hallam, S., Cross, I. and Thaut, M., *The Oxford Handbook of Music Psychology*, pages 403-412. Oxford University Press, Oxford.

Ingold, T. (2007). Against Soundscape. In Carlye, A., editor, *Autumn Leaves: Sound and the Environment in Artistic Practice*, pages 10-13. Double Entendre, Paris.

Miranda, E. R., editor (2011). *A-Life for Music: Music and Computer Models of Living Systems*. A-R Editions, Middleton, WI.

Miranda, E. R. and Biles, J., editors (2007). *Evolutionary Computer Music*. Springer, London.

Moore, B (2002). Interference effects and phase sensitivity in hearing *Philosophical Transaction of the Royal Society. London A*, 360: 833-858.

Ikeda, R. and Nicolai, C. (2011). *cyclo_id*. Die Gestalten, Berlin.

Viola, B. (2013). The Sound of One Line Scanning. In Lander, D. and Lexier, M., editors, *Sound by Artists*, pages 39-62. Charivari Press, Etobicoke, Ontario.