Interactive Sound: Generative approaches from Computation and Cognition

Roger T. Dean

*austraLYSIS, Sydney, and MARCS Institute, University of Western Sydney, Australia

roger.dean@uws.edu.au

Abstract

I discuss processes of interactive generation of sound which are common to settings with and without acoustic instruments. I range from processes which realize mathematical or compositional formulations as music, to empirical models of music and their generative use. I particularly emphasise cognitive models, and their current and potential value as generative tools. Consideration is given to both live algorithms and live coding.

Keywords: cognitive models, live algorithms, live coding

1. Introduction

How can we usefully structure real-time interactive music generation systems? I am considering here two kinds of situation in which I and many others practice music-making(Dean, 2003). First, the juxtaposition of acoustic instruments (the piano in my own case) with computational sound generation processes. Second, purely electronic computer-interactive performance. In the first case, a key feature is that performer-generated or –realised musical strands feed into the performance, and potentially to the computer-interactive system. And here it is important that a performer who has facility on their instrument (itself a kind of interface) can participate. In the second case, a musician can develop substantial facility with the computer-interactive interface itself, using gesture, controllers or voice, but only with the voice will their input be itself a musical stream, and one which can contribute immediately to the overall sound stream.

In both situations, many real-time computational approaches are applicable; and a key distinction seems to me, as just implied, to be the nature of the input stream that the performer provides: whether primarily musical, or primarily a stream of generative or controller information. After the stage of musical stream input, the opportunities of the two situations are essentially identical, and these opportunities are my topic. So I focus on algorithms which may be used in acoustic, purely electronic/computational, or hybrid circumstance. I consider in particular 'live algorithms'(Bown, 2011), those whose path can be perturbed in flight. One might say, those with chinks in their black-box armour. I make some contrasts with 'live coding'(Collins, McLean, Rohrhuber, & Ward, 2003), where there is at the outset generally little or no armour, no carapace, rather a jelly of wobbling potential.

2. *Traditional algorithmic approaches to music creation: the realization of mathematical formulations as music*

The pages of the prestigious Computer Music Journal are strewn with articles on specific mathematical ideas that have been embodied in music: for example, power functions, fractals, chaotic functions, and mathematics derived from kinematics or other aspects of physics and organismal movement. Nierhaus(Nierhaus, 2009) provides a recent extensive review of such algorithmic approaches, though he seems to conclude that they may be restricted in impact and/or longevity. On the other hand, Brown and Sorensen(Brown & Sorensen, 2009) indicate that they regularly find a particular set of algorithms valuable in their generative work in live coding: "probability, linear and higher order polynomials, periodic functions and modular arithmetic, set and graph theory, and recursion and iteration". Interesting amongst these contributions are those by Jeff Pressing(Pressing, 1990), whose appreciation of motor and cognitive science issues always informed the way in which he applied the algorithms to make music, foreshadowing the more recent approaches I turn to as the article proceeds.

3. *Embodying a compositional approach, which is not explicitly mathematical, in an empirically derived algorithm*

This endeavor has a long history, including the work contributed by people such as Ebcioglu(Ebcioğlu, 1988) and Cope(D. Cope, 2001) . Both created models of prior music, but had to extract or formulate the model, and then implement it for generative use. Ebcioglu formulated a probabilistic mechanism to generate harmony sequences allied to those of Bach Chorales. Cope also sought to recreate styles of earlier composers, from Mozart to Prokofiev and Joplin, by statistical transitions. His patent(D. H. Cope, 2010) also seeks 'retrograde recombination', in which patterns are recombined, normally in reverse, to generate a kind of remix, a new piece, mainly focused on pitch and rhythmic structure. This is of course somewhat like mosaicing, or remixing, but perhaps with an emphasis on larger, more unified blocks of material.

4. *Computational approaches based on a compositional technique*

There have been some attempts, particularly very recently, to use the procedures of serial composition, as developed by Schoenberg, Webern and Berg, a century ago. This is a particularly rigorous system, in which pitch sequences (and subsequently, in total serialism, sequences of any other chosen musical parameter) are repeated and transformed. For example, a 'Prime' sequence of non-repeating pitch-classes, commonly 12 notes, may be reversed (the Retrograde), or Inverted, and the Inversion may also be reversed. A class at Columbia University has recently written code for these serial pitch transformations, which are quite easy to achieve. But what is interesting about serial composition with pitch is the way that different realisations of the sequential progression (the 'horizontal' temporal pitch succession) is combined with vertical integration. In this integration, chords are formed, and several strands of the Prime and its derivatives may occur in any possible juxtaposition; note repetitions, and transpositions of the series are also important. I have written a live algorithm, the *Serial Collaborator* to make multipart piano music using these principles, but with a range of interactivity permitted, for example controlling the degree of overlap of different series version, by varying the note density of chords and their frequency (Dean, submitted). These principles have many interesting future applications for live interactive performance, focusing on pitch, rhythm, timbre, spectral density, spatialisation, or a virtual infinity of other salient musical features. So far, I have used this software in live performance, and to provide material within soundscapes in installation work in collaboration with Keith Armstrong and colleagues (visit http://embodiedmedia.com : see the works *Finitude* (2011), and *Long Time, No See* (2013)).

It seems that quite often considerable benefits from generative code occur when the code is squeezed into a new context or use. This is the case with the *Serial Collaborator,* as it can operate on any note sequence, not solely genuine 12-note series. When presented with entirely major scale melodies, for example, the transforms are also entirely within a related major scale : thus the C major scale ascending from C, when inverted becomes the Ab major scale descending, starting from C. This is a feature well know to certain medieval composers such as Obrecht, but not commonly exploited since, and it provides contrasting outputs from the Serial Collaborator, especially when probabilistic small note transpositions are also allowed with the major scale materials, providing inflections somewhat equivalent to passing and grace notes.

5. *Statistical models of music or musical structure*

An alternative approach to generative principles is based on statistical analysis of musical features, at the micro- (event by event), meso- (section by section), or macro- (whole piece) levels. The micro- approach is dominated by information about statistical pitch structures. For example, Bayesian and information theoretic analyses allow the analysis of sequential patterns of pitches, in principle using sequences of any length (not more than 10 are usually informative). These analyses may then be used generatively. For example, using their IDyoM model, my colleagues Marcus Pearce and Geraint Wiggins have obtained statistical data on various symbolic corpora of classical tonal music, and used it to predict musical perceptual and performative features, but also to generate chorales and other musical forms, in keeping with earlier styles(Pearce & Wiggins, 2006; Wiggins, Pearce, & Mullensiefen, 2009). These data were obtained 'unsupervised', that is, without musical information being provided separately, and they are productive. OMAX, developed under the IRCAM auspices, is a counterpart endeavour which allows real-time use of related principles, and operates on audio data streams as well as symbolic data (note description, midi information).

The meso- approach is one which can flow for example from our own analyses of acoustic intensity profiles in realized music. We found that on event, metrical unit, phrase, or even longer time frames, intensity rises and falls follow a statistically dominant pattern(Dean & Bailes, 2008, 2010a, 2010b). In this pattern, rises are shorter and involve faster changes in acoustic sound pressure levels, than do falls. Given that there are finite limits to feasible (or survivable) acoustic intensity levels, generally the rises and falls are overall balanced, so that the intensity mean level is fairly constant. There are potential explanations in terms of psychoacoustic or evolutionary adaptation phenomena: for example, an increasing intensity with time is more likely a signal of danger in the natural environment than is a decreasing pattern. There are also analogies with movement kinematics, and devices built into, or accessible within, performance parameter control software such as Director Musices(Friberg, Colombo, Frydén, & Sundberg, 2000).

I have used this feature of intensity profiles as a structuring device in some live algorithms. As soon as one does this, the issue of temporal scale is raised. If the observed temporal asymmetry in duration of rises and falls is applicable over durations up to say, 20-30 seconds, then what of longer periods? To obtain the same strength of empirical data on this is difficult if not impossible with music, since the precision of an analysis increases with the number of data points obtained: so to obtain good precision about musical patterns operating over say 10 minutes, one might need to analyse multiple (say 50) musical works each of which is at least 16 hours long (so that each piece would provide at least 100 data points for the analysis: 100x10min > 16h)! The problem is of course not the analysis, but the availability of material. Consequently, a composer has to take quite arbitrary decisions in applying such statistical data about temporal profiles of intensity to composition and improvisation. Arguably, this is analogous to the 'squeezing into a new context' I discussed in the preceding section. But can one find an alternative approach that limits this arbitrariness?? We will return to this issue later.

Another 'meso' approach I am pursuing involves time series analysis models of music and its performance(Bailes & Dean, 2012; Dean & Bailes, 2010c). These have generative application. Most if not all musical events show what is called 'serial correlation', which means sequential temporal correlation (and does not refer to serial music). For example, a high note is most often followed by another, and not by a very low one; a loud note by a loud note; and movement patterns, similarly, necessarily are continuous, so that the position of the movement at one instant is the strongest predictor of the (nearby) position at the next. Without going into any technical details here, a time series model of a continuous process is a mathematical formulation, taking into account serial correlation, of the impact of the controlling factors (say acoustic intensity) on the other continuous processes (say note duration, or pitch). Furthermore this influence may of course be reciprocal, rather than unidirectional. If one builds such a model of an ongoing musical stream, such as the events entering MAXMSP, then that model, expressed in terms of the relevant musical features, be they timbral power spectrum, or acoustic intensity, can be used to generate future events. The model can be updated regularly; or it can be static, representing an image of the whole continuous process as a single entity.

Currently, Geraint Wiggins and I are writing Max externals that incorporate time series models, which can be updated, and which act generatively in real-time. We have working prototypes, which are already useful, and plan to integrate this work with that mentioned on real-time IDyOM implementations: the two approaches are potentially highly complementary, as IDyOM is based on the mico- and sometimes meso- levels, while time series analysis is essentially macro- (though it too can be broken down into successive meso-level sequences).

6. *Sonification*

Somewhere between the categories and continua described above lies sonification, as used in compositional contexts. Sonification is primarily the representation of data in sound, with a view to facilitating the recognition and comprehension of meaning and pattern in the data(D. Worrall, Bylstra, Barrass, & Dean, 2007; David Worrall, 2009). A wide range of techniques are used, from the quite literal (where for example frequency of occurrence of some event might determine the pitch or loudness or the sound representing it), to highly filtered mappings. In the context of composition and sonic interaction, sonification bears consideration because gesture or other performative components may be mapped to sonic outputs by the same processes. In addition, music can be made in relation to real-time data streams arriving from any process on (or outside) the planet. I will not pursue the topic further here, but merely note that most of the issues of interaction discussed elsewhere in the article are relevant to its use.

7. *Algorithmic compositional approaches based on cognitive studies, including computational models of cognition*

I want to suggest here that an interesting way forward towards such an end is provided by the empirical sciences of cognition, and by computational modeling of cognition.

Let me give first an example based on a disputed aspect of psychoacoustics: the degree to which we can locate low frequencies in space(Hill, Lewis, & Hawksford, 2012). The literature on this is in conflict, partly because the two conventionally understood cues to location, inter-aural temporal and intensity differences, do not suggest good theoretical mechanisms for low frequencies. On the other hand, some empirical evidence, especially with listeners in off-centre positions, and with subwoofers situated on the floor, is strongly positive: localization seems effective. One possible line of investigation, which we are considering, is that timbral differences between sounds arriving at the two ears, consequent on asymmetries in room reflections and absorption contribute. Another possible factor is the 'seismic' like vibrations which can be transmitted from floor speakers, and may have separable vestibular sensor mechanisms (for example, see(Todd, Rosengren, & Colebatch, 2009)).

I raise this topic primarily to mention low frequency spatialisation in interactive composition and performance, something which is probably underutilized, since most speaker arrays only have one subwoofer. It is also interesting in the light of the enthusiasm some composers, notably Robert Normandeau have for spatial organization of frequency distribution in their compositions for multi-speaker systems. Normandeau (personal communication) and I share a liking for elevated subwoofers, suspended in mid-air, where the floor vibrations are at least likely to be altered. With a pair of subwoofers, one on the floor and one elevated, many interesting possibilities may exist. And to return to generative algorithms, clearly spatialisation is an important aspect of electroacoustic work, and given suitable performing spaces and speaker arrays, algorithmic control of low frequency distribution is a tempting possibility, partially pursued in the creative commercial software Kenaxis, an effective interface based on MaxMSP. Kenaxis has an interesting spectral distribution option, developed by its author, Stefan Smulovitz.

Let us turn next to computational modeling of cognition (see(Farrell & Lewandowsky, 2011) for a practical introduction). This can mean at least two main kinds of model. One kind deals with empirical cognitive data, and seeks a parsimonious representation of it, usually in terms of the putative cognitive components. The other uses data about basic cognitive functions, such as the speed or the spiking and transmission pattern of nerve impulses, hence of the interaction between sensory and motor systems. Given such neurophysiological data, large scale 'system models' of the cognitive system have been built, such as ACT-R(Anderson et al., 2004). These system models then can predict the speeds of various response processes, and sometimes their nature and effect.

Given either kind of model, it is not difficult to envisage a path by which it can be translated into musical action, given some assumptions about the nature of the musical substrate: whether pitch based or sound based, metrical or otherwise, and so on. I can point to few examples of the rigorous application of these ideas as yet. On the other hand, historically the idea of using continuous electrophysiological data from human brains to drive music synthesis has been powerful (from Rosenboom in the 1960s onwards(Rosenboom, 1976) to Miranda and others at present). Given this, I suggest there is reason to hope that future work will provide further controlled, systematic application of computational cognitive models to music making.

8. *Interaction with live algorithms: and the opportunities of live coding.*

It has been suggested that live coding is a particular practice 'that balances the capabilities of the computer and the human' partly because programming 'can become intuitive', and hence 'an extension of the musical imagination'(Brown & Sorensen, 2009). Many practitioners also believe that the field has considerable as yet unfulfilled potential, for example depending on more mature programming skills, as just implied. A degree of engagement, even embodiment, is suggested, and it is quite possible that this can be different in nature or extent from that with live algorithms. Live coding may exploit and develop implicit understanding and processes, whereas to be effective with a live algorithm, a user at least has normally to retain explicit knowledge of how it functions. Arguably, the distinction between these two cases, live algorithms and live coding, is not absolute, but rather they represent separable zones in one or more continua. Future potentials in both fields are appealing.

9. *Conclusion*

I have argued that interactive and generative algorithms are a developing component of music making. The algorithms may be based on mathematical structures applied to music, on statistical analyses of prior or ongoing music, or on cognitive understanding and modeling. I suggest that it is in the last of these areas that some major developments may occur, perhaps bringing together the strands of live algorithm and live coding, of sonification and neurophysiology, of engagement and embodiment, and of acoustic and electronic sound sources.

10. *Acknowledgements.*

I would like to thank Alex McLean (slub and the University of Leeds, UK) for insightful discussion, and our editor, PerMagnus Lindborg, for valuable suggestions.

References

Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036.

Bailes, F., & Dean, R.T. (2012). Comparative time series analysis of perceptual responses to electroacoustic music. *Music Perception, 29*, 359-375.

Bown, Oliver. (2011). Experiments in modular design for the creative composition of live algorithms. *Computer Music Journal, 35*(3), 73-85.

Brown, Andrew R, & Sorensen, Andrew. (2009). Interacting with generative music through live coding. *Contemporary Music Review, 28*(1), 17-29.

Collins, Nick, McLean, Alex, Rohrhuber, Julian, & Ward, Adrian. (2003). Live coding in laptop performance. *Organised Sound, 8*(03), 321-330.

Cope, D. (2001). *Virtual Music. Computer Synthesis of Musical Style.* Cambridge, Mass.: MIT Press.

Cope, David H. (2010). Recombinant music composition algorithm and method of using the same: Google Patents.

Dean, R.T. (2003). *Hyperimprovisation: Computer Interactive Sound Improvisation; with CD-Rom*. Madison, WI: A-R Editions.

Dean, R.T., & Bailes, F. (2008). Is there a 'Rise-fall temporal archetype' of intensity in electroacoustic music? *Canadian Acoustics, 36*(3), 112-113.

Dean, R.T., & Bailes, F. (2010a). The control of acoustic intensity during jazz and free improvisation performance: possible transcultural implications for social discourse and community. *Critical Studies in Improvisation, 6*(2), 1-22.

Dean, R.T., & Bailes, F. (2010b). A rise-fall temporal asymmetry of intensity in composed and improvised electroacoustic music. *Organised Sound, 15*(2), 148-159.

Dean, R.T., & Bailes, F. (2010c). Time Series Analysis as a Method to Examine Acoustical Influences on Real-time Perception of Music. *Empirical Musicology Review, 5*, 152-175.

Ebcioğlu, Kemal. (1988). An expert system for harmonizing four-part chorales. *Computer Music Journal, 12*(3), 43-51.

Farrell, S., & Lewandowsky, S. (2011). *Computational Modeling in Cognition: Principles and Practice*. Thousand Oaks, CA.: Sage Publications, Inc.

Friberg, Anders, Colombo, Vittorio, Frydén, Lars, & Sundberg, Johan. (2000). Generating musical performances with Director Musices. *Computer Music Journal, 24*(3), 23-29.

Hill, Adam J, Lewis, Simon P, & Hawksford, Malcolm OJ. (2012). Towards a generalized theory of low-frequency sound source localization. *Proceeding of the Institute of Acoustics, 34*, 138-149.

Nierhaus, Gerhard. (2009). *Algorithmic composition: paradigms of automated music generation*: Springer.

Pearce, M.T., & Wiggins, G.A. (2006). Expectation in melody: the influence of context and learning. *Music Perception, 23*, 377-405.

Pressing, J. (1990). Cybernetic issues in interactive performance systems. *Computer Music J., 14/1*, 12-25.

Rosenboom, D. (Ed.). (1976). *Biofeedback and the Arts : Results of Early Experiments*: Aesthetic Research Centre of Canada.

Todd, Neil PM, Rosengren, Sally M, & Colebatch, James G. (2009). A utricular origin of frequency tuning to low-frequency vibration in the human vestibular system? *Neuroscience letters, 451*(3), 175-180.

Wiggins, G. A., Pearce, M. T., & Mullensiefen, D. (2009). Computational modeling of music cognition and musical creativity. In R. T. Dean (Ed.), *The Oxford Handbook of Computer Music* (pp. 383-420). New York: Oxford University Press.

Worrall, D., Bylstra, M., Barrass, S., & Dean, R. (2007). *SoniPy: the design of an extendable software framework for sonification research and auditory display.* Paper presented at the 13th International Conference on Auditory Display, Montreal.

Worrall, David. (2009). An introduction to data sonification. In R. T. Dean (Ed.), *The Oxford Handbook of Computer Music* (pp. 312-333). New York: Oxford University Press.