Artificial Intelligence in *Organised Sound*

EDUARDO R. MIRANDA and DUNCAN WILLIAMS

Interdisciplinary Centre for Computer Music Research (ICCMR), Faculty of Arts and Humanities, Plymouth University, Plymouth PL4 8AA, UK
Email: eduardo.miranda@plymouth.ac.uk; duncan.williams@plymouth.ac.uk

Artificial Intelligence is a rich and still-developing field with a number of musical applications. This paper surveys the use of Artificial Intelligence in music in the pages of *Organised Sound*, from the first issue to the latest, at the time of writing. Traditionally, Artificial Intelligence systems for music have been designed with note-based composition in mind, but the research we present here finds that Artificial Intelligence has also had a significant impact in electroacoustic music, with contributions in the fields of sound analysis, real-time sonic interaction and interactive performance-driven composition, to cite but three. Two distinct categories emerged in the *Organised Sound* papers: on the one hand, philosophically and/or psychologically inspired, symbolic approaches and, on the other hand, biologically inspired approaches, also referred to as Artificial Life approaches. The two approaches are not mutually exclusive in their use, and in some cases are combined to achieve 'best of both' solutions. That said, as *Organised Sound* is uniquely positioned in the electroacoustic music community, it is somewhat surprising that work addressing important compositional issues such as musical form and structure, which Artificial Intelligence can be readily applied to, is not more present in these pages.

1. INTRODUCTION

Artificial Intelligence (AI) concerns the development of human-like intelligence, typically in software. The ultimate goal of many AI approaches is to produce optimal or super-human solutions; this might be to augment or indeed to entirely replace the role of composer, performer or listener. In the context of electroacoustic music, this might mean assisted composition, human-like performance rendering of scores, sound organisation and/or advanced synthesis control. Philosophical questions are often raised by those working with AI in music. For example, in the context of the above applications, a question that is often asked is this: Who is the composer if AI has been used in the creation of new music?

Organised Sound is well placed as the leading platform for electroacoustic composers, sound designers, sonic artists and the like to explore such type of questions in their own practice. This article surveys the approaches to AI that have been taken by researchers in the pages of *Organised Sound* during the last twenty years. The reader is advised that although we are aware of other significant progress in the field – for example, AI has many applications in musical analysis and music education – the objective of this exercise is to evaluate what the journal has captured regarding the developments in this important field of music, and as such these developments are mainly focused on composition and/or interactive performance.

We will assume that the reader has some knowledge of the most prominent AI terminology. Therefore, for the sake of brevity we provide details of working processes only when they are absolutely essential to the context of the paper. The timeline is loosely chronological, and sees two main camps emerge amongst the approaches. We examine these in more detail and compare the approaches, and suggest the likely next steps in this fertile and evolving field.

2. THEN AND NOW

> Composers, musicians and computer scientists have begun to use software-based agents to create music and sound art in both linear and non-linear … idioms, with some robust approaches now drawing on various disciplines. (Whalley 2009: 156)

Much of the work found in the timeline of AI in *Organised Sound* can be classified by approach. For example, work towards compositional models using AI begins with symbolic approaches, using machine learning or human input to determine rules for the creation of tonal music. Such systems have appeal to composers who are familiar with symbolic approaches to music representation, such as composing using traditional classical music notation. Conversely, connectionist approaches, based on 'black box' neural networks with no symbolic musical representation, have seemingly not made huge inroads into the pages of *Organised Sound*. Although the chronology is not strict, broadly speaking work using AI such that the system can self-program begins to surface later in the timeline, and often features some comparison or combination of the two. In any case, such systems include distributed autonomous agents, genetic algorithms, flocking or swarming simulations, and neural networks. Essentially, the latter category includes anything that falls under the banner of simulating

Organised Sound **20**(1): 76–81 © Cambridge University Press, 2015. doi:10.1017/S1355771814000454

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press

some sort of biological process above and beyond that of an initial structural rule-set, which includes the simulation of living beings with cognitive learning and evolving behaviour. In this article, we consider that the former category is a symbolic approach to AI in music, while the latter is a biological approach.

2.1. Symbolic approaches

Symbolic approaches to machine learning in AI are synonymously referred to as traditional AI and GOFAI: Good Old-Fashioned AI. They are concerned with hard-coding a set of rules, which prescribe the behaviour of the machine. The choices for developing the rules are manifold, as Collins points out:

> In case it is still unclear, any algorithmic method might be applied, and this potentially includes all artificial intelligence techniques. The extent to which such algorithms have yet to be harvested makes this an open research area; there are favourite techniques, controlled probabilistic expert systems being a typical route. (Collins 2008: 239)

One such probabilistic example would be a series of rules or constraints whereby the machine is taught to generate note structures within various degrees of aleatoric likelihood. Casey (2001) presents and discusses a system for selection and subsequent organisation of sounds from a database, determined by automated acoustic classification. This system uses a machine learning method to differentiate between musical sounds, genre, speech and environmental sounds, and can be used creatively by matching target sounds to other sounds from the database. This forms the core of a sound-generating method referred to as 'concatenative synthesis'. The application of a particular type of probabilistic logic is illustrated by Casey: that of hidden Markov models, or HMM. These HMM models are common in algorithmic composition tasks, and Casey's application to sound selection and organisation further illustrates that/how they are also suitable for use in electroacoustic composition. HMM represent the likely behaviour of events where a given event can adopt any one of a range of states; in a note-based algorithmic composition these might be note sequences, rhythms and so on. Changes in states are known as transitions, and it is the probability of given transitions that can be used to algorithmically generate sequences of notes based on an analysis of the transitions between states in some selected source material. Visell (2004) introduced the application of a spontaneously organising HMM to the analysis and synthesis of statistically driven stochastic music, based on pattern theory analysis. This is an analysis technique that derives the principles for a generative model from specified source signals. In both cases, material that follows the general structural rules of the source signals can be generated whilst still allowing for variation and variability in the result. Further, Visell also gives some consideration to Artificial Life alternatives, which we will discuss in the next section, specifically addressing the alleged advantage of the (HMM) system thus:

> [the] essential difference is that the standard artificial neural networks do not attempt to model directly the native domain of the signal (time, in the case of sound signals). Consequently, the possibilities for structural refinement based on the analysis of output relative to natural signals are more limited. (Visell 2004: 154)

Indeed, several of the articles we have surveyed document some comparison between these two streams of AI – that is, traditional/symbolic and Artificial Life/biological systems – but consideration of the capability of AI to address structural issues is less common. This is particularly surprising given the well-documented applications of AI to music information retrieval, and something which might well be addressed in the future as a practical research question by those working in the field of electroacoustic music.

Other complex probabilistic systems of sound selection and organisation can also be found, such as the fuzzy-logic-based approach presented by Eigenfeldt and Pasquier (2010). Fuzzy logic allows for degrees of 'reasoning' in the AI, rather than exact Boolean values or a series of gate-based logics for decision making. Eigenfeldt and Pasquier also present a method for sound organisation with a self-organising map, abbreviated as SOM, which is essentially an artificial neural network. In this system, perceptual proximities in sounds timbres are assessed on the Bark scale: twenty-four critical bands of frequencies which are correlated to various psychoacoustic responses; perceived brightness, sharpness and so on. This analysis provides values for similarity that are used to build connections in the SOM, which can then be navigated in a real-time process of sound organisation. Nevertheless, these rules remain predefined and do not evolve autonomically. Once the rules are established and the system has been given the input parameters (in the examples above, a database of sounds), musical results can then be then evaluated by the user. Again, autonomic evaluation (machine learning, genetic algorithms and the like) are not present in traditional symbolic approaches.

2.1.1. Evaluating music produced with AI

Clearly, seemingly simple rule-based systems are able to create new sequence of musical notes or ordered sounds and, in the case of aleatoric rule-sets, a considerable amount of variety. But how can we evaluate the music produced by these approaches? A traditional way to evaluate the success of any artificial intelligence system is the so-called 'Turing Test'. This test allegedly

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press

evaluates whether an AI system has created material which is indistinguishable from material created by a human. The original Turing Test was developed to evaluate computer–generated text, but the more popular procedure which is commonly applied in musical AI is much simplified. The test is often adaptable to music simply by asking listeners the following question: Do you think this piece of music been composed by a human or by a machine? However, beyond the evaluation of rule-based success in symbolic AI systems, the use of AI in music composition tasks raises several further aesthetic and philosophical questions. How do we determine what is good or bad when evaluating the output of such systems? And, indeed, who is the author? Aesthetic issues are far from universal and are not readily evaluated in a systematic or repeatable way. How do we determine authorship in the case of creating new music with such systems? Does the authorship rest with the rule-maker, the rule-programmer or a casual user who chooses new seed material as an input for the system? Or does this duty begin at the selection stage of the process? Is the author the decision maker who evaluates the generated material? Or is the author the machine itself?

Philosophically, Jacob summarises the evaluation of success in such AI as:

> how to program a computer to differentiate between 'good' and 'bad' music. The philosophical issues reduce to the question *who or what is responsible for the music produced?* (Jacob 1996: 157; emphasis in original)

We might then conclude that the issue of aesthetic musical quality in the creation of AI-assisted music is moot. Implicit in Jacob's summary is that we have seen the development of AI in practical terms reach a level where it is possible to model the knowledge base of a human composer, and that the questions regarding such work are solely to do with authorship, authenticity and creativity. When discussing his own AI-based approach to real-time composition, Eigenfeldt (2011), addresses this more directly when he suggests that the act of designing the complexity of interactions between agents is a compositional act in itself.

Eigenfeldt is not alone in wishing to stress the ownership and authorship of the music when AI is involved, though Dahlstedt feels more conflicted in this regard:

> I have a slight feeling I did not write that music, and yet I am quite sure no one else did. I designed the algorithm, implemented it and chose the parameters, and still I feel alienated. (Dahlstedt 2001: 122)

Certainly, many involved in algorithmic composition can find themselves in agreement with either end of this scale of ownership, though the real-time nature of these particular systems is something of a special case. It nevertheless highlights that the difference between structural and performing rules is perhaps a smaller one in the field of *Organised Sound* than one might expect in more traditional music composition, where AI is often employed solely to create human-like performances of music sequenced or scored in an otherwise traditional manner. One conclusion we might draw is that the whole, regardless of parts-composition and parts-performance, is of central importance to those working with electroacoustic music.

2.1.2. Imitation of style with symbolic AI

Let us consider a number of other applications for this type of symbolic AI in such music. When the rule-set can be derived from another input, for example an existing piece of music, rather than being pre-determined in some other fashion (as in the HMM examples given above), the effectiveness of the learning may be judged as a measure of successful imitation in the output.

Systems for mimicking a composers style by training the rule-set in this manner exist and have been used successfully, as documented in Ron Geesin's review of David Cope's *The Algorithmic Composer* (Orton 2000) wherein *Alice* (Algorithmically Integrated Composing Environment) is able to extrapolate rules from source material (and thus, compositional 'style' from material contained in the source database) without the need for the composer to specify a rule-set in advance:

> Cope warns [that] the user should not imagine that composing with Alice is necessarily easier than composing without its aid. The choice of the musical material for the database … is critical … A poorly matched database can only give poor results. (Orton 2000: 155)

In this example the database of source material clearly becomes an important part of the generative music process, and, implicitly, the evaluation of success is in the ear of the beholder.

Another real-time example of style imitation by means of AI can be found in the automatic generation of a musical accompaniment. However, learning in real time requires a more complicated approach to the AI than symbolic approaches alone can afford. For example, Cunha and Ramalho (1999) proposed a system for generic automatic accompaniment that achieves good results by combining a symbolic approach with a neural network. As with the work with SOM by Eigenfeldt and Pasquier, we consider that Cunha and Ramalho's system also falls in to the second category of AI: that of biologically inspired, Artificial Life approaches.

2.2. Biological approaches

Biologically inspired AI include systems using neural networks, distributed agents, genetic algorithms and

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press

flocking simulations, all of which have made their way into the pages of *Organised Sound*. One of the fundamental differences between these and the symbolic approaches documented above is in the learning process. Unlike symbolic approaches, these systems can often continue to adjust their rule-sets, potentially developing further without continued human intervention. To some extent this approach gives the user a way to address the issue of creativity that symbolic approaches found philosophically challenging.

Dahlstedt (2001) introduced a system called *MutaSynth*, which fosters a way to explore interactive composition by modelling basic evolutionary processes through sounds. Here, operators referred to as genetic modifiers are applied to create mutations and variations from parent sounds, before the user selects the outputs for which they have a preference. This preference is analogous to a fitness function in evolutionary biology, and as such can be automated with AI evolutionary models. Indeed, this is an issue that is important to consider in any such approach.

Whalley (2004) describes two possible ways to consider the issue of fitness – evolutionary systems and intelligent software agent systems – with the goal of developing an intelligent machine capable of having a musical, interactive conversation. Whalley settled on intelligent agents, which are devices that can make informed decisions, move within a network, and learn in response to their environment over a software-based evolutionary system. In order for this approach to be conversationally interactive, the system must be able to listen and respond appropriately to human input, as well as to initiate conversations of its own accord. Each interaction the agent experiences will thus enhance its own learning. Musical parameters including tempo, dynamics and acoustic features (e.g. panning or audio effects) are then mapped to performance gestures. An interesting aspect to such a system is the continuous exchange of ideas between human users and the AI, unlike systems which only allow for human interaction at the beginning or the end of the process, such as setting parameters, selecting source materials, evaluating results and making aesthetic decisions about good or bad.

2.2.1. Self-organisation

Self-organisation implies a degree of cognitive ability, in the case of multiple agents, to interact, respond, and create structure on a localised level. Blackwell and Young (2004) described their own system for creating self-organised music by interpreting musical parameters from swarm dynamics, as might be exhibited by flocks of birds, herds of animals or groups of cooperating insects. Swarms are modelled using local interactions between agents or particles, rather than a higher-level control. Moreover, users can interact with the particles of the swarms to influence their behaviour. Again, this shows the use of biologically inspired AI to create a system that can adjust its behaviour in real time, in continuous response to human input.

2.2.2. Imitation of styles with neural networks

As we briefly mentioned in the symbolic approaches category, another use case for such a system is that of the creation of automatic musical accompaniments for a human performer. Cunha and Ramalho (1999) described their application of a neural network for this task. Their system, which has already been trained in harmonic development, is capable of generating real-time accompaniment to songs it has not previously been exposed to by means of a prediction model. Neural networks are well documented in AI as rough models of biological neuronal function and can accommodate a high level of complexity in their functioning. Neural networks developed in response to music – for example, those developed in response to specific source material – present a conceptually different solution to symbolic approaches for generating probabilistic rule-sets. The distinction is that the neural network develops connections rather than strict rules, which may give a unique perspective to systems operating outside of the note-based approach to music creation often taken in the work we survey here. Nevertheless, Cunha and Ramalho note that by the addition of a rule-based tracker to their neural network predictor, the performance of the resulting hybrid model was improved.

2.2.3. Genetic algorithms and other combined approaches

Brown (2004) directly compared the aesthetics of melodies produced by both the symbolic (rule-based) and biological (genetic algorithm) approaches, and suggested that a combination of techniques yielded the most 'aesthetically appropriate' (sic) musical results. Other applications of genetic algorithm techniques to algorithmic composition have also been explored elsewhere in *Organised Sound* by Collins (2002), and Manzolli, Moroni, VonZuben and Gudwin (1999). The former looked into providing control of sound synthesis parameters; the latter, into generating and evaluating chord progressions. Metaphorical musical genes, coding musical or sonic phenomena, are mutated and then evaluated via a fitness function. In the case of Collins's synthesis-driving system, the fitness function ultimately remains the choice of the user, which the author refers to as 'the fitness bottleneck of the human decision-maker'. Manzolli and colleagues also acknowledged the fitness function, but instead they created a statistical function based on an analysis of existing memories, showing another way to incorporate the probabilistic rule-based approaches used in the symbolic AI stream.

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press

Notwithstanding the above approaches, there is also a growing movement in the AI community towards embodiment for such strategies. In *Organised Sound* this can manifest itself, for example, in listening and real-time improvisation, as in the 'prosthetic culture' of Impett (2000), the gesture-driven sonorous objects of Godøy, (2006), and the singing cyborgs of Bosma (2003). The similarity in all of these approaches is the physical embodiment of the AI in a third-party agent which the composer or performer can then interact with. This process in turn affords opportunities to explore the realm of embodied cognition and the interplay between AI and human in musical contexts.

3. CONCLUDING REMARKS

Whilst consciously being non-exhaustive in this review, we nevertheless find that AI has had a definite presence in the pages of *Organised Sound*, with both symbolic rule-based systems and biologically inspired Artificial Life approaches being used to create new work by the electroacoustic community, as well as a number of combined approaches. Artificial Intelligence gives a rich pool for those interested in algorithmic composition to develop new systems and indeed to evaluate the musical effectiveness of their output.

The philosophical questions raised by the use of AI in creating music are also not overlooked in these pages, though the traditional questions which might be used to evaluate the success of AI in such applications are perhaps less relevant to the electroacoustic community. The electroacoustic community seems less sensitive to issues of authentic human performance of classical music and more concerned with the aesthetic results which might be obtained. The issue of whether or not the material generated is readily distinguishable from human output when carrying out such processes entirely by hand is seemingly not relevant. Thus we suggest that the majority of authors consider that the training of the AI, from rule demarcation to source-material selection, still constitutes the process of composition.

Given that *Organised Sound* has become arguably the foremost journal of the electroacoustic community, it is surprising that it lacks papers tackling the problem of overall musical form in composition, though some such work exists: for example David Cope's aforementioned *Alice* system used a musical phrase classification algorithm to enable the generation of music with formal structure and coherence in the compositions. Artificial Intelligence has been well used as an analysis tool to determine and describe musical structure by means of structural representations or acoustic analysis. So we might speculate that the absence of other such structural analysis by AI is because the electroacoustic community is not always so interested in directly addressing note-based music. Iannis Xenakis' UPIC system (Xenakis 1996) for graphic scoring allows for structure in the linking of its pages, which become analogous to the score and the structure in note-based music. UPIC has already been shown in *Organised Sound* to be well suited to learning applications (Nelson 1997; Bourotte and Delhaye 2013) so perhaps a method of training AI with UPIC as the interface would be welcomed by practitioners from the electroacoustic community.

More recently, Artificial Life approaches show that although the application of AI to electroacoustic music creation has yet to reach maturity, it is still an open and growing field of research.

It is not a trivial task to predict what AI might contribute to electroacoustic music within the next twenty years of *Organised Sound*. But we suspect that AI informed and inspired by biology will continue to evolve, in particular developments pertaining to computational neuroscience, where scientists are developing increasingly more sophisticated models of the brain. This research is providing us with better understandings of how our brain works. Such understanding is bound to result in new technological and also theoretical developments for music. Unfortunately, scientific progress on this front has so far been largely insignificant for music, as most of this progress has been on visual processing. The truth is that auditory processing turns out to be fiendishly more complicated than had been previously thought. Consequently, our current understanding of how the brain processes music lags far behind our understanding of other brain functions. Despite a fair amount of research that is being developed within the emerging field of cognitive neuroscience of music, progress so far has been disappointing and profoundly irrelevant to musicians in general, and in particular to the electroacoustic community. One of the problems that we can identify here is that the great majority of scientists working in this field – and, consequently, their respective peer-reviewing community lack knowledge of music, thus the premises for the corresponding experiments are often flawed in terms of criteria for investigating musical functionality. Still, we believe that the future is bright. A better understanding of how the brain listens to sound is bound to lead to new technology for the analysis of electroacoustic music based on neurophysiologic models of our auditory system. For instance, in addition to today's cochleogram, which is based on how our inner ear analyses the spectra of sounds, in the future we might be able to build tools along the lines of a *thalamogram* (Miranda 2010). This analysis tool would give information related to the activity of a functional region of the brain referred to as the thalamus. The thalamus plays an important role in controlling attention: it enables the brain to suppress information in order to focus on particular aspects of incoming sensorial information, including sounds. The thalamogram would reveal salient sound

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press

attributes that might be deemed more important than others in specific musical contexts or under specific conditions. We would envisage the possibility of being able to specify such contexts as analysis parameters for simulating the focus of the thalamus under different contexts or conditions.

We surely need to see more musicians walking in the corridors of neuroscience laboratories if AI contributions to electroacoustic music are to continue to be reported in the pages of *Organised Sound*.

REFERENCES

Blackwell, T. and Young, M. 2004. Self-Organised Music. *Organised Sound* **9**(2): 123–36.

Bosma, H. 2003. Bodies of Evidence, Singing Cyborgs and Other Gender Issues in Electrovocal Music. *Organised Sound* **8**(1): 5–17, doi:10.1017/S135577180300102X.

Bourotte, R. and Delhaye, C. 2013. Learn to Think for Yourself: Impelled by UPIC to Open New Ways of Composing. *Organised Sound* **18**(2): 134–45, doi:10.1017/S1355771813000058.

Brown, A. R. 2004. An Aesthetic Comparison of Rule-Based and Genetic Algorithms for Generating Melodies. *Organised Sound* **9**(2): 191–7, doi:10.1017/S1355771804000275.

Casey, M. 2001. General Sound Classification and Similarity in MPEG-7. *Organised Sound* **6**(2): 153–64.

Collins, N. 2002. Experiments with a New Customisable Interactive Evolution Framework. *Organised Sound* **7**(3): 267–73.

Collins, N. 2008. The Analysis of Generative Music Programs. *Organised Sound* **13**(3): 237–48, doi:10.1017/S1355771808000332.

Cunha, U. S. and Ramalho, G. 1999. An Intelligent Hybrid Model for Chord Prediction. *Organised Sound* **4**(2): 115–19.

Dahlstedt, P. 2001. A MutaSynth in Parameter Space: Interactive Composition Through Evolution. *Organised Sound* **6**(2): 121–4.

Eigenfeldt, A. 2011. Real-time Composition as Performance Ecosystem. *Organised Sound* **16**(2): 145–53, doi:10.1017/S1355771811000094.

Eigenfeldt, A. and Pasquier, P. 2010. Real-Time Timbral Organisation: Selecting Samples Based upon Similarity. *Organised Sound* **15**(2): 159–66, doi:10.1017/S1355771810000154.

Godøy, R. I. 2006. Gestural-Sonorous Objects: Embodied Extensions of Schaeffer's Conceptual Apparatus. *Organised Sound* **11**(2): 149–57, doi:10.1017/S1355771806001439.

Impett, J. 2000. Situating the Invention in Interactive Music. *Organised Sound* **5**(1): 27–34.

Jacob, B. L. 1996. Algorithmic Composition as a Model of Creativity. *Organised Sound* **1**(3): 157–65, doi:10.1017/S1355771896000222.

Manzolli, J., Moroni, A., Von Zuben, F. and Gudwin, R. 1999. An Evolutionary Approach to Algorithmic Composition. *Organised Sound* **4**(2): 121–5.

Miranda, E. R. 2010. Organised Sound, Mental Imageries and the Future of Music Technology: A Neuroscience Outlook. *Organised Sound* **15**(1): 13–25.

Nelson, P. 1997. The UPIC System as an Instrument of Learning. *Organised Sound* **2**(1): 35–42.

Orton, R. 2000. Review of *David Cope: The Algorithmic Composer*. *Organised Sound* **5**(2): 111–16.

Visell, Y. 2004. Spontaneous Organisation, Pattern Models, and Music. *Organised Sound* **9**(2): 151–65.

Whalley, I. 2004. PIWeCS: Enhancing Human/Machine Agency in an Interactive Composition System. *Organised Sound* **9**(2): 167–74.

Whalley, I. 2009. Software Agents in Music and Sound Art Research/Creative Work: Current State and a Possible Direction. *Organised Sound* **14**(2): 156–67, doi:10.1017/S1355771809000260.

Xenakis, I. 1996. Tutorial Article: Determinacy and Indeterminacy. *Organised Sound* **1**(3): 143–55.

https://doi.org/10.1017/S1355771814000454 Published online by Cambridge University Press