Artificial Intelligence through the eyes of Organised Sound

Eduardo Miranda, Duncan Williams

Interdisciplinary Centre for Computer Music Reseach, Plymouth University, Devon, UK

Keywords: Artificial Intelligence, symbolic and artificial life approaches, algorithmic composition

Abstract

Artificial intelligence is a rich and still-developing field with many musical applications. This article surveys the use of artificial intelligence approaches in the pages of Organised Sound, from the first issue to the present day. Often, these approaches are designed with note-based composition in mind, but the research we present finds that artificial intelligence has also had a significant impact in electroacoustic music such as sound analysis and organisation, real-time compostion, and interactive performance-driven composition. Two distinct categories emerge in the literature; philosophically inspired, symbolic approaches, and biologically inspired artificial life approaches, though the two 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 tackling issues of form and structure, which artificial intelligence can

be readily adaptable to, is not more present in these pages. We consider this as well as further hybrid models, as likely avenues for future work.

Introduction

Artificial intelligence (hereafter, 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, sound organisation, or advanced synthesis control. Philosophical questions are often raised by those working with AI, for example, in the context of the above applications, *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 questions in their own practice. This article surveys the approaches to AI that have been taken by researchers in the pages of Organised Sound through the last twenty years. The reader should be 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 focussed on composition and/or interactive performance. We will assume that the reader has some knowledge of the most prominent AI terminology and will provide details of working processes only when 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 (both procedurally and philosophically) before looking forward to likely next steps in this fertile and evolving field.

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)

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 artificial intelligence 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 (scoring in MIDI, for example). Connectionist approaches (for example, black box neural networks with no symbolic musical representation) have seemingly not made huge inroads in 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 feature 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, anything that falls under the banner of simulating 'artificial life', with cognitive learning and evoluving behaviour, above and beyond that of an initial structural rule-set. We consider that the former category is a *philosophically* inspired use of AI, whilst the latter is *biologically* inspired.

Symbolic approaches

Symbolic approaches to machine learning in AI are synonymously referred to as 'traditional' AI, and 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)

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 (Casey, 2001) presents and discusses a system for selection of sound based on automated classification and further for subsequent sound organisation. This system uses machine-learning to differentiate between musical sounds, genre, speech, and environmental sounds, and can be used creatively by matching target sounds to other sounds from a database (concatenative synthesis). The application of a particular type of probabilistic logic is illustrated by Casey, that of hidden Markov models (HMM). Markov models are common in algorithmic composition tasks and Casey's application to sound selection and organisation further illustrates they are also suitable for use in electroacoustic composition. Markov chains represent the likely behaviour of events (in a note-based algorithmic composition these might be note sequences, rhythms and so on), where a given event can adopt any one of a range of states. Changes in states are known as *transitions*, and it is the probability of given transitions which can be used to algorithmically generate sequences of notes based on an analysis of the transitions between states in some selected source material. Visell (Visell, 2004) describes the application of a spontaneously organising HMM to the analysis and synthesis of statistically-driven stochastic music, based on pattern theory analysis (an analysis technique which derives the principles for a generative model from natural signals). In both cases, material which follows the general structural rules of the source signals can be generated whilst still allowing for variation and variability in the result. Further, Vissell also gives some consideration to artificial life alternatives which we will discuss in the next section, specifically addressing the 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)

q. Do we agree with Visell on this point?

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

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 (Eigenfeldt & 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 (SOM), which is essentially an artificial neural network. In this system, perceptual proximities in sounds timbres are assed on the Bark scale (24 '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 pre-defined, 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.

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 near-infinte 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 famous *Turing Test*. This test evaluates whether the artificial intelligence system has created material which is indistinguishable from human material. Originally developed to evaluate computer generated text, the test is adaptable to music by asking the question “has this piece of music been composed (or performed) 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-inputter? 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 in fact the decision maker who evaluates the generated material? Or is the author in fact the machine itself?

Philosophically, Jacob summarizes 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?” (emphasis: original author)

(Jacob, 1996)

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 (Eigenfeldt, 2011), addresses this directly:

"Designing the complexity of interactions between agents is a compositional act" (emphasis: ours)

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)

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 performative 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' sounding 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 of us working with electroacoustic music.

Style 'mimickry' by 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 'apeing' 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)

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

Another real-time example of 'style mimickry' by means of AI can be found in the automatic generation of a musical accompaniment – though learning in real-time requires a more complicated approach to the AI than symbolic approaches alone can afford. For example, Cunha and Ramalho's system for generic automatic accompaniment achieves good results by combining a symbolic approach with a neural network (Cunha & Ramalho, 1999), which, like Eigenfeldt and Pasquier's SOM, falls in to the second category of AI which we commonly find in the field, that of biologically inspired, artificial life approaches.

Artificial life approaches

Biologically inspired AI include systems using neural networks, distributed agents, genetic algorithms, and flocking simulations – all of which have made there 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 gives a way to tackle the issue of creativity that symbolic approaches found philosophically challenging (though combining a symbolic approach with a human 'editor' provided satisfactory results for many, too). Thus, artificial life approaches provide fertile material for algorithmic composers to work with. Dahlstedt introduces *MutaSynth* as a way to explore interactive composition by modelling basic evolutionary processes through sounds (Dahlstedt, 2001). Here, genetic modifiers are applied to create mutations and variations from parent sounds, before the user selects the outputs they have a preference for. This preference is analogous to a fitness function in evolutionary biology, and as such can also automated with AI in evolutionary models, but is an issue which is important to consider in any such approach.

Whalley describes two possible approaches (Whalley, 2004), evolutionary systems and intelligent software agent systems, with the goal of developing a cognisant machine capable of having a musical, interactive conversation. Whalley settled on intelligent agents – devices who can make informed decisions, move within a network, and learn in response to their environment over an 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 other acoustic features (panning, audio effects) are then mapped to performance gestures. An interesting aspect to this system is in 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 (setting parameters, selecting source materials, evaluating results, and making aesthetic decisions about ‘good’ or ‘bad’, for example).

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 describe their own system for creating self-organised music by interpreting musical parameters from swarm dynamics (Blackwell & Young, 2004), as might be exhibited by flocks of birds, herds of animals, or groups of co-operating insects. Swarms are modelled by local interactions between agents (particles), rather than a higher level control, and humans 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 on-the-fly, in continuous response to human input.

Style mimickry part two: 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 described their application of a neural network to this task (Cunha & Ramalho, 1999), such that the system (which has already been trained in harmonic development), is capable of generating novel 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 approximations of brain function and can accommodate a high level of complexity in their dependent layers. Neural networks developed in response to music, for example developed in response to specific source material, would 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 that operate 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.

Genetic algorithms and other combined approaches

Similarly, Brown directly compared the aesthetics of melodies produced by both the symbolic (rule-based) and biological (genetic algorithm) approaches (Brown, 2004), and determined that a combination of techniques yielded the most aesthetically appropriate musical results. Other applications of genetic algorithm (GA) techniques to algorithmic composition have also been explored elsewhere in these pages by Collins (Collins, 2002), and Manzolli et al (Manzolli, Moroni, Von Zuben, & Gudwin, 1999), the former to provide control of sound synthesis parameters, and the latter to generate and evaluate chord progressions. Genes (with a musical or sonic mapping) are mutated (or in the case of the 'fittest' genes, left unmodified) and then evaluated via a fitness function. In the case of Collins' 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 et al also acknowledge the fitness function, but instead create 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.

Looking Forward

Body text goes here

Approaches look into the mind in a neuroscience approach; cognition

How does the brain work – looking to the brain for inspiration "cognitive neuroscience"

A lot of work going on but not much published as yet; ERM suggested that by understanding the way we process sounds we can get new types of analysis what next; which worked which didn't, signal processing tools are often v successful/interesting to the electroacoustic community

We have seen that there are a huge variety of applications for AI in music. Symbolic or traditional approaches (probabilistic algorithmic composition techniques, for example), are often focussed on ‘note-based’ music but some of these developments do have an impact on electroacoustic music and many of the papers in this survey have shown how, for example, HMM can be used for sound organisation and the creation of novel electroacoustic works. Furthermore, artificial life or biologically inspired approaches often allowed for new methods of interactive music creation which might otherwise have been completely impossible to realise. We also find that there is a crossover between both streams and that many of the researchers surveyed have given consideration to one or the other as the platform (and in some cases, found the best aesthetic results by combining rule-based and artificial life systems in a hybrid model).

An area that hasn’t been explored in the survey is the use of neural networks in the fitness function of GA systems..?

Philosophical question: Using AI as a source of *inspiration*? (has not been brought up in the OS papers)

Concluding remarks

Whilst consciously remaining non-exhaustive in this article, we nevertheless find that AI has 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 which document good results. AI 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 – the electroacoustic community is perhaps less sensitive to issues of authentic 'human' performance 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 find that the training of the AI, from rule demarcation to source material selection still constitutes the process of composition.

Given that Organised Sound is arguably the foremost journal in the electroacoustic community, it is surprising that there are no papers tackling the problem of overall musical form in composition, which AI tools could be readily adaptable to, though some such work exists, for example 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. AI 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. 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 these pages to be well-suited to learning applications (Bourotte & Delhaye, 2013; Nelson, 1997) 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 the application of AI to electroacoustic music creation has yet to reach saturation – it is still an open, and growing field of research. The *neuroscience of music* (Miranda, 2010) may provide a fertile area for future work.

[other combined approaches…?]

References

References go here

Blackwell, T., & Young, M. (2004). Self-organised music. Organised Sound, 9(02). doi:10.1017/S1355771804000214

Bourotte, R., & Delhaye, C. (2013). Learn to Think for Yourself: Impelled by UPIC to open new ways of composing. Organised Sound, 18(02), 134–145. doi:10.1017/S1355771813000058

Brown, A. R. (2004). An aesthetic comparison of rule-based and genetic algorithms for generating melodies. Organised Sound, 9(02). doi:10.1017/S1355771804000275

Casey, M. (2001). General sound classification and similarity in MPEG-7. Organised Sound, 6(02). doi:10.1017/S1355771801002126

Collins, N. (2002). Experiments with a new customisable interactive evolution framework. Organised Sound, 7(03), 267–273.

Collins, N. (2008). The Analysis of Generative Music Programs. Organised Sound, 13(03), 237. doi:10.1017/S1355771808000332

Cunha, U. S., & Ramalho, G. (1999). An intelligent hybrid model for chord prediction. Organised Sound, 4(02), 115–119. doi:null

Dahlstedt, P. (2001). A MutaSynth in parameter space: interactive composition through evolution. Organised Sound, 6(02). doi:10.1017/S1355771801002084

Eigenfeldt, A. (2011). Real-time Composition as Performance Ecosystem. Organised Sound, 16(02), 145–153. doi:10.1017/S1355771811000094

Eigenfeldt, A., & Pasquier, P. (2010). Real-Time Timbral Organisation: Selecting samples based upon similarity. Organised Sound, 15(02), 159–166. doi:10.1017/S1355771810000154

Jacob, B. L. (1996). Algorithmic composition as a model of creativity. Organised Sound, 1(3), 157–165. doi:10.1017/S1355771896000222

Manzolli, J., Moroni, A., Von Zuben, F., & Gudwin, R. (1999). An evolutionary approach to algorithmic composition. Organised Sound, 4(2), 121–125.

Miranda, E. R. (2010). Organised Sound, Mental Imageries and the Future of Music Technology: a neuroscience outlook. Organised Sound, 15(01), 13. doi:10.1017/S1355771809990227

Nelson, P. (1997). The UPIC system as an instrument of learning. Organised Sound, 2(1), 35–42.

Orton, R. (2000). David Cope, The Algorithmic Composer. Madison, WI: A-R Editions Inc., 2000. 302 pp. with CD-ROM (Mac format). ISBN 0-89579-454-3. RRP $49.95. Organised Sound, 5(02). doi:10.1017/S1355771800232081

Visell, Y. (2004). Spontaneous organisation, pattern models, and music. Organised Sound, 9(02). doi:10.1017/S1355771804000238

Whalley, I. (2004). PIWeCS: enhancing human/machine agency in an interactive composition system. Organised Sound, 9(02). doi:10.1017/S135577180400024X

Whalley, I. (2009). Software Agents in Music and Sound Art Research/Creative Work: current state and a possible direction. Organised Sound, 14(02), 156. doi:10.1017/S1355771809000260

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