A-Life and Musical Composition: A Brief Survey

Eduardo R Miranda[1] and Peter M Todd[2]

[1]Computer Music Research, Faculty of Technology, University of Plymouth, Drake Circus, Plymouth, Devon PL4 8AA, United Kindgom.

[2]Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Lentzeallee 94, 14195 Berlin, Germany.

eduardo.miranda@plymouth.ac.uk, ptodd@mpib-berlin.mpg.de

Abstract: *There have been a number of interesting applications of A-Life in music, ranging from associating musical notes to the cells of cellular automata, to forging genotypes of musical parameters for generating music using genetic algorithms. From the three approaches surveyed in this paper, only the cultural approach allows for the study of the circumstances and mechanisms whereby music might originate and evolve in virtual communities of musicians and listeners. This approach considers musical systems in the context of the origins and evolution of cultural conventions that may emerge under a number of constraints, such as psychological, physiological and ecological constraints.*

1. Introduction

A-Life (or Artificial Life) is a discipline that studies natural living systems by simulating some of their biological aspects on computers (Langton 1997). The attempt to mimic biological phenomena on computers is proving to be a viable route for a better theoretical understanding of living organisms, let alone the practical applications of biological principles for technology (robotics, nanotechnology, etc.) and medicine. Because A-Life deals with such complex phenomena, it has fostered the development of a pool of research tools for studying complexity, most notably modelling tools based upon the notion of interacting agents. It is interesting though, that these tools are also proving to be useful in fields other than Biology, most notably Social Sciences (Gilbert and Troitzsch 1999), Linguistics (Kirby 2002; Cangelosi and Parisi 2001) and Computer Music (Degazio 1999; Todd 2000; Dahlstedt and Nordhal 2001; Miranda 2002a).

The interactive agent-based modelling tolls developed by the A-Life community provides a rich framework within with to build systems of socially interacting individuals, but not all approaches take on board what is perhaps the most determinant aspect of musical development, namely *social dynamics*. In this paper we review three approaches to using models of interacting agents in music composition with special focus on those systems that do take social dynamics into account.

2. Approaches to using models of interacting agents in music composition

There have been a number of interesting applications of A-Life models of interactive agents in music, ranging from associating musical notes to the cells of cellular automata (Hunt et al. 1991) to forging genotypes of musical parameters for generating music using genetic algorithms (Degazio 1999). We identify at least three approaches to the use of these models for composition: (a) rendering of extra-musical behaviour, (b) genetic algorithm-inspired and (c) cultural.

2.1 Rendering of extra-musical behaviour

First, we can construct models of artificial agents going about their business in their simulated world (e.g., moving around, looking for food, avoiding bumping into rocks and each other, and so on) and as they behave, we convert some aspect of their behavior into sound and listen to them. These agents are not musical in the sense that they are not designed with any musical task in mind. Rather, some sort of "sonification" or "musification" to their behavior patterns is applied in order to see (or hear) what emerges. Their social interactions will affect the music we hear, but the music being produced will not affect their social interactions, nor anything else about their lives; instead, the music is a side-effect of whatever the agents are doing.

A system called *Music Insects*, by Toshio Iwai (1992), is well-known example of this approach. It incorporates a small set of insect-like creatures moving over a two-dimensional landscape onto which a user can place patches of different colours. When an insect crosses a patch of a particular colour, it plays a particular associated note. Thus, once an environment of colour-note patches has been set up, the movements of the insects are translated into sound. By appropriate placement of patches and choice of behavioral parameters of the insects (e.g., their speed and timbre), different musical performances can be created.

In a related but more abstract vein, Miranda (1993), Bilotta and Pantano (2001), and others have explored "musification" of the dynamic spatial patterns created by cellular automata; for a review, see (Miranda 2001). In a cellular automaton, cells (or locations) in a grid (e.g., a two-dimensional environment) can have different states (e.g., the "on" state could be interpreted as "this cell contains an agent"), and the states of cells at one point in time affect the states of nearby cells at the next point in time (e.g., an "on" cell at time t can make a neighboring cell turn "on" at time t+1). As different cells in a two-dimensional field are turned on by the states of neighboring cells according to particular production rules, the overall activity pattern of the cells in this "world" can be converted to sound by further musification rules. Because cellular automata are commonly used to study the creation of complexity and dynamic patterns, their behavior can produce interesting musical patterns as well when sonified.

2.2 The genetic algorithms-inspired approach

A second, more directly musical approach is to let each individual produce its own music or tune as it goes about its existence, and to use this music to determine the survival or reproduction of each agent. The songs present in the population can evolve, as more successful songs lead to greater survival and reproduction of the individuals singing those songs, and hence to more copies of versions of those songs in the next generation. This artificial evolutionary process can lead to more complex or interesting pieces of music if allowed to go on long enough. In models of this type, music production is intrinsic to each individual, rather than merely being a consequence of non-musical behavior as in the previous approach. The music an individual produces has material consequences for its own life in turn, so that in some sense the music matters to the agents. The music produced by an individual in this case is not heard and reacted to by other individuals in the population. Some external almighty critic evaluates the outcome. This critic can be an artificially-designed judge, such as an expert system looking for particular melodic or harmonic developments. Or it can be a human user, listening to songs one at a time or to the music composed by whole the population at once, and rewarding those individuals who produce more pleasing songs, or musical parts, with more offspring. So, although a population of individuals is creating music here, each individual still remains blissfully unaware of what the others are singing, and the truly social element remains lacking from the musical process.

There have been a number systems in which a population of musical agents has been reduced to its bare bones, or rather “genes”: each individual is simply a musical phrase or passage, mapped more or less directly from the individual’s genetic representation, or genotypes. These genotypes are in turn used in an artificial evolutionary system that reproduces modified (mutated and shuffled) versions of the musical passages in the population's next generation, according to how “fit” each particular individual is. Fitness can be determined either by a human listener, as in the *Vox Populi* system (Moroni et al. 1994), or by an artificial critic, as in Spector and Alpern’s (1995) use of a hybrid rule-based and neural network critic to assess evolving jazz responses. Whereas in the former higher fitness are assigned to solos that sound better, in the latter assigns higher fitness to responses that more closely match learned examples or rules.

When human critics are used, these evolutionary systems can produce pleasing and sometimes surprising music, but usually after many tiresome generations of feedback. Fixed artificial critics take the human out of the loop, but have had little musical success so far. What would happen if we unfix the critics and/or replace the human critic by other agents in the artificial world? This is one of the central ideas of the cultural approach, where individuals become both producers and receivers of music.

2.3 The cultural approach

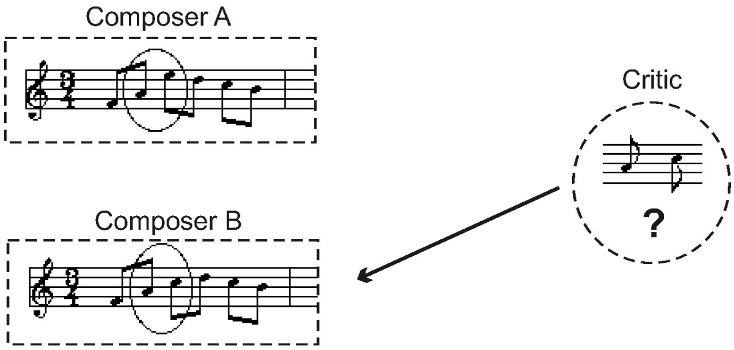
The cultural approach involves actual social interaction on the basis of the music created by individuals. In this case, agents produce musical signals that are heard and reacted to by other agents, influencing for instance the songs that they themselves sing, or their proclivity to mate, or their vigilance in defending their territory. Consequently, the music created in this system affects the behavior of the agents living in this system, giving it a social role. This role is not necessarily the one that this music would have in the human social world—that is, the agents are creating music that is meaningful and effective for their own world, but perhaps not for ours. However, because this system creates music through a social process that is richer than that in the previous two less-social approaches, it could be that the creative products have the potential to be more musically interesting to us, too, as a result.

Inspired by the notion that some species of birds use tunes to attract a partner for mating, Todd and Werner (1999) designed a model that employs mating selective pressure to foster the evolution of fit composers of courting tunes. The model can co-evolve male *composers* who play tunes (i.e., sequences of notes) along with female *critics* who judge those songs and decide whom to mate with in order to produce the next generation of composers and critics. Offspring were then created with a combination of the traits of their parents, and over time both songs and preferences coevolved to explore regions of “melody space” without any human intervention.

Each composer here holds a tune of 32 musical notes from a set of 24 different notes spanning two octaves. The critics encode a Markov chain that rates the transitions from one note to another in a heard tune. The chain is a 24-by-24 matrix, where each entry represents the female’s expectation of the probability of one pitch following another in a song. Given these expectations a critic can decide how well she likes a particular tune. When she listens to a composer, she considers the transition from the previous pitch to the current pitch for each note of the tune, gives each transition a score based on her transition table, and adds those scores to come up with her final evaluation of the tune. Each critic listens to the tunes of a certain number of composers who are randomly selected; all critics hear the same number of composers. After listening to all the composers in her courting-choir, the critic selects as her mate the composer who produces the tune with the highest score. This selective process ensures that all critics will have exactly one mate, but a composer can have a range of mates from none to many, depending on whether his tune is unpopular with everyone, or if he has a song that is universally liked by the critics. Each critic has one child per generation created via crossover and mutation with her chosen mate. This child will have a mix of the musical traits and preferences encoded in its mother and father. The sex of the child is randomly determined and a third of the population is killed at random after a mating session in order not to reach a population overflow.

From the many different scoring methods proposed to judge the tunes, the one that seems to produce the most interesting results is the method whereby critics enjoy being surprised. Here the critic listens to each transition in the tune individually, computes how much she expected the transition, and subtracts this value from the probability that she attached to the transition she most expected to hear. For example, if a critic has a value 0.8 stored in her Markov chain for the A-E transition, whenever she hears a note A in a tune, she would expect a note E to follow it 80% of the time. If she hears an A-C transition, then this transition will be taken as a surprise because it violates the A-E expectation. A score is calculated for all the transitions in the tune and the final sum registers how much surprise the critic experienced; that is, how much she likes the tune. What is interesting here is that this does not result in the composers generating random tunes all over the place. It turns out that in order to get a high surprise score, a tune must first build up expectations, by making transitions to notes that have highly anticipated notes following them, and then violate these expectations, by not using the highly anticipated. Thus there is constant tension between doing what is expected and what is unexpected in each tune, but only highly surprising tunes are rewarded (Figure 1).

Figure 1: The critic selects composer at the bottom because it produced the most surprising tune.



The composers are initiated with random tunes and the critics with Markov tables set with probabilities calculated from a collection of folk-tune melodies. Overall, this model has shown that selection of co-evolving male composers who generate attracting tunes, and female critics who assess these tunes according to their preferences, can lead to the evolution of tunes and the maintenance and continual turnover of tune diversity over time.

Currently the model initializes their Markov chains with coefficients computed from samples of existing tunes. Would it be possible to evolve such expectations from scratch?

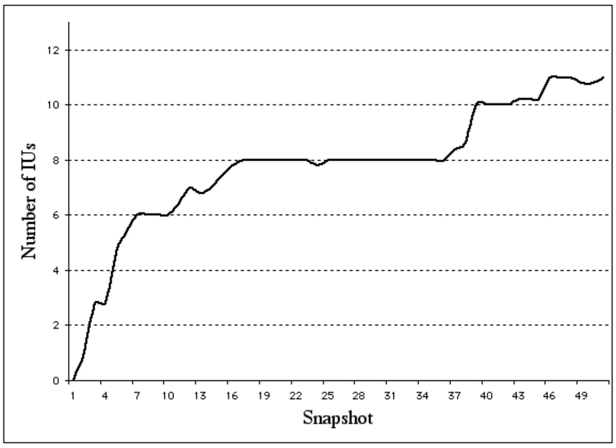
Miranda's (2002b) *mimetic model* may be a plausible alternative to address this question. It demonstrates that a small community of interactive distributed agents furnished with appropriate motor, auditory and cognitive skills can evolve a shared repertoire of melodies, or tunes, from scratch after a period of spontaneous creation, adjustment and memory reinforcement.

The motivation of the agents is to form a repertoire of tunes in their memories and foster social bonding. In order to be sociable, an agent must form a repertoire that is similar to the repertoire of its peers. Sociability is therefore assessed in terms of the similarity of the agents' repertoires. In addition to the ability to produce and hear sounds, the agents are born with a basic instinct: to imitate what they hear.

The agents are equipped with a voice synthesiser, a hearing apparatus, a memory device and an enacting script. The voice synthesiser is essentially implemented as a physical model of the human vocal mechanism. The agent's memory stores its sound repertoire and other data such as probabilities, thresholds and other parameters such as creative willingness, forgetfulness disposition, reinforcement threshold and degree of attention. They have a dual representation of tunes in their memories: a motor map (synthesis) and a perceptual representation (analysis). At each round, each of the agents in a pair from the community plays one of two different roles: the *agent-player* and the *agent-imitator*. The agent-player starts the interaction by producing a tune p_r, randomly chosen from its repertoire. If its repertoire is empty, then it produces a random tune. The agent-imitator then analyses the tune p_r, searches for a similar tune in its repertoire, i_n, and produces it. The agent-player in turn analyses the tune i_n and compares it with all other tunes in its own repertoire. If its repertoire holds no other tune p_n that is more perceptibly similar to i_n than p_r is, then the agent-player replays p_r as a reassuring feedback for the agent-imitator; in this case the imitation would be acceptable. Conversely, if the agent-player finds another tune p_n that is more perceptibly similar to i_n than p_r is, then the imitation is unsatisfactory and in this case the agent-player would halt the interaction without emitting the reassuring feedback; the agent-imitator realizes that no feedback means imitation failure. If the agent-imitator hears the reassuring feedback, then it will reinforce the existence of i_n in its repertoire and will change its perceptual parameters slightly in an attempt to make the tune even more similar to p_r - i.e., only if they are not already identical (refer to the algorithm in the Appendix). Over time the society builds up a repertoire of common musical (or vocal) phrases through their interaction, creating a sort of effective language which, when extended, could provide the basis for composition.

The graph in Figure 2 shows the evolution of the average repertoire of a community of 5 agents after a total of 5000 interactions, with snapshots taken after every 100 interactions. The agents quickly increase their repertoire to an average of between six and eight tunes per agent. At about 4000 interactions, more tunes appear, but at a lower rate. Identical behaviour has been observed in many such simulations with varied settings. The general tendency is to quickly settle into a repertoire of a certain size, which occasionally increases at lower rates. The pressure to increase the repertoire is mostly due to the creativity willingness parameter combined with the rate of new inclusions due to imitation failures. Please refer to (Miranda 2002a and 2002b) more information.

Figure 2: The evolution of the average size of the repertoire of the whole community.



3 Conclusion

A-Life techniques may have varied applications in computer music research. Perhaps the most interesting application is for the study of the circumstances and mechanisms whereby music might originate and evolve in artificially designed worlds inhabited by virtual communities of musicians and listeners. From the three approaches surveyed in this paper, only the cultural approach allows for this study. The cultural approach considers musical systems in the context of the origins and evolution of cultural conventions that may emerge under a number of constraints, such as psychological, physiological and ecological constraints.

References

Bilotta, E., and Pantano, P. (2001). Artificial life music tells of complexity. In E. Bilotta, E. R. Miranda, P. Pantano, and P.M. Todd (Eds.), *ALMMA 2001: Proceedings of the workshop on artificial life models for musical applications* (pp. 17-28). Cosenza, Italy: Editoriale Bios.

Cangelosi, A. and Parisi, D. (Eds.) (2001). *Simulating the Evolution of Language*. London, UK: Springer Verlag.

Dahlstedt, P. and Nordhal, M. G. (2001). Living Melodies: Coevolution of Sonic Communication. *Leonardo*, 34:3, 243-248.

Degazio, B. (1999). La evolución de los organismos musicales. In E. R. Miranda (Ed.), *Música y nueavas tecnologías: Perspectivas para el siglo XXI*. Barcelona, Spain: L'Angelot.

Gilbert, G. N. and Troitzsch, K. G. (1999). *Simulations for the Social Scientist*. Buckingham, UK: Open University Press.

Hunt, A., Kirk, R. and Orton, R. (1991). Musical Applications of a Cellular Automata Workstation. *Proceedings of the International Computer Music Conference – ICMC'91* (pp. 165-166). San Francisco, USA: ICMA.

Iwai, T. (1992). *Music insects*. Installation at the Exploratorium, San Francisco, USA. http://www.iamas.ac.jp/%7Eiwai/artworks/music_insects.html (Commercially available as SimTunes in the SimMania package from Maxis.)

Kirby, S. (2002). Natural Language from Artificial Life. Artificial Life, 8:2,185-215

Langton, C. G., (Ed.) (1997). *Artificial Life: an Overview*, Cambridge, USA: The MIT Press.

Miranda, E. R. (2002a). Emergent sound repertoires in virtual societies. *Computer Music Journal*, 26:2, 77-90.

Miranda, E. R. (2002b). Mimetic Development of Intonation. C. Anagnostopoulou and Smaill, A. (Eds.) *Music and Artificial Intelligence – Second International Conference*, Lecture Notes in Computer Science, Vol. 2445 (pp. 107-118). London, UK: Springer-Verlag.

Miranda, E. R. (2001). *Composing Music with Computers*. Oxford, UK: Focal Press.

Miranda, E. R. (1993). Cellular Automata Music: An Interdisciplinary Music Project. *Interface* (now *Journal of New Music Research*) 22:1, 3-21.

Moroni, A., Manzolli, J., van Zuben, F. and Godwin, R. (1990). Vox Populi: An interactive evolutionary system for algorithmic music composition. *Leonardo Music Journal*, 10, 49-54.

Spector, L., and Alpern, A. (1995). Induction and recapitulation of deep musical structure. *Working Notes of the IJCAI-95 Workshop on Artificial Intelligence and Music* (pp. 41-48).

Todd, P.M., and Werner, G.M. (1999). Frankensteinian methods for evolutionary music composition. In N. Griffith and P.M. Todd (Eds.), *Musical networks: Parallel distributed perception and performance* (pp. 313-339). Cambridge, MA: MIT Press/Bradford Books.

Todd, P. M. (2000). Simulating the evolution of musical behavior. In N. Wallin, B. Merker and S. Brown (Eds.), *The origins of music*. Cambridge, MA: MIT Press.

Appendix: The main algorithm of the enacting script

agent-player - AP	**agent-imitator - AI**
{ IF *repertoire(AP)* not empty pick motor control for p_d; produce p_d; ELSE generate random motor control for p_d; add p_d to *repertoire(AP)*; produce p_d; }	{ analyse p_d } { build perceptual representation; } { IF *rep(AI)* not empty i_n = most perceptually similar to p_d; ELSE generate random motor control for i_n; add i_n to *repertoire(AI);* produce i_n; }
{ analyse i_n; } { build perceptual representation; } { p_n = most perceptually similar to i_n; } { IF $p_n = p_d$ send positive feedback to AI; reinforce p_d in *repertoire(AP)*; ELSE send negative feedback to AI; }	{ IF feedback = positive approximate i_n to p_d perceptually; generate appropriate motor control; reinforce i_n in *repertoire(AI)*; } { IF feedback = negative IF i_n scores good H_T; execute *add_new_similar(snd)*; ELSE Modify motor representation of i_n towards p_d; }
{ execute *final_updates(AP)*; }	{ execute *final_updates(AI)*; }

The *add_new _similar()* function works as follows: the agent produces a number of random intonations and then it picks the one that is perceptually most similar p_d to include in the repertoire.