Emergent Musical Environments: An Artificial Life Approach

Marcelo Gimenes, Eduardo R. Miranda and Chris Johnson

Interdisciplinary Centre for Computer Music Research, Faculty of Technology
University of Plymouth, United Kingdom

{marcelo.gimenes, eduardo.miranda, c.johnson}@plymouth.ac.uk

Abstract. Our research is aimed at investigating the genesis and development of musical styles in artificial worlds. Focusing on the analysis of piano improvisation, we have designed and implemented a computer system (Interactive Musical Environments - iMe) with which we analyse processes involved in music perception and cognition in order to evaluate how musical influence can lead to particular musical worldviews. iMe also entails interaction between software agents and human pianists playing improvised music. In this paper we discuss the main components and algorithms that comprise the system.

Keywords: Musical style, artificial life musicianship, music perception and cognition.

1 Introduction

We are investigating the emergence and development of musical styles in environments inhabited by artificial agents, under the perspective of human perception and cognition. This subject contains important issues belonging both to real and to artificial worlds which relate to the understanding of music-related abilities, such as the interaction and collaboration between man and machine.

Artificial life (*Alife*) tools and techniques are an interesting paradigm that deals with such complex phenomena. Miranda explains that "*Alife* is a discipline that studies natural living systems by simulating their biological phenomena in silico" and "the attempt to mimic biological phenomena on computers is proving to be a viable route for a better theoretical understanding of living organisms..." [10]

A growing number of researchers are developing computer models to study cultural evolution [2], including music and specially how musical structures can originate and evolve in artificially created environments inhabited by virtual communities of musicians and listeners [11]. Impett [7], for instance, generated musical compositions with a system in which agents co-exist and interact in the same world, adapting themselves to the changing environment to which they belong.

Genetic Algorithm procedures are also used for evolving musical materials such as melodies, rhythms and chords. One such example is Vox Populi [8], a system that evolves populations of chords of four notes, through operations of crossover and mutation. GenJam [1] is another known model that simulates a student learning to improvise jazz solos, in which populations of melodic ideas are improved via feedback given by a human mentor, which is used to derive fitness values.

Evolutionary Computation models are also being used in many models. In one of them, CAMUS [9], the emergent behaviour of Cellular Automata (CA) is used to generate musical compositions in which the distances between the notes (in a set of three musical notes) are associated with the co-ordinates of the cells.

We introduce our research in this area through a computer model (Interactive Musical Environments - iMe) that simulates environments where software agents interact among themselves as well as with external agents, such as other systems and human beings. The main purpose of this system is to observe how different musical influences lead to particular musical worldviews in artificial worlds.

In Philosophy of Science, ontogenesis refers to the sequence of events involved in the development of an individual organism from its birth to its death. We therefore use the term musical ontogenesis to refer to the sequence of events involved in the development of the musicality of an individual. In contrast to the models mentioned above, iMe implements a musical ontogenesis model that is useful in studying the influence of musical material learned by artificial musicians, especially in systems for musical composition and improvisation.

iMe's general characteristics were obviously inspired by the real world: agents perform previously defined tasks such as listening to and improvising music, for which they possess perceptive and cognitive abilities. Generally speaking, agents perceive and are influenced by music. This influence is transmitted to other agents as long as they generate new music that is then perceived by other agents, and so forth.

The model enables the design and/or observation of chains of musical influence as is found in human musical apprenticeship. In jazz improvisation, for instance, besides the regular instrumental practice, students are encouraged to listen to a variety of musicians from which they are expected to receive all sorts of "musical material" and that will contribute to their own idiom. On the other hand, by listening to someone's improvisation it is frequently possible to differentiate musical elements that could be attributed to the influences they received during their apprenticeship time.

The ability to distinguish the different music elements that are characteristic of different performers seems to be very natural to human beings. Actually, computer systems also achieve good results at this type of recognition, especially with the use of neural networks ([12], [1], [6]). Similarly, as will be described below, iMe has tools to assess the stylistic distance between pieces of music, and with which the development of musical styles can be evaluated.

Nevertheless this is just one of the topics that we aim to tackle with iMe, the most central of which being the perceptive and cognitive issues involved in musical influence. The way we perceive music and how we organize it in our memory has direct connections with the music we produce. The more we get exposed to certain types of elements, the more these elements get meaningful representations in our memory. The result of this exposure and interaction is that our memory is constantly changing with new elements being added and old elements being forgotten.

Bill Evans, the influential American pianist once commented on the ways he pursued his musical style:

"First of all, I never strive for identity. That is something that just has happened automatically as a result, I think, of just putting things together, tearing things apart and putting it together my own way, and somehow I guess the individual comes through eventually." [4]

In our system it is precisely the description of a certain number of musical elements and the balance between them (differences of relative importance) that define a musical style or, as we prefer to call it, the musical worldview: the musical aesthetics of an individual or of a group of like-minded individuals. All these subjects will be addressed in the next sections.

2 Improvisation

According to a traditional definition, musical improvisation is the spontaneous creative process of making music while it is being performed and, to use a famous analogy, is like speaking or having a conversation as opposed to reciting a written text.

As it encompasses musical performance, it is natural to observe that improvisation has a direct connection with performance related issues such as instrument design and technique. Considering the universe of musical elements played by improvisers, it is known that certain musical ideas are more adapted to be played with polyphonic (e.g., piano, guitar) as opposed to monophonic instruments (e.g., saxophone, flute) or with keyboards as opposed to blown instruments, and so forth.

Since instrument design and technique affects the ease or difficulty of performing certain musical ideas, we deduce that different musical elements must affect the cognition of different players in different ways.

The technical or "performance part" of a musical improvisation is, at the same time, passionate and extremely complex but, although we acknowledge the importance of its role in defining one's musical worldview, our research (and this paper) is focused primarily on how: (i) music is perceived, (ii) music is represented in memory and (iii) the resulting cognitive processes relevant to musical creation in general and, more specifically, to improvisation, conveys the emergence and development of musical worldviews.

Regarding specifically the creative issue, it is important to remember that improvisation, at least in its most generalised form, follows a protocol that consists of developing musical ideas "on top" of pre-existing directives, which are in general a musical theme that comprises, among other elements, melody and harmonic structure. Therefore, in this particular case, which happens to be the most common, one doesn't need to create specific strategies for each individual improvisational session but rather follow the generally accepted protocol.

In spite of the fact that this may seem somehow limiting under the viewpoint of more complex compositional strategies, one of the major interests to the research of improvisational music relies on the fact that once a musical idea has been played, one cannot erase it. Therefore, each individual idea is an "imposition" in itself that requires completion that leads to other ideas, which themselves require completion.

Newly played elements complete and re-signify previous ones in such ways that the improviser's musical worldview is revealed.

In this continuous process two concurrent and different plans play inter-dependent roles: on one hand, a pathway (the "lead sheet") to which the generated ideas have to adapt and, on the other hand, the "flow of musical ideas" that is particular to each individual at each given moment and that imply (once more) their musical worldview.

The general concepts introduced so far are all an integral part of iMe and will be further clarified as we describe the system in the following sections.

3 The System

iMe simulates an environment in which software agents perform music related tasks that convey musical influence. By executing basic tasks such as reading pre-existent music (MIDI files), agents only receive influence while more complex tasks such as improvising music lead to musical influence in both directions: to and from an agent. Interaction can also be established directly with other agents in the system as well as with external systems, including human beings. The main outcome of these interactions is the emergence and development of the agents' musical styles as well as the musical style of the environment as a whole.

In this research we are working within the scope of piano-improvised music. Due to the construction of the perceptive algorithms, iMe is better fitted to a genre of performance (musical texture) in which the pianist uses the left hand to play a series of chords while the right hand plays the melodic line. iMe addresses this type of complex music but also accepts music that could be considered a subset of it, e.g., a series of chords, a single melody or any combination of the two. Obviously any music (genre, etc.) that fits into these categories would generate an optimal response by the system but we are also experimenting with other polyphonic music that goes beyond these constraints in order to consider other aspects of music making and evaluate future improvements to iMe.

Agents should normally act autonomously and decide if and when to interact. Nevertheless, in the current implementation of the system, we decided to constrain this skill in order to have better control over the development of the musical styles: agents can choose which music they interact with but not how many times or when they will interact.

In a very basic scenario, simulations could be designed by simply specifying:

a) a number of agents,
b) a number of tasks for each agent and
c) some initial music material for the interactions.

If one wishes to have more control over the simulations and to observe different developmental routes, some criteria (such as the name of the composer, year of composition, etc.) can be added to constrain the agents' choices. If the agents perform an improvisational task, the new piece is delivered to the system and can be used for further interactions.

iMe generates a series of consecutive numbers that correspond to an abstract time control (cycle). Once the system is started, each cycle number is sent to the agents, which then execute the tasks that were scheduled to that particular cycle.

As a general rule, when an agent chooses a piece of music to read (in the form of a MIDI file) or is connected to another agent to listen to it's music, it receives a data stream which is initially decomposed into several parallel layers, and then segmented as is described below.

4 Perception & Segmentation

Agents initially perceive the music data stream according to a number of "filters" that correspond to basic human sensitive information (e.g., melodic direction or melodic inter-onset intervals). This perception results in a parallel stream of data that is subsequently used for segmentation, storage (memory) and style definition purposes (Figure 1).

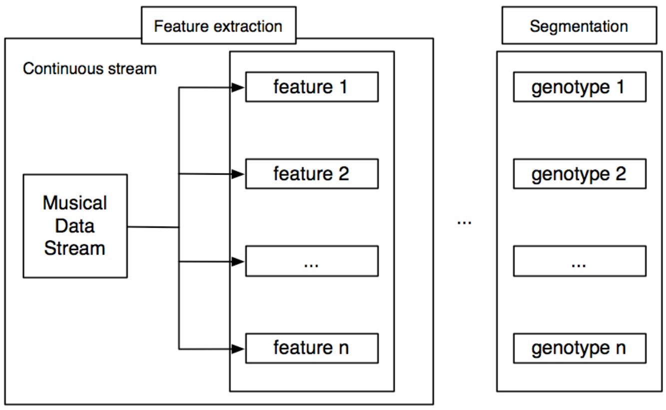

Fig. 1. Feature extraction and segmentation.

The behaviour of each individual parallel stream of data that was perceived by an agent is an input to the segmentation algorithm that takes into account principles taken from Gestalt psychology [5]. Gestalt psychologists argued that perception is driven from the whole to the parts by the application of concepts that involve simplicity and uniformity in organising perceptual information. Proximity, closure, similarity and good continuation are some of these concepts.

We adopted the expression "musical meme" or simply "meme" to refer the segmented musical structures in iMe, a term that has been introduced by R. Dawkins to describe basic units of cultural transmission in the same way that genes, in biology, are units of genetic information. "Examples of memes are tunes, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm and eggs, so memes propagate in the meme pool by leaping from brain to brain via a process which, in a broad sense, can be called imitation." [3]

The idea of employing this concept is attractive because it covers both the concept of structural elements and processes of cultural development, which fits well with the purpose of our research. However, note that we do not embrace all aspects of the original memetic theory, as we believe that it needs to be expanded and/or adapted for our purposes.

In iMe a meme is defined as "the shortest possible musical structure that cannot be segmented without loosing musical meaning", hence Memes are generally small structures in the time dimension. They can have any number of notes vertically, though. By "vertical structure" we mean all music elements that happen at the same time.

Figure 2 shows a possible psychological organisation and consequent segmentation for a musical line (from J.S. Bach's Two Voices First Invention) according to Gestalt laws, in which the inter-onset note interval is taken into account. In this case the same time length separates all except for the first and the last notes, which are disconnected from the previous and the following notes by rests. This implies the application of similarity and proximity rules.

Fig. 2. Musical meme.

Figure 3 shows a meme (from the same piece) and its' genotype representation following the application of three filters: melodic direction (1 is upward and -1 is downward movement), melodic leap (number of semitones) and melodic duration (ratio between current and previous duration).

```
1, 1, 1, -1, 1, -1, 1, -1  (mel. dir.)
2, 2, 1,  3, 2,  4, 7, 12  (mel. leap)
1, 1, 1,  1, 1,  1, 2,  1  (mel. dur.)
```

Fig. 3. Meme and corresponding genotype representation.

Since the memes were previously separated into channels of basic information, they can be described as a set of genotypes, each corresponding to a particular musical feature. By genotype we mean a very basic musical characteristic that is a sub-component of a meme. The meme is therefore represented by 'n' genotypes, in which ‘n’ is the number of musical features that one wants to configure in a given simulation.

So far, we have implemented 10 filters that extract the melodic line (direction, leap, inter-onset interval, duration and intensity) and non-melodic notes (vertical number of notes, note intervals from the melody, inter-onset interval, duration and intensity). The more filters are used, the more accurate is the representation. In any meme the number of elements of all the genotypes is the same and corresponds to the number of vertical structures minus one.

5 Memory

As explained above, the execution of any of the musical tasks requires the perception and segmentation of the musical flow. The next step consists of storing this information in the agent’s memory by comparing with the elements that were previously received. This is a continuous process that constantly changes the state of the agents’ memory.

According to Snyder [13], “the organisation of memory and the limits of our ability to remember have a profound effect on how we perceive patterns of events and boundaries in time. Memory influences how we decide when groups of events end and other groups of events begin, and how these events are related. It also allows us to comprehend time sequences of events in their totality, and to have expectations about what will happen next. Thus, in music that has communication as its goal, the structure of the music must take into consideration the structure of memory - even if we want to work against that structure”.

In our model, the agents’ memory comprises a Long Term Memory (LTM) and a Short Term Memory (STM). The STM is the simplest of the two and stores the ‘x’ memes (‘x’ is defined “a priori” by the user) that were most recently received into the memory. They represent the focus of the “conscious awareness” of the agents.

A much more complex structure, the LTM is a series of specialized “Feature Tables” (FTs) in which all the genotypes are stored according to their category. FTs are formed by “Feature Lines” (FLs) that keep a record of the genotype, the dates of interaction (first reading, last reading and number readings), weight and “connection pointers”. In Figure 4 we present the excerpt of a hypothetical FT (for melody leaps) in which there are 11 FLs. The information between brackets in this figure corresponds to the genotype and the numbers in front of the colon correspond to the connection pointers.

Feature n. 2 (melody leaps):
Line 0: [0 0]: 0 0 0 0 0 0 0 0 0 0
Line 1: [2 2 0 1 0 1 2 5 0]: 1
Line 2: [1 0 0 3 2 2 0]: 2 20 10 10
Line 3: [1 0 0 0 1 2 2 4]: 3
Line 4: [2 0 2 0 4 1 3 0]: 4
Line 5: [0 3 2 7 0 2 0 4]: 5 8 10
Line 6: [3 0 2 0 3 2 4]: 6 5 3
Line 7: [1 0 1 2 2 0 3]: 7 3
Line 8: [2 0 2 0 2 0 0]: 8 31 8
Line 9: [2 0]: 47 4 9 9 4 9 9
Line 10: [5 0 8 2 1 2]: 10

Fig. 4. A Feature Table excerpt.

5.1 Transformation

When a new meme is received in the memory, if the genotype is not present in the corresponding FT, a new FL is created and added to the FT. The same applies to all the FTs in the LTM. The other information in the FLs (dates, weight and pointers) is then (re)calculated.

Figure 5 shows a hypothetical situation in which different genotypes in three different FTs are interconnected via connection pointers. The meaning of these pointers is different depending on the FT. In one of them (chosen by the System's user), called the "first FT", the pointers point to the genotype that was listened to next in the musical piece. In the other FTs, they point to the index of the genotype in the first FT to which they were connected in the original meme. This information will be used during the execution of the improvisation.

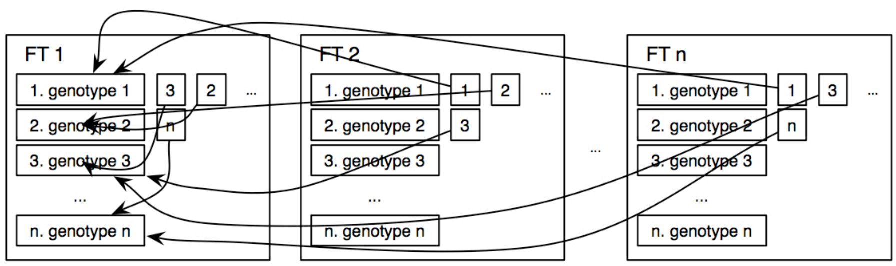

Fig. 5. Interconnection of Feature Tables.

The weight of the genotypes increases or decreases depending on whether it is received by the agent's memory during the tasks. The genotype weight is increased every time it is received and decreases if, at the end of the cycle, it is not received. We can say thereafter that the weights represent the relative importance of the genotypes in relation to each other in a FT.

The calculation of the weights is also affected by a "coefficient of permeability" that the agent's memory possesses at any given moment. This coefficient is defined by a group of other variables (attentiveness, character and emotiveness), the motivation being that some tasks entail more or less transformation to the agent's memory depending on the required level of attentiveness (e.g., a reading task requires less attention than an improvisation task). On the other hand, attributes such as character and emotiveness can also influence the level of "permeability" of the memory.

6 Generative Processes

Agents execute two types of generative tasks: solo and collective improvisations.

Solo improvisations:
During solo improvisations, agents play alone, following these steps:

1. generate a new meme according to the current "meme generation mode"
2. transform the memory with the chosen meme
3. adapt the meme to the Compositional and Performance Map (CPM)
4. play the meme (if in real time mode)
5. store the meme in the composition
6. repeat previous steps until the end of the CPM

The CPM, iMe's equivalent to a "lead sheet", possesses instructions regarding a number of parameters that address both aspects of the improvisation: the generation of new musical ideas and the performance of these ideas. Examples of the former are: the meme generation mode, transformations to the meme, local scales and chords, note ranges for right and left hand. Examples of the latter are: ratio of loudness between melodic and non-melodic notes, shifts for note onset, loudness and duration both for melodic and non-melodic notes. Instructions regarding the performance only affect the sound that is generated by the audio output of the system and is not stored with the composition.

The instructions (or "constraints") contained in the CPM are distributed on a timeline. The agent checks the constraints that are applicable at the "compositional pointer", a variable that controls the position of the composition on the timeline, and acts accordingly.

Collective improvisations:

Collective improvisations are very similar to solo improvisations except for the fact that agents play along with a human being. A simplified way to understand this task would be to consider it as a double task: a listening task and a solo improvisation task executed at the same time. Memes are chosen and transformed as in solo improvisation. The agent's memory is equally affected by the memes agents choose as well as by the memes that they listen to in the musical data that comes from the external improviser. Both agent and external improviser follow the same CPM.

At the end of the improvisation (solo or interactive), the composition is stored in the system in order to be used in further interactions.

7 Results and applications

So far, we have seen that the agent's memory is complex and dynamic, comprising all genotypes, their weights and connection pointers. The execution of musical tasks affects memory state in proportion to the appearance of different memes. As mentioned in the introduction, a particular musical ontogenesis can be objectively associated with the development of an agent's musicality.

Although we understand that it can be virtually impossible to assess the aesthetic value of a musical composition, the assessment of an agent's musical worldview can be done directly by measuring the distance between two different memories. An initial attempt to measure the distance between the memories of two agents was implemented as follows. Firstly, for each genotype in turn, the difference in weights of the genotype in the two memories is computed. These differences are then combined using Euclidean distance to produce a measure of distance between the two memories. Other measures of distance and criteria can be adopted and we are currently experimenting with other possibilities such as the edit (Levenshtein) distance between the most important genotypes (in terms of weight) and the number of connection pointers between the genotypes.

If we run a simulation in which an agent starts with an empty memory, and we give it only one reading task, then at the end of the simulation its memory is its musical style and, also, the musical style of the piece.

Taking this a step further, if we give the agent new reading tasks, the agent's musical style will be transformed so that, after each interaction the new memory state will represent the agent's accumulated experience through the reinforcement or forgetting of the memes. To say that a meme was reinforced or forgotten is actually a simplification of what really happens in our model because, as we have seen above, what is reinforced is not the meme itself but the different elements (e.g., connection pointers) that it comprises.

In another simulation, we could give two different pieces of music to be read by two agents. At the end of the simulation, if their memories were compared, we would have the differences between them and, as a result, between the musical styles of these pieces.

Having these considerations in mind, we can conceive several scenarios and, thus, design different simulations in order to observe the stylistic development during several phases of one (or more) agent's lifetime, or just simply the differences of style between different pieces or sets of pieces of music.

In order to exemplify this process, Figure 6 shows the development of the distance between the musical styles of two agents during 100 cycles in which they listened to the same set of compositions (12 Inventions by J.S. Bach); the horizontal axis represents time cycles and the vertical axis represents style distance. The agents chose the pieces from this set randomly. During the initial cycles we expected that they could interact with different pieces and that there would be major differences in their memories but after a number of cycles (approximately 35) the tendency was towards stability.

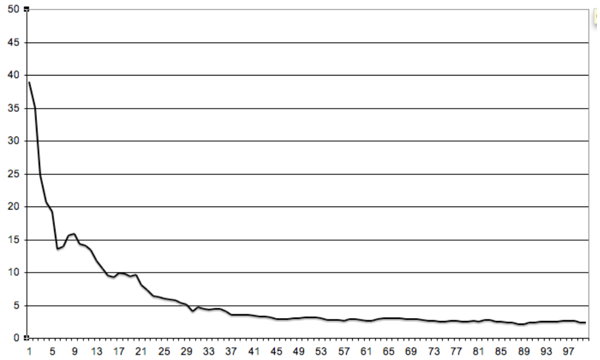

Fig. 6. Style differences between two agents (same set of pieces).

In another experiment (Figure 7 below), during the same number of cycles, one agent only listened to the set of pieces mentioned in the previous paragraph. The other agent, however, listened to a different set comprised by 10 ragtime piano pieces. Figure 8 shows the development of the distance between the musical styles of these agents. The graphs have the same tendency towards stabilization but, of course, because the agents were listening to different sets of music, the final difference is higher in the latter case. We could suggest that the difference of the final distances (approximately 13.5) could be viewed as the difference of the style of these two sets of pieces.

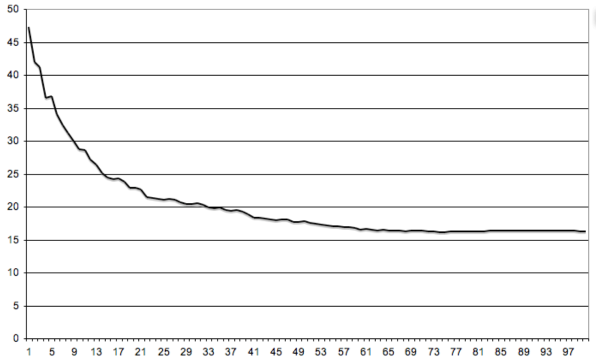

Fig. 7. Style differences between two agents (different sets of pieces).

8 Conclusion

In this paper we have reported the implementation of a computer system (Interactive Musical Environments - iMe) for the investigation of the emergence and development of musical styles in environments inhabited by artificial agents, under the perspective of human perception and cognition.

The initial experiments discussed above are encouraging in the sense that they give good indications of the potential of iME for the study of musical styles and influences. Besides the study of the development of musical styles in artificial worlds, we are also conducting experiments with human subjects in order to assess iMe's effectiveness to evaluate musical influences in inter-human interaction.

The study of creativity and interactive music in artificial and real worlds could also benefit with a number of iMe's features which we are currently evaluating as well.

9 Acknowledgments

This research is funded by the Brazilian Government's Fundacao Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES).

References

1. Biles, J.A. (1994) GenJam: A Genetic Algorithm for GeneratingJazz Solos. Proceedings of Computer Music Conference (ICMC '94), pp. 131-137.
2. Blackmore, S. (1999). The Meme Machine. Oxford: Oxford University Press.
3. Dawkins, R. The Selfish Gene. Oxford, Oxford University Press, 1989.
4. Enstice, W. and Rubin, P. Jazz Spoken Here: Conversations with Twenty-two Musicians. Baton Rouge, LA: LSU Press, 1992. (Bill Evans)
5. Eysenck, M.W. and Keane, M.T. Cognitive Psychology: A Student's Handbook. Psychology Press, 2005.
6. Harford, S. (2003). Automatic segmentation, learning and retrieval of melodies using a self-organizing neural network. In Proceedings of the 4th International Conference on Music Information Retrieval.
7. Impett, J. (2001). Interaction, simulation and invention: a model for interactive music. Proceedings of ALMMA 2001 Workshop on Artificial Models for Musical Applications (E. Bilotta, E. R. Miranda, P. Pantano, and P. M. Todd, eds.), (Cosenza, Italy), pp. 108– 119, Editoriale Bios.
8. Manzolli, J., Moroni, A., von Zuben, F. & Gudwin, R. (2000). Vox Populi: An Interactive Evolutionary System for Algorithmic Music Composition. Leonardo Music Journal, vol. 10, p. 49-54.
9. Miranda, E.R. (1993). Cellular automata music: An inter-disciplinary music project. Interface (Journal of New Music Research), vol. 22, no. 1, pp. 03–21.
10. Miranda, E.R. (1999). The artificial life route to the origins of music. Scientia, 10(1):5-33.
11. Miranda, E.R. (2002). Emergent Sound Repertoires in Virtual Societies. Computer Music Journal, 26(2):77-90.
12. Rowe, R. Interactive Music Systems. MIT Press, 1993.
13. Snyder, B. Bob Snyder. Music and Memory: An Introduction. Cambridge, Massachusetts: MIT Press. 2000.