ON THE LEARNING STAGES OF AN INTELLIGENT RHYTHMIC GENERATOR

Marcelo Gimenes, Eduardo Reck Miranda, Chris Johnson
Computer Music Research, School of Computing, Communications and Electronics,
University of Plymouth, UK
{marcelo.gimenes, eduardo.miranda, c.johnson}@plymouth.ac.uk

ABSTRACT

RGeme (Rhythmic Meme Generator) is an artificial intelligence system for the composition of rhythmic streams. The system is inspired by Richard Dawkin's theory of memes and is based on software agents.

This paper focuses mainly on the first of two broad stages of the system, the learning stage, in which Agents are trained with examples of musical pieces in order to evolve a "musical worldview". During the second stage, the production stage, Agents are able to learn from each other's "compositions" and capable of evolving a new rhythmic style by adapting to each other's rhythms.

The dynamics of this evolution is studied by analysing the behaviour of the memes logged during the learning and the interaction processes. In this paper we present the learning stage of a simulation of the system that uses rhythmic information taken from music compositions by three Brazilian composers, namely Chiquinha Gonzaga, Ernesto Nazareth and Jacob do Bandolim.

Only the learning stage is discussed here. The production stage and its connexions with the learning stage will be introduced in a future paper.

1. INTRODUCTION

RGeme is an artificial intelligence system that uses software Agents for the composition of rhythmic passages. Autonomous Agents are computer systems that inhabit a dynamic and complex environment, sense and act autonomously in this environment executing a series of goals and tasks for which they were devised [12]. These computational entities are designed to have the ability to perceive and to act in their environment in order to achieve certain targets [19].

The framework provided by the theory of memes created by Richard Dawkins is a key component in our system. Memes [4] are 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" ... "Just as genes propagate themselves" ... "memes propagate in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation." ([5], p. 206)

The meme concept entails in RGeme the possibility of observation of the different processes involved in musical influence, the emergence and evolution of different musical styles. According to Meyer, style is "a replication of patterning, whether in human behaviour or in the artefacts produced by human behaviour, that results from a series of choices made within some set of constraints" [15]. By controlling the way software agents interact with each other it is possible to ultimately observe the influence they have over each other.

The "memetic hypothesis" is based on the concept that the understanding that someone has on sounds comes from the comparison with the sounds that this person already produced [3]. The process of comparison involves tacit imitation, or memetic participation that is based on the previous personal experience on the production of the sound.

In the same way that information patterns evolve through biological processes, mental representation, or memes, evolve through the adaptive exploration and transformation of an informational space through variation, selection and transmission [7]. Our minds perform tasks on its replication through an aptitude landscape that reflects internal movements and a worldview that is continuously being updated through the renovation of memes.

Besides this aspect, RGeme was also devised as an aid to composition through the generation of rhythmic streams. Computers have long being used for aiding musical composition in a number of possible ways. Some composers ([2], [6], [20], [21]) use mathematical models such as combinatorial systems, grammars, probabilities and fractals to create new pieces of music. Other systems apply standard Genetic Algorithm procedures for evolving musical materials such as melodies, rhythms, chords, and so on. One such example is Vox Populi [13] which evolves populations of chords of four notes, through the operations of crossover and mutation.

Evolutionary Computation models are also being used in many models. In one of them, CAMUS [16], the emergent behaviour of Cellular Automata (CA) is used to generate musical compositions in which case the co-ordinates of the cells are associated with the distances between the notes of a set of three musical notes.

Impett [9] uses an Agent system to generate musical compositions. Through the interaction of embodied behaviours that co-exist and interact in the same world, Agents are adaptive to the changing environment to which they belong.

A growing number of researchers are developing computer models to study cultural evolution, including musical evolution [1]. For instance, Miranda [17] investigates how musical structures can originate and

evolve in artificially created environments and inhabited by virtual communities of musicians and listeners.

Some rhythmic generating systems have already been proposed [8]. Pachet [18] describes an evolutionary model where a group of agents play rhythms together in real time without prior knowledge about the music to play. Agents play in cycles to which transformation rules are applied in order to produce new variations.

Therefore, our system contributes to this trend by means of computational modelling of a memetic environment and the generation of rhythmic streams.

2. THE SYSTEM

In our system, Agents are able to look for the existence of music compositions and to choose the ones with which they will interact. Later on Agents parse and extract the rhythmic information. Conversely, Agents are also able to actuate in the system through the generation of new rhythmic streams.

At the beginning of a given simulation, a number of Agents are created to which are given an identity (name), a number of tasks ("Goal Matrix") and the criteria ("Evaluation Matrix") they will apply to choose the compositions for interaction ("Candidate Compositions").

Agents can listen to, practice and compose music. As Listeners, Agents only execute listening tasks. In the Student phase, Agents can listen to and practice rhythms. Finally, as Composers, Agents can execute listening, practicing and composition tasks. Broadly speaking these stages and tasks split the model into two general concepts: the learning and the production phases. Evidently, listening and practicing tasks focus mainly on the learning phase whereas composition tasks focus mainly on the production phase.

Before the execution of listening and practicing tasks the Agents choose the Candidate Music according to the Evaluation Matrix (composer's name and/or year of composition). An Evaluation Matrix can determine the same rules for the Agent's entire lifetime or can establish different ones according to the stage in which the Agent is at a specific moment. This last possibility will be employed in the simulation described in the next Section.

Once the Candidate Music is chosen, Agents parse it in order to extract rhythmic memes (Candidate Memes). In the real world, the definition of the exact length/boundaries of a musical meme is a very complex subject for a number of reasons ([10], [11]). Roughly speaking, different individuals can identify different memes in the same or in different pieces of music in accord with, among other factors, their previous personal musical background. Our model, however, was designed to produce musical material in artificially inhabited environments, although it has many features that were inspired in real life situations. Therefore, in order to keep it reasonably simple in the first steps of implementation, currently each rhythmic meme has a fixed length that corresponds to a music bar.

Agents store their musical knowledge in a Style Matrix in which every entry is related to a unique rhythmic structure (rhythmic meme) with the following information:

- the dates (represented in terms of a counter that calculates each interaction cycle) in which the memes were first and last listened to,
- the number of times the memes were listened to,
- the weight (importance) the memes hold due to the various interactions with the Candidate Memes and
- the Candidate Music the meme was listened from.

Style Matrices also hold 'Composition Maps', which correspond to the ways the Candidate Memes are interconnected in the Candidate Compositions.

RGeme represents rhythmic memes coded as vectors whose entries are 0s and 1s (Figure 1), where 1 means the trigger of sounds and 0s are used to represent rests and as time placeholders.

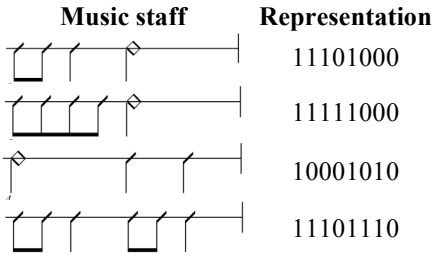


Figure 1. Musical staff and corresponding meme representation

One of the drawbacks of this representation is that information such as the position of the meme in the musical stream, the intensity of each sound and the articulation (duration of the sounds) are not taken into account. Nevertheless, this representation is still useful for the initial implementation of the system and can be extended to include the above-mentioned aspects .

Every time a composition is chosen and the Candidate Memes are parsed, a transformation algorithm is applied as follows. In the beginning of a simulation the Style Matrix is empty and receives the first parsed Candidate Meme. Its weight is set to 1 and the dates of first and last listening are set according to the general time controlled by the system. The second Candidate Meme is then compared with the first meme in the Style Matrix. If they are different the Candidate Meme is copied to the Style Matrix and its weight is set to 1. Subsequently, the other memes in the Agent's Style Matrix have their weight upgraded according to their distance to the Candidate Meme being compared.

The distance between two given memes a = [a1, a2, ... an] and b = [b1, b2, ... bn], is defined as:

$$d(a,b) = \frac{1}{n}\sum_{i=1}^{n} | a_i - b_i | \quad (1)$$

Equation 1: Measure of distance.

For example, the distance between the memes a = 01011101 and b = 11011101 is d(a, b) = 0.125 while the distance between the memes a and c = 11111111 is d(a, c) = 0.375. This is, roughly speaking, the so-called block distance.

Once all Candidate Memes are compared with the memes in the Style Matrix, the new memes are copied and the remaining ones have their weight upgraded, a forgetting effect is applied to the memes in the Style Matrix that don't appear in the current Candidate Memes set. The next section provides a better understanding of how the transformation algorithm alters the musical knowledge possessed by the Agents as a result of the execution of the listening and practicing activities.

Since the learning phase is the main focus of this paper, it suffices to say about the production phase that Agents execute composition tasks mainly through the reassignment of the various Composition Maps according to the information previously stored in the learning phase. Composition tasks, beyond the production of new material, also have a transformation effect on the Style Matrix where all memes are updated according to the musical material used in the newly produced rhythms.

The model has the potential to execute intricate simulations with several Agents learning at the same time from rhythms by composers from inside and outside the system's environment.

3. A SIMULATION

A group of 29 pieces by Brazilian composers Chiquinha Gonzaga, Ernesto Nazareth and Jacob do Bandolim was selected. RGeme was configured to create only one Agent (Agent 'L') to which a series of 100 tasks was given as shown in Table 1:

	Listen	Practice	Compose
Listener	33	n/a	n/a
Student	33	34	n/a
Composer	0	0	0

Table 1. Agent's 'L' Goal Matrix

The Evaluation Matrix for the first 33 tasks established the choice of only Gonzaga's works. During the following 33 tasks only Nazareth works should be chosen and in the following 34, only Bandolim's works, as shown in Table 2:

	Composer	Year begin	Year end
Listener	CGonzaga	-	-
Student	ENazareth	-	-
Composer	JBandolim	n/a	n/a

Table 2. Agent's 'L' Evaluation Matrix

"Year begin" and "Year end" can usually be employed in an Evaluation Matrix to define a date interval. In this case they were not specified which means that the algorithm returned all the compositions by each one of the mentioned composers. In each time period Agent 'L' performed a task consisting of: one Candidate piece of Music was chosen, the Candidate Memes were parsed and the Style Matrix was transformed according with the transformation algorithm. The system generated a new Style Matrix after the accomplishment of each task and all the resulting Style Matrices were logged in the system in order to observe the behaviour of each meme during the interaction processes.

In order to make these concepts more clear, we present in the following paragraphs two extracts from the Style Matrices generated by the system. In Table 3 we show the Style Matrix after an Agent listened to the first music ('Lua Branca', by composer Chiquinha Gonzaga):

#	Meme	dFL	dLL	nL	W
1	00000111	1	1	1	1.026
2	11111111	1	1	18	1.036
3	10100111	1	1	15	1.035
4	10100011	1	1	3	1.030
5	11111010	1	1	6	1.025
6	10000111	1	1	6	1.023
7	10000000	1	1	1	1.000

dFL: date of first listening
dLL: date of last listening
nL: number of listening
W: weight

Table 3: Extract from 1st Style Matrix

Table 4 shows the corresponding data in the Style Matrix after Agent 'L' listened to the second music ('Gaúcho', by same composer):

#	Meme	dFL	dLL	nL	W
1	00000111	1	1	1	1.024
2	11111111	1	2	28	1.062
3	10100111	1	1	15	1.034
4	10100011	1	1	3	1.030
5	11111010	1	2	10	1.053
6	10000111	1	1	6	1.024
7	10000000	1	2	4	1.020
8	01111111	2	2	2	1.023

9	01011111	2	2	5	1.021
10	11011101	2	2	1	1.021
11	10011000	2	2	1	1.022
12	10001010	2	2	1	1.016
13	10001000	2	2	9	1.015
14	01011010	2	2	5	1.011
15	10101010	2	2	1	1.008
16	11011111	2	2	1	1.005
17	10000010	2	2	1	1.004

Table 4: Extract from 2nd Style Matrix

Notice, for example, that:

- After the first cycle of interaction (Style Matrix 1 or SM1), meme 11111111 (second in the list) had been listened (nL) 18 times and its weight (W) was 1.036. After the second cycle of interactions (Style Matrix 2 or SM2) its number of listening was 28 and its weight had been raised to 1.062.
- In SM1 meme 10100111 (3rd in the list) was listened 15 times and its weight had been set to 1.035. In SM2, the number of listening had been kept the same (it was not listened at the second composition) and the weight dropped to 1.034 meaning that a 'forgetting' effect was applied because it was not listened in two consecutive interaction cycles.
- Meme 10000010 (last in Table 4) only appears in SM2 and its weight was set to 1.004, which means that, after its first appearance (in which the weight had been set to 1), the weight raised due to the comparisons made with the other memes that were listened to afterwards.

After the completion of the simulation, we observed that, during the first 33 tasks, Agent 'L' learned a total of 31 memes from the music by Gonzaga. In the second third of the simulation, 63 memes were learned, which indicates that 32 new memes originated from the music by Nazareth. In the last third of the simulation, Agent 'L' learned 24 new memes from the music by Bandolim. Figure 2 shows the evolution in time of the number of memes that were learned by Agent 'L'.

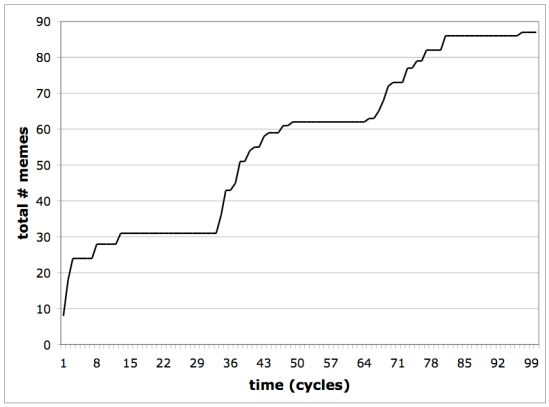


Figure 2. Number of memes learned in time

It was also possible to observe the number of times that each one of the memes was listened to by Agent 'L'. The next Figure shows this number for the first learned 20 memes.

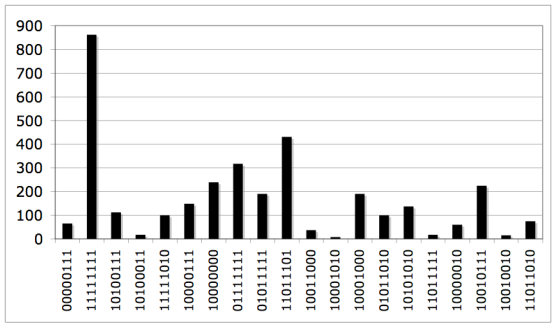


Figure 3. First 20 learned memes: number of listening.

One of the most interesting features that RGeme generates is the track of the evolution of the importance (weight) of each one of the memes during the learning phase of an Agent. The increase or decrease of the importance of the memes is the direct result of the number of times and the date they were listened to and/or practiced. The next Figure shows this analysis during Agent 'L's whole learning phase. Every time a new meme was learned a new line appeared. If a meme was not heard during a certain time, its curve started to fall (forgetting effect).

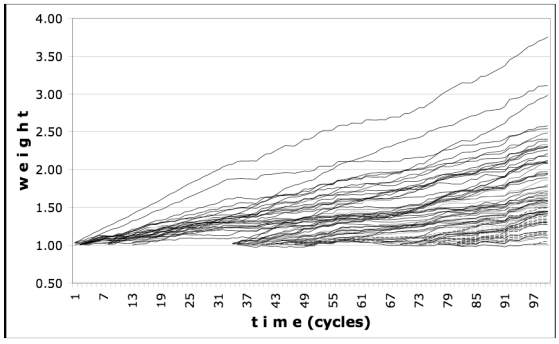


Figure 4. Memes curve of importance in time.

As it is obviously very difficult do visualize the evolution of all the 87 memes in the same graph, in Figure 5 we made a selection of a few of them. Some typical behaviour that emerged from the interactions is described in the paragraphs below.

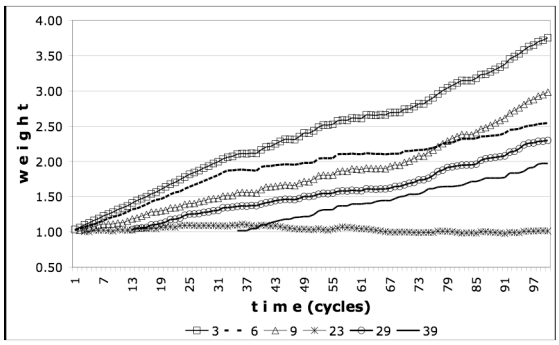


Figure 5. Memes curve of importance in time (selection)

In Figure 5, each series corresponds to the memes described in Table 5:

#	Meme	dFL	dLL	nL	W
3	11111111	1	100	862	3.753
6	11111010	1	91	100	2.543
9	01111111	2	100	318	2.982
23	00100010	3	57	14	1.013
29	10111111	13	95	51	2.297
39	11011000	35	98	69	1.970

Table 5. Description of memes

Agent 'L' listened to memes 3 and 6 in the first interaction (dFL = date of first listening) with music 'Lua Branca', by Gonzaga. Meme 9 appeared in the second interaction (music 'Gaucho', same composer), meme 23 in the third (music 'Annita', same composer), meme 29 in the 13th (music 'Atraente', same composer) and meme 39 in the 35th (music 'Tenebroso', by Nazareth).

Although meme 23 begun to be listened to in time 3, its relative importance comparing to the other memes was never very high. On the other hand, meme 39 was only listened to in time 35 and, at the end of simulation was victorious over meme 23.

Meme 6 was relatively important in the music by Gonzaga but its performance was less significant after Agent 'L' begun to listen to the music by Nazareth. For this reason, at the end of the simulation, meme 9 was victorious over meme 6. Memes 3 and 9 had a steady and comparable performance during the whole simulation. At the end, meme 3 was the winner over all the others. Table 6 shows the 10 most relevant memes at the end of the simulation.

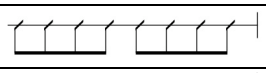
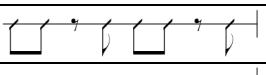
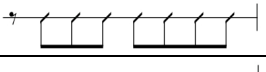
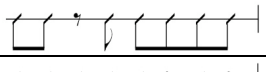
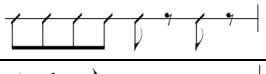
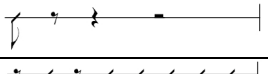
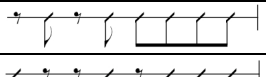
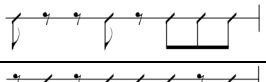
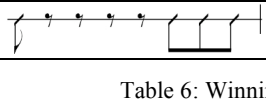
#	Meme	dFL	dLL	nL	W
3		1	100	862	3.753
11		2	98	431	3.111
9		2	100	318	2.982
17		2	92	18	2.577
6		1	91	100	2.543
8		1	100	240	2.482
10		2	98	191	2.422
19		3	94	225	2.402
38		35	98	161	2.381
7		1	99	149	2.329

Table 6: Winning memes

As previously mentioned, if at any given moment Agent 'L' was requested to perform a composition task, the relative importance of the various rhythmic elements should be, together with the corresponding composition maps, a decisive element in the choice of the musical material.

4. CONCLUSION

In this paper we introduced the learning stages of RGeme, an artificial intelligence system for the composition of rhythmic streams.

Besides the production stage that will be covered in a future paper, RGeme has already proved to be an efficient tool to evolve rhythmic worldviews in artificially inhabited environments. Through the description of a simulation we demonstrated how the exposure to different rhythmic material could ultimately shape the musical "knowledge" of an agent.

Experiments are being carried out with different sources of data according to musical genres and styles.

In the future, besides the rhythm information that is being currently employed, the system will deal with more complex musical structures that consider note information (pitches and vertical structures). A better parsing algorithm is being tested in order to extract memes of varied length and a new measure of distance is also being implemented.

5. ACKNOWLEDGEMENTS

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

6. REFERENCES

[1] Blackmore, S. The Meme Machine. Oxford: Oxford University Press. 1999.

[2] Cope, D. Computers and Musical Style. Oxford, UK: Oxford University Press, 1991.

[3] Cox, A. The mimetic hypothesis and embodied musical meaning. MusicæScientiæ, V(2):195–212. 2001.

[4] Dawkins, R. The Blind Watchmaker. Penguin Books, London, 1991.

[5] Dawkins, R. The Selfish Gene. Oxford University Press, 1989.

[6] Dodge, C. and Jerse, T. Computer Music. London, UK: Schimer Books, 1985.

[7] Gabora, L. M. The origin and evolution of culture and creativity. Journal of Memetics–Evolutionary Models of Information Transmission

[http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/gabora l.html]. 1997.

[8] Horowitz, D., “Generating Rhythms with Genetic Algorithms”, Proceedings of the 1994 International Computer Music Conference, San Francisco, ICMA, 1994.

[9] Impett, J. “Interaction, simulation and invention: a model for interactive music,” in 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, 2001.

[10] Jan, S. “Replicating sonorities: towards a memetics of music”, Journal of Memetics– Evolutionary Models of Information Transmission, 4. 2000.

[11] Jan, S. Meme Hunting with the Humdrum Toolkit: Principles, Problems and Prospects. Computer Music Journal (2004) 28(4):68-84.

[12] Maes, P. Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back, MIT Press, 1991.

[13] Manzolli, J., Moroni, A., von Zuben, F. and Gudwin, R. Vox Populi: An Interactive Evolutionary System for Algorithmic Music Composition. Leonardo Music Journal, vol. 10, p. 49-54. San Francisco, USA, 2000.

[14] Martins, J. M., Gimenes, M., Manzolli, J. and Maia Jr, A. "Similarity Measures for Rhythmic Sequences", Proceedings of the 10th Brazilian Symposium of Musical Computation (SBCM). Belo Horizonte (Brazil), 2005.

[15] Meyer, L.B. Style and Music: Theory, History, and Ideology. Philadelphia: University of Pennsylvania Press, 1989.

[16] Miranda, E. R. “Cellular automata music: An inter-disciplinary music project,” Interface (Journal of New Music Research), vol. 22, no. 1, pp. 03–21, 1993.

[17] Miranda, Eduardo R. “The artificial life route to the origins of music”. Scientia, 10(1):5-33. 1999.

[18] Pachet, F. “Rhythms as emerging structures”, Proceedings of 2000 International Computer Music Conference, Berlin, ICMA, 2000.

[19] Russell, S.J. and Norvig, P. Artificial Intelligence: A Modern Approach. Prentice Hall, 2002.

[20] Worral, D. “Studies in metamusical methods for sound image and composition,” Organised Sound, vol. 1, no. 3, pp. 183–194, 2001.

[21] Xenakis, I. Formalized Music: Thought and Mathematics in Composition. Bloomington (IN), USA: Indiana University Press, 1971.