University of Huddersfield Repository

Velardo, Valerio and Vallati, Mauro

On the Stylistic Evolution of a Society of Virtual Melody Composers

Original Citation

Velardo, Valerio and Vallati, Mauro (2015) On the Stylistic Evolution of a Society of Virtual Melody Composers. In: Evolutionary and Biologically Inspired Music, Sound, Art and Design: 4th International Conference, EvoMUSART 2015. Springer, pp. 249-260. ISBN 9783319164977

On the Stylistic Evolution of a Society of Virtual Melody Composers

Valerio Velardo[1] and Mauro Vallati[2]

[1] School of Music, Humanities and Media, University of Huddersfield, UK
[2] School of Computing and Engineering, University of Huddersfield, UK

Abstract. In the field of computational creativity, the area of automatic music generation deals with techniques that are able to automatically compose human-enjoyable music. Although investigations in the area started recently, numerous techniques based on artificial intelligence have been proposed. Some of them produce pleasant results, but none is able to effectively evolve the style of the musical pieces generated.
In this paper, we fill this gap by proposing an evolutionary memetic system that composes melodies, exploiting a society of virtual composers. An extensive validation, performed by using both quantitative and qualitative analyses, confirms that the system is able to evolve its compositional style over time.

Keywords: stylistic evolution, melody generation, memetic approach, computational creativity

1 Introduction

Automatic generation of music is a new exciting area of computational creativity. Many techniques to generate music have been developed, which draw upon several approaches of artificial intelligence, such as evolutionary algorithms, machine learning and expert systems [9]. Even though some of these methods produce music which can be deemed as pleasant by human listeners, none of them is actually capable to convincingly evolve its compositional style. The main problem is that systems are usually based on a single agent, whose compositional process is predetermined by the programmer, and cannot be changed. To avoid that pitfall, a number of systems characterised by a society of software agents, which exchange information have been implemented. Pachet [21] designed a system in which a society of agents play rhythms together, creating new variations of rhythmic passages, according to transformation rules. Gimenes et al. [11] expanded the initial idea of Pachet, proposing a society of software agents, which generate rhythmic passages, following an evolutionary process based on memetics. However, both methods described are focused only on rhythm, so that their outputs cannot be defined as music. Miranda [20] solved the issue designing a society of composers, which interact with one another and develop a shared repertoire of melodies. The system is effective, but the agents do not evolve the way they produce melodies.

In this paper, we introduce an evolutionary memetic system, which generates melodies and is capable of evolving its compositional style over time. To obtain that, we developed a mixed approach, which simulates both the psychological and the social levels of human composers [22]. The system is the result of a number of virtual composers connected together, which generate melodies, exchange them with their fellow composers, and evolve their compositional style, thanks to the influence of other agents.

The method has been extensively validated using quantitative analysis and music experts. Results suggest that the system effectively evolve its own compositional style. This not only is the first case of stylistic evolution we are aware of, but it also demonstrates that computers are potentially capable of evolving the style of the creative artifacts they generate. In this regard, the system is one of the first computational techniques characterised by a primitive form of transformational creativity [1], i.e., the ability to generate artifacts that completely transcend a given conceptual space. Moreover, some of the findings that emerge from the experiments can be easily extended to the real-world musical environment. This is the case of *musical attractors*, which are specific stylistic configurations shaped by cultural and cognitive constraints, likely to be positively assessed by human listeners.

The remaining paper is organised as follows. First, we provide the musical background and describe the system. Then, we discuss the experimental setup, report results and interpret them. Finally, we give conclusions.

2 Musical background

This section provides introductory information about memes, musical memes (i.e., *musemes*), and musical style; necessary to understand the system.

Memes and musemes. Memes are cultural replicators that spread from person to person within a society [5]. Examples of these are ideas, fashion and technologies. Each meme carries a unit of cultural information, that can be passed from person to person by the means of writing, speech, gestures and rituals. Memes can be regarded as sociocultural analogues to genes [12]. Indeed, just like genetic information undergoes a continuous process of evolution, so memetic information does. In particular, the memetic evolutionary process is characterised by three distinct steps: variation, replication and selection [7]. Random variation introduces novelty within a meme pool. Replication allows a single element to be copied and spread within a population. Selection guarantees that only the fittest memes survive within a specific cultural environment. Since music is a subset of human culture, it is possible to extend the concept of meme to the musical domain. As Jan suggests [15], a musical meme or museme is:

> a replicated pattern in some syntactic/digital elements of music - principally pitch and, to a lesser extent, rhythm - transmitted between individuals by imitation as part of a neo-Darwinian process of cultural transmission and evolution.

For the purpose of this paper, we consider melodic musemes as small monophonic phrases containing 5 to 9 notes. The length of musemes has been chosen accordingly to short-term memory constraints [19]. Musemes are stored within the brain and are actively used for processing/composing music [14]. Considering the finite memory capacity of human brain, a continuous fight for survival happens among musemes. Each person unconsciously selects musemes, based on a personal musical fitness function, influenced both by musical universals [2], and by the musical environment she happens to live in.

Musical style. Musical style is a very loose concept, which has no single agreed definition among academics. Indeed, there are many - sometimes contradictory - theoretical definitions of musical style. For Fabbri [8], musical style is the recurring presence of specific musical events, which is typical of a composer, a place, or a period of time. This definition does not explain how/why stylistic change happens, as well as how/why a composer internalises specific musical events. Meyer solves the issue, by proposing an interdisciplinary definition of style, at the intersection of information theory, cognitive science and music theory [18]. For Meyer, style emerges from the replication of specific musical patterns chosen from a potentially infinite repertoire of musical constructs. The choice of specific patterns over others results from the combination of composer temperament, cognitive constraints and cultural constraints Stylistic change happens thanks to the tension between composer temperament and the cultural environment, which leads to instability, and therefore to the invention of new compositional strategies. Even though the definition proposed by Meyer is powerful, it is difficult to implement into a computational system, in order to extract useful stylistic information from musical pieces. For this reason, numerous operational definitions of musical style have been developed by researchers interested in classifying pieces, based on their style [3, 4, 17, 13, 6]. These definitions consider style as a synthetic metric, which is the combined result of a number of musical features, such as pitch distribution, types of intervals and rhythmic structure.

3 Framework

The system we propose in this paper simulates the compositional process of human composers at two different levels, i.e., psychological and social [22]. The former encompasses all those processes and musical elements that univocally identify the generation process of a single composer. Compositional rules, style and aesthetic judgement are all instances of the psychological level, which differ from musician to musician. The social level considers a composer as a node within a network of composers. This level analyses the way composers change their psychological elements while interacting with each other, thanks to a constant exchange of musical information. The psychological level of the system models how composers generate music, while the social level simulates how composers influence and are influenced by the musical environment.

To implement the psychological level, the systems relies on a mixed top-down bottom-up approach, that was proposed in [22]. The top-down element provides

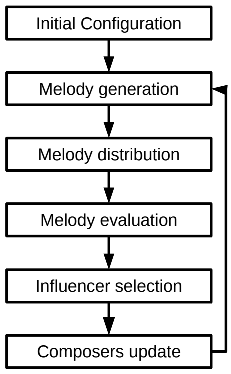


Fig. 1. Steps followed by the system to generate melodies and to evolve its own compositional style.

a coherent hierarchical musical structure for a melody, which is then filled by a bottom-up technique. To emulate the behaviour of a human composer, virtual composers are characterised by three elements: *musical content*, *musical grammar* and *evaluation function*. Musical content is provided under the form of a set of musemes, generated using Markov Chains, trained on 5000 German folksongs of the Essen database. The musical grammar is based on a generative grammar, which is responsible for the development of the musical structure, as well as for the variation of musemes. The evaluation function assesses the quality of a melody according to a linear combination of three musical parameters: *pitch range*, *rhythmic homogeneity* and *step ratio*. Pitch range indicates the difference between the highest and the lowest pitch of a melody. Rhythmic homogeneity measures the overall rhythmic coherence of a melody on a note-by-note basis. Step ratio returns the percentage of steps within a melody. The process of generation of a melody consists of four steps: initiating musical content, generating musical structure, filling the structure, evaluating a melody. After the generation of a symbolic hierarchical structure, a number of musemes are chosen to fill the backbone. These musical phrases can either be left unchanged or modified, depending on the structure. If the quality of the melody is not satisfying, according to the evaluation function, the algorithm goes back to step two and generates a new structure.

All elements of a virtual composer can potentially be modified over time, thus its style can evolve. Indeed, the values of the parameters of the musical grammar and the evaluation function can change, as well as the raw musical content of a composer. In order to model musical evolution as it happens in the real world, a network of virtual composers is necessary. To implement the social level, we use a multi-agent system. A number of virtual composers are tied together, and build up a fully connected network. Each composer can be affected by all other composers of the network. The mutual influence of nodes is guaranteed by a continuous exchange of melodies among elements. Indeed, agents "listen" to melodies created by other agents, and are usually influenced by the music of the composer they think is best. The social level is implemented through a multi-step algorithm (Figure 1). The specific steps are: configuration of composers, generation of melodies, melodies distribution, evaluation of melodies, choice of influencer, update of composers' style.

During the initial configuration, the value of some parameters of the musical grammar and the evaluation function of a composer are randomly seeded. For the evaluation function, the weights of all three parameters of the function (i.e., pitch range, rhythmic homogeneity, step ratio) are randomly chosen. The same happens for the value of three parameters of the musical grammar called *pitch change ratio*, *duration change ratio* and *delete ratio*. These parameters are responsible for the variation of a museme during the bottom-up filling process. Specifically, they indicate the probability that a note contained in a museme can change pitch/duration or can be deleted, when modifying a museme. Furthermore, a set of 500 musemes is generated for each composer, using Markov chains. These musical phrases are the actual musical content composers will use to fill the musical structure of their melodies.

After configuration is completed, each composer generates a melody according to the mixed top-down bottom-up approach described above. Melodies are then distributed across the network, and "heard" by composers. This means that the values of pitch range, rhythmic homogeneity and step ratio are retrieved from each melody. During the evaluation phase, each composer judges all melodies according to its specific configuration of the evaluation function, and stores the melody with the best score. Afterwards, agents choose another agent as a source of influence. The likelihood of the composer of the best melody to be chosen is by far higher than that of other agents. This procedure guarantees that composers are generally influenced by those nodes, whose music they like best.

The final step of the algorithm is the update phase. This process involves both musical content and musical grammar. Each composer randomly picks a museme from the melody it likes best, and substitutes it to one of its own musemes. This ensures a steady change in the musical content, while respecting memory constraints of human composers. In order to guarantee a clear influence in the compositional process, the musical grammar is modified as well. One parameter is randomly chosen among pitch change ratio, duration change ratio and delete ratio. The value of the chosen parameter is changed, and gets closer to the value of the same parameter of the influencer. After completion of the first iteration, the algorithm goes back to step two, in which all composers generate new melodies.

4 Methods

Experiments with different network size have been conducted, to investigate how the number of virtual composers affects stylistic evolution. Precisely, we tested the system with 10, 30 and 100 composers. For each size of the network, 100 experiments were run. The initial configuration of the parameters of the musical grammar and the evaluation function for each composer was derived randomly. When updating the compositional grammar, the value of parameters can be modified up to 30% of their current value, in order to avoid radical changes in compositional style, which are uncommon in the real world.

All experiments aimed to measure the average stylistic evolution of the system over time. However, evaluating stylistic evolution of a composer is a challenging task, since in the first place, there is no agreement on the definition of musical style among scholars. To avoid this issue, we considered two approaches: a quantitative-based analysis, and an expert-based analysis. The latter relies on judgement of music experts, while the former processes large amount of musical data. In particular, we used quantitative-based analysis to track style from three different perspectives, i.e., the evolution of melodic output, musical content and musical grammar.

Style is considered as an emergent quality of a melody, which can be reduced to a handful of musical features. To track the stylistic evolution of melodies over time, we introduce the *melodic style phase space*, which is a 3-dimensional space, whose axes are pitch range, rhythmic homogeneity and step ratio. Specifically, pitch range carries timbral information, rhythmic homogeneity is a synthetic measure for the duration of notes at a local level, and step ratio provides a synthetic information about the intervallic content of a melody. A single point within the melodic style phase space corresponds to a specific melodic style, and conversely, a specific melodic style corresponds to a single point within this space. Of course, musical style cannot be reduced to three musical features. However, working with only three features facilitates the visual representation of stylistic evolution, otherwise impossible, as well as it provides an effective method to reduce/manage the complexity of musical style. To have a dynamic view of the stylistic evolution of the melodies over time, we introduce the *melodic style event space*. In this space, all axes are the same as those of the phase space, except for the pitch range, which is replaced by the number of iterations of the algorithm. The same spaces - event and phase - are used for visualising the evolution of musemes. While melodic spaces consider the actual melodies produced by composers, museme spaces consider musemes stored by composers.

In order to be informed about the evolution of the musical grammar, we introduce the *grammar phase space* and the *grammar event space*. These spaces reflects the internal state of the system, by considering those parameters of the grammar that change over time: pitch change ratio, duration change ratio and delete ratio. For the grammar event space, a combined parameter called *change ratio* is used, calculating the mean of pitch change ratio and duration change ratio, so that an axis can be available to the parameter number of iterations; without losing too much information about the musical grammar. Figure 2 shows the typical evolution of melodic style, musemes and grammar, through the aforementioned spaces, in our experiments. For all types of spaces, the actual metrics we track are obtained considering the average values of all musical features, calculated over the entire population of composers at the end of each iteration.

Consequently, a trajectory within any of the event spaces expresses the average value of the system considered as a whole, and each point represents a picture of the system at a specific point in time.

5 Results

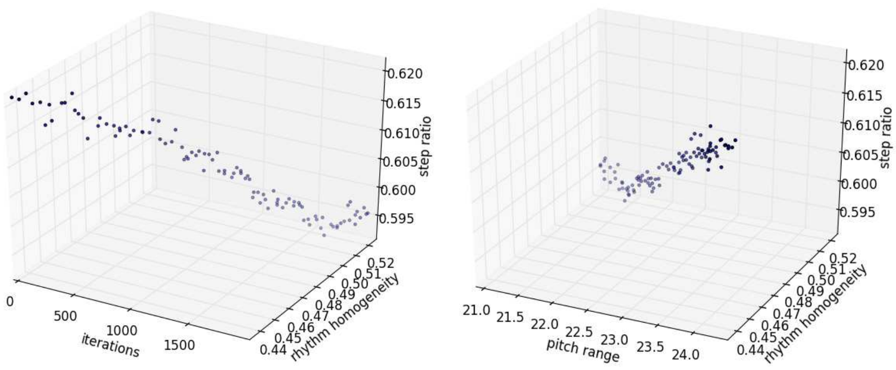
A relevant aspect that emerges from the experiments is the limited impact of differently sized networks of composers. The system with 10, 30 and 100 composers behave the same way.

A quantitative analysis of the style of the melodies generated by the system reveals that there is a continuous coherent change in style after each iteration. A typical example of melodic evolution can be seen in Figure 2. No large jumps have been noticed within the melodic phase space and within the melodic event space. Melodies seems to change a little at each iteration, causing style to evolve slowly over time. Over 60% of the experiments have the stle of melodies converging towards a precise point of the melodic style phase space. However, before converging melodies visit small regions of the phase space. Different runs of the algorithm always visit different regions of the phase space, even though overlaps are frequent. Also, some regions of the phase space are clearly preferred over others, while others have never been visited. This is the case of regions that lie at the extremes of the domain of the musical features considered.

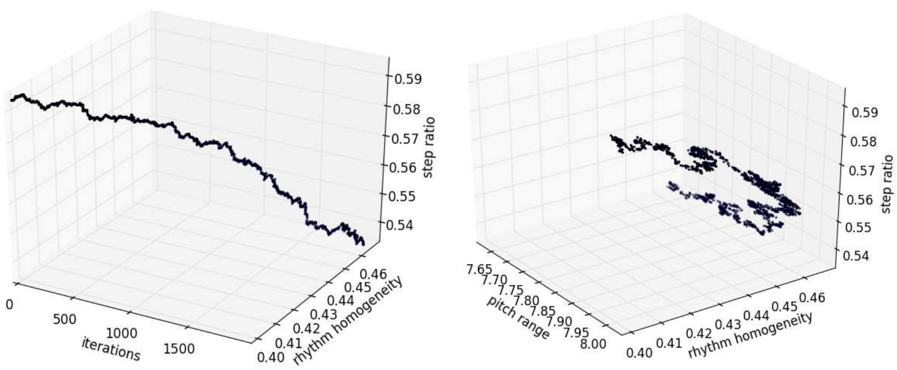
A similar result is obtained by analysing the style of musemes stored by composers. Stylistic evolution of musemes is a slow coherent process, which never shows large jumps (Figure 2). Less than 40% of the experiments converge towards a single point in the museme style phase space. Another major difference with the case of melodies, is that musemes usually anticipate the points of the phase space, that are visited by melodies after a number of iterations. Musemes visit small regions of the phase space, as melodies do. Also, no two different runs of the system visit the same regions of the phase space, and some subspaces such as the extreme of the axes have never been visited.

Results are different with regards to the evolution of the musical grammar. As can be seen in Figure 2, the configuration of the features of the grammar tends to change conspicuously over time, until it converges towards a precise point of the space. Large jumps are always present during initial iterations. These tends to get smaller over time, until the system falls in a certain configuration. After the system converges towards a point in the space, it stays stable for the remaining iterations. Different runs of the algorithm explore different regions of the phase space, and again there are some subspaces which are never visited.

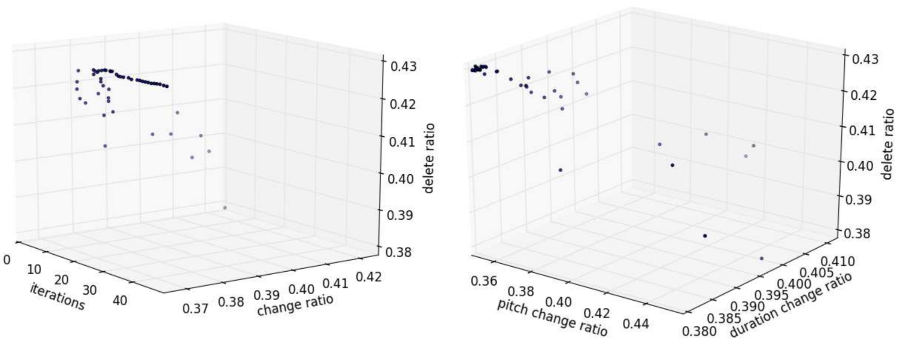
A general question that encompasses many runs of the algorithm at once is: do melodies/grammar/musemes converge always in the same points of their respective phase spaces? The answer is no, since there seem to be preferred regions of the phase space, where the system is attracted into. As it is shown in Figure 3, the region of attraction for the grammar phase space is bigger than that for the museme style phase space.



Melodic style event space (left) and phase space (right).



Musemes style event space (left) and phase space (right).



Grammar event space (left) and phase space (right).

Fig. 2. An example of stylistic evolution for melodies, musemes and grammar; obtained from an experiment with 30 composers.

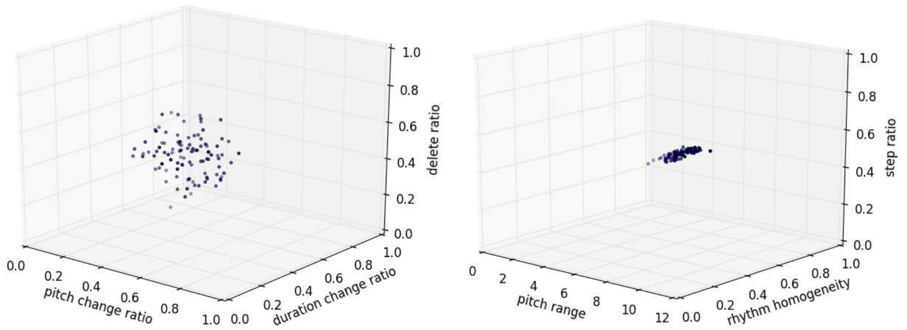


Fig. 3. Regions of convergence for grammar (left) and musemes (right), obtained by plotting the last configuration of the system for 100 experiments with 30 composers.

Since all the metrics we are using are based on the mean of values calculated over the entire population of composers, we have no information about possible local variations of the values for single agents. To have an idea of the dispersion from the average, we calculated the mean of the standard deviations of all nine musical features used to build all three style phase spaces. Figure 4 clearly shows that the standard deviation tends to approach zero after 50 iterations. Moreover, there is no clear difference between the systems with 10, 30 and 100 composers, since they always drop to zero following the same curve.

To assess whether or not the system evolves its compositional style, we also used an expert-based analytic approach. Four music experts (i.e., two composers, one music theorist and one musicologist) assessed ten pairs of melodies produced by the system during ten different runs.[3] Each pair of melodies consists of a melody randomly chosen between those generated at the beginning of a run, and another picked among those composed during the last iteration of a run. Music experts had to evaluate both the aesthetic quality of the melodies and the stylistic difference between members of the same pair. To do that, they used two scales ranged from 1 to 5, where 1 respectively means "no aesthetic value" / "absolute no stylistic difference", and 5 stands for "very high aesthetic values" / "melodies completely different". Stylistic difference obtained a score of 3.25, while aesthetic scored 2.78. Experts said that some melodies are musically interesting, presenting an overall good directionality with nice distributions of pitches and rhythms; while others sound uncreative or even "dull". We used Fleiss Kappa to evaluate inter-rater reliability among experts. Indeed, if experts agree on an assessment, it is likely that their evaluation might be valid. We obtained Fleiss Kappa [10] values of 0.26 for stylistic difference and of 0.32 for aesthetic quality. According to the classical interpretation of Fleiss Kappa, both values indicate a fair agreements among raters [16].

[3] One of the pairs of melodies is available at http://goo.gl/9nHVKl

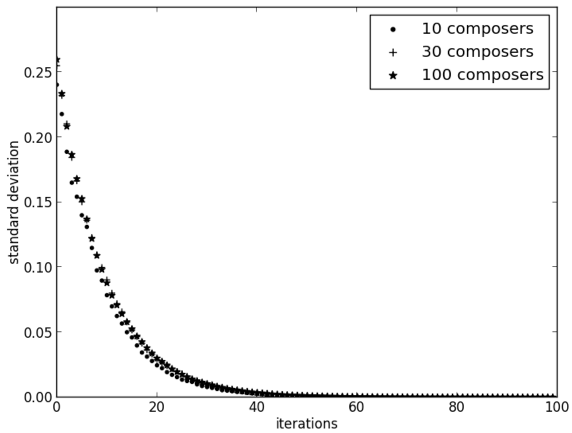


Fig. 4. The mean of the standard deviations of all nine musical features used to build the style phase spaces, for experiments with 10, 30 and 100 composers.

6 Discussion

Both quantitative-based and expert-based analyses confirm that the proposed system is likely to produce a change in the style of the melodies. The specific term to define this process is evolution, since the change is blindly guided by a process of memetic selection, which happens both at the level of the grammar and the musical contents. The system encapsulates the most important aspects needed to have an evolutionary process, i.e., a population of composers which replicate, vary and select a number of musical patterns, while modifying their grammars accordingly to the musical content they like. Indeed, for stylistic evolution to happen, it is necessary to have at least a group of agents which exchange musical information with one another. This process can be regarded as a musical analogue to the biological case. Just as in nature for genetic evolution to occur, a huge number of animals is needed to carry, share, reproduce and select genetic information; so for musical style to evolve, a group of interconnected composers is needed to allow selection and replication of musical ideas. To obtain the same type of evolution with computer generators, the same basic elements are needed.

We would like to emphasise the importance of a complex social structure, with regards to the evolution of musical style. If the compositional goal is to create a melody which has the same melodic style of our system at a specific time, then one can use a single composer with all parameters configured specifically to obtain that precise melodic style. But, if the compositional goal is to simulate a steady stylistic evolutionary process, that does not solely randomly explore areas of the style spaces, there is no way traditional search/evolutionary algorithms can succeed. All nuances given by non-linear interactions of many agents, whose grammars and musical contents are slightly different from one another, cannot be reproduced by a single agent. Indeed, the value of a many-composers approach lies in the musical journey, not in the single stages. The social level of the system adds an extra layer of complexity that cannot be reduced.

As experimental analysis suggests, the systems tends to converge rapidly. The plot of the standard deviation highlights that after 50 iterations all composers have only small differences in all their parameters. The fast convergence rate as well the fact that there is no difference in the way the system with different numbers of composers behaves, are probably due to the fully connected network used to model the musical environment. To provide unpredictability and instability, which are necessary elements of creativity, a scale-free network is needed.

However, since all experiments of the system converge within specific regions of the three phase spaces, it is probable that those regions are attractors for the system. In other words, not all regions of the phase spaces are equally likely to be visited. Even considering all the simplifications made to develop the system, we can extend this idea to the real world, since the system is based on sound cognitive elements. Our hypothesis is that the real-world musical style phase space is divided into a number of attractorsThese regions are privileged portions of the phase space, whose shape depends on a number of cognitive and social factors such as musical universals, exposure to music, and cognitive constraints.

The system we proposed has two main limitations. First, only three parameters of the musical grammar change over time. As a consequence, composers can hardly completely revolutionise their musical style, since they are severely constrained by many other parameters that do not evolve. However, if many parameters could be modified, then the musical results of the system would be completely unpredictable, and it would be really difficult to trace the contribution of one parameter to a specific compositional behaviour. Also, the complexity of the system would increase a lot, making the system less manageable. Second, the evaluation function is simple and its parameters are stationary. This implies that composers cannot have sophisticated aesthetic judgement on melodies they "listen to". However, a complex evaluation function for simulating the aesthetic preferences of human composers, would be extremely difficult to develop.

7 Conclusion

Automatic melodic generation is a challenging task, and a large number of approaches that are able to produce human-enjoyable melodies have been proposed. However, no system has been capable of evolving its compositional style until now. In this paper, we presented such a system. The method we introduced relies on a memetic evolutionary approach, whose backbone is a multi-agent society of composers organised into a network. Composers evolve their musical grammar and the musical contents they use to generate music, by constantly exchanging musical information. Both quantitative analysis and music experts confirm that the system is able to effectively evolve its own compositional style over time. As a consequence, the paper has also indirectly demonstrated that computers can show at least a low degree of transformational creativity.

Future work includes the improvement of the evaluation function of composers, the exploitation of a scale-free network for simulating the society, and the generation of small polyphonic pieces rather than simple monophonic lines.

References

1. Boden, M.A.: The creative mind: Myths and mechanisms. Psychology Press (2004)
2. Brown, S., Jordania, J.: Universals in the world's musics. Psychology of Music 41, 229–248 (2013)
3. Cope, D.: Computers and musical style (1991)
4. Dannenberg, R.B., Thom, B., Watson, D.: A machine learning approach to musical style recognition (1997)
5. Dawkins, R.: The selfish gene (1976)
6. De Leon, P.J.P., Inesta, J.M.: Musical style classification from symbolic data: A two-styles case study. In: Computer Music Modeling and Retrieval, pp. 167–178. Springer (2004)
7. Dennett, D.C., Mittwoch, U.: Darwin's dangerous idea: Evolution and the meanings of life. Annals of Human Genetics 60(3), 267–267 (1996)
8. Fabbri, F.: Browsing music spaces: Categories and the musical mind (1999)
9. Fernández, J.D., Vico, F.J.: AI methods in algorithmic composition: A comprehensive survey. J. Artif. Intell. Res. (JAIR) 48, 513–582 (2013)
10. Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychological bulletin 76(5), 378 (1971)
11. Gimenes, M., Miranda, E.R., Johnson, C.: A memetic approach to the evolution of rhythms in a society of software agents. In: Proceedings of the 10th Brazilian Symposium on Computer Music (SBCM). vol. 16 (2005)
12. Graham, G.: Genes: a philosophical inquiry. Psychology Press (2002)
13. Hontanilla, M., Pérez-Sancho, C., Iñesta, J.M.: Modeling musical style with language models for composer recognition. In: Pattern Recognition and Image Analysis, pp. 740–748. Springer (2013)
14. Jan, S.: Music, memory, and memes in the light of calvinian neuroscience. Music Theory Online 17(2), 3–50 (2011)
15. Jan, S.B.: The memetics of music: A neo-Darwinian view of musical structure and culture (2007)
16. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. biometrics pp. 159–174 (1977)
17. Lartillot, O., Dubnov, S., Assayag, G., Bejerano, G.: Automatic modeling of musical style. In: Proceedings of the 2001 International Computer Music Conference. pp. 447–454 (2001)
18. Meyer, L.B.: Style and music: Theory, history, and ideology. University of Chicago Press (1989)
19. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological review 63
20. Miranda, E.R.: On the evolution of music in a society of self-taught digital creatures. Digital Creativity 14(1), 29–42 (2003)
21. Pachet, F.: Rhythms as emerging structures. In: Proceedings of 2000 International Computer Music Conference, Berlin, ICMA (2000)
22. Velardo, V., Vallati, M.: Automatic melody composition and evolution: a cognitive-based approach. In: Proceedings of the Conference on Interdisciplinary Musicology (CIM) (2014)