

GenoMeMeMusic: a Memetic-based Framework for Discovering the Musical Genome

Valerio Velardo

University of Huddersfield
U1370329@hud.ac.uk

Mauro Vallati

University of Huddersfield
m.vallati@hud.ac.uk

ABSTRACT

The paper introduces G3M, a framework that aims to outline the musical genome through a memetic analysis of large musical databases. The generated knowledge provides meaningful information about the evolution of musical structures, styles and compositional techniques over time and space. Researchers interested in music and socio-cultural evolution can fruitfully use the proposed system to perform extensive inter-opus analysis of musical works as well as to understand the evolution occurring within the musical domain.

1. INTRODUCTION

Music is a highly structured phenomenon which can be easily analysed through computational techniques. Nowadays, a large amount of data and information are freely available on the Internet. That is the case of music as well. Indeed, the ready availability of musical data can be exploited by extracting relevant information directly from the structure of musical compositions, in order to discover unknown relationships between musical utterances, pieces, and composers. Furthermore, this process could unveil the inner evolutionary process of music, which is responsible for the change of musical style, taste and compositional techniques over time. Surprisingly, very few projects exploited the increasing availability of big data in music for performing extensive structural analysis of musical works.

In this paper we propose a framework that aims to automatically discover the musical genome: GenoMeMeMusic (G3M). The task is performed by identifying and finding the occurrences of musical memes [1] within large musical databases. Musical memes (musemes) are cognitively relevant chunks of musical information which can be copied from one brain to another. Indeed, G3M considers music as a cultural evolutionary process, thus it extracts fundamental components which make up music, and traces their evolution over time and space. The framework addresses the following research questions:

- How does music evolve over time and space?
- What does the musical genome consist of?

- What are the stylistic relationships between composers?
- What are the best strategies for identifying musical memes and tracing their mutations?

The knowledge inferred by the G3M framework provides useful insights on musical structure, style and evolution, to researchers interested both in music and sociocultural evolution.

The remainder of this paper is organised as follows. First we summarise relevant related works, then we provide the necessary background on memes and musical pattern discovery. Section 4 describes the high level structure of G3M. Section 5 provides a working definition of musical memes as used by the framework. Sections 6 and 7 describe the main modules of G3M as well as the provided outputs. Finally, section 8 gives the conclusions.

2. RELATED WORKS

Artistic, biological and sociological phenomena such as pieces of music, DNA or literary movements usually show extremely complex structures. One of the most exploited approaches to handle such complexity is to reduce it by splitting the phenomenon into a sequence of constituents that encode bits of information. When some of those constituents are arranged together through generative rules, an instance of the phenomenon arises, showing high level of complexity. Therefore, in order to understand and describe complex systems it is necessary to unveil the single parts of the structure and discover the generative rules that allow their combination. From a high-level point of view, this approach can be regarded as discovering the genome of a complex system.

The Human Genome Project is the main example of the process of discovering and categorising the components of a complex system [2]. In particular, the Human Genome Project had the goal of determining the sequence of chemical base pairs which constitute human DNA, as well as to finding and mapping all the genes of the human genome. The project, which was completed in 2003, found 20,500 genes and analysed more than 3.3 billion base pairs.

After the Human Genome Project, a number of projects attempted to create a map of constituents of complex phenomena functionally similar to the genetic one. The most interesting examples consider either artistic or sociological phenomena.

The Book Genome Project proposes an intelligent system that identifies and measures the salient aspects which

make up a book.¹ Different components, such as language, characters and themes, are analysed in order to organise and categorise books. Books are separated one from another and put into an abstract complex space of books named the “booksphere”. The Book Genome has three primary gene structures (i.e. language, story, characters), which contain a specific subset of measurements. The system tracks and quantifies the different measurements and put the results into an online database. The final outcome is a genome which systematically encodes and categorises different possible manifestations of a book.

A similar study covers the visual art domain. The Art Genome Project aims to categorise artists and artworks by providing a unique genome for each of them.² Particularly, every artistic genome is made up of about 400 genes which are organised in coherent categories such as medium, time period and style. The result is an abstract space that organises and structures the visual art domain coherently.

The same approach has been used also to categorise music. The Music Genome Project proposes a specific genome that uniquely describes a musical composition [3]. The musical genome is made up of 450 different genes which reflect salient characteristics of a piece of music such as tempo, key and gender of the lead vocalist. The process of categorisation is carried out by musical experts, who listen to a musical work and give a score to each of the different 450 genes. Every genome is then stored in an online database. The project has also a web application, called Pandora. Pandora is a web radio which suggests pieces of music to the listeners. The suggestions are based on listeners, musical preferences, and are made by exploiting the database of the Music Genome Project.

Although the Music Genome Project has demonstrated itself to be effective, it is possible to identify some issues. It relies on music experts to extract information from a piece of music and, therefore, to compile the musical genome. This interactive process shows two major flaws. First, there could be significant differences between experts in how they judge music and score genes. Secondly, there is a substantial problem of scalability. Indeed, the greater the number of pieces the project wants to analyse, the greater the number of music experts needed. Moreover, the Music Genome Project uses very broad categories to define the genome of a musical composition. Thus, the project focuses on high-level descriptions, ignoring the raw musical content that actually makes up a piece of music, such as rhythms, notes and melodies.

To overcome some of these problems, Hawkett proposed an automatic extraction system which identifies musical patterns and performs research based on pattern similarity on a group of different pieces [4]. The outcome is a form of musical genome that encodes the melodic materials that make up a set of string quartets. Hawkett exploits a brute-force approach, which considers every musical pattern defined as a group of notes containing from 3 to 11 tones. Also, the study attempts to demonstrate the existence of musical memes by analysing the evolution and the

properties of music patterns extracted from the string quartets. This work has a significant weakness. The algorithm of extraction ignores the cognitive relevance of the musical patterns, since it focuses on every possible pattern of 3 to 11 notes. Therefore, the system overlooks the musical relevance of the patterns identified.

3. BACKGROUND

This section introduces the concepts of meme, museme (i.e. musical meme) and the existing relevant techniques of pattern matching used in music.

3.1 Memes and Musemes

Memes are cultural traits that can be passed on from one person to another by non-genetic means such as imitation and teaching [5]. They can be habits, ideas, stories, songs or tunes [6]. Memes are selfish replicators like genes, since they are bits of information that are copied with variation and selection. They can be encoded in different ways, as pieces of information in the human brain or on DVDs, and they compete for survival evolving in a meme pool. Although memes and genes are quite similar, there are some major differences between them. Genes are made of DNA, whereas memes are not. Furthermore, there is no equivalent to a base pair for memes. Finally, genes are more stable than memes, since they experience a radically slower rate of mutation than memes.

Nonetheless, memes and genes share some basic properties, such as *copying-fidelity*, *fecundity* and *longevity* [6]. *Copying-fidelity* assures that replicators are copied accurately and remain recognisable over time. This process does not exclude variation, rather it indirectly fosters the dynamic process of selection that memes undergo within the meme pool. *Fecundity* refers to how rapidly a meme can be replicated and spread. This property is of primary importance: it guarantees a clear competitive advantage to replicators which have large number of copies. *Longevity* measures how long a meme can survive and evolve. The greater the amount of time a meme remains active, the greater the possibility of spreading. Fidelity, fecundity and longevity are complementary properties of replicators which contribute to define the success of memes.

Memes evolve over time and respond to selective pressure. The memetic-evolutionary process is analogous to the genetic-evolutionary process. Dennett identifies three elements of an algorithm that guarantee evolution: *variation*, *heredity or replication* and *differential fitness* [7]. *Variation* refers to a huge amount of different elements within a pool of replicators. *Heredity or replication* refers to the capacity of element to create copies of themselves. *Differential fitness* provides a selective process guaranteed by the interaction of the elements at a certain time with the environment. Variation, heredity or replication and differential fitness are conditions that appear both within the memetic and genetic domains.

It is worth saying that group of memes can be organised, so that they replicate and adapt together. Such complex memetic structures might be termed *memplexes* [8].

¹ <http://bookgenome.com> (last accessed 05/05/2014)

² <https://artsy.net/theartgenomeproject> (last accessed 05/05/2014)

Memes that live within a memplex benefit from the success of the memplex itself. Examples of memplexes are religions and cultures which consist of a set of coherently organised memes that spread and replicate together.

Memes can also play a fundamental role in analysing music. As suggested by Jan, it is possible to consider music from a memetic point of view [1]. This approach is compatible with applications of Darwinian theories of evolution, and provides a useful theoretical framework to understand relevant questions such as why some musical structures and procedures are more common than others at certain times.

Jan defines a musical meme or *museme* as 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.

Musemes are cognitive relevant musical structures and listeners can identify them partly through bottom-up innate cognitive processes, and partly through top-down learned listening strategies. Moreover, musemes exist at several structural hierarchical levels of a musical piece and are usually multi-parametric instances of pitch and duration. Several musemes constitute musical memplexes across many hierarchical musical structures, up to the level of the piece as a whole. Musemes manifest the basic meme properties of longevity, fecundity and copying-fidelity. Additionally, they undergo the same algorithmic evolutionary process which consists of the three steps of variation, replication and differential fitness.

As far as we know, there are few studies attempting to identify musemes in musical compositions. In 2004, Jan [9] tried to track and identify musemes in the Adagio in C Major for Glass Harmonica, KV 356 by Mozart, exploiting the Humdrum Toolkit. Even though the work opens new avenues of research, its methodology inherently lacks scalability. Indeed, the patterns had to be manually inserted into the system in order to discover the occurrences of musical memes within a single piece. Therefore, an application to large musical databases would be impractical. Rather, an intelligent system that could autonomously identify and confront musemes within a large set of musical works is needed.

3.2 Musical Pattern Discovery

Pattern discovery is a fundamental part of symbolic music processing [10], which has numerous applications such as music analysis, music information retrieval and music classification. There are several algorithms that perform pattern discovery exploiting different strategies.

Conklin [10] proposes an approach that considers inter-opus pattern discovering, i.e. the process of discovering recurring patterns within a corpus of musical pieces. The system addresses the issue of pattern ranking by focusing on *distinctive patterns*, which are defined as frequent patterns that are over-represented in the corpus, as compared to an anticorpus of random generated musical pieces. Even

though the system proposed by Conklin manages to find occurrences of the patterns across several pieces, it does not consider the evolution of the patterns over time and their structural organization.

Lartillot [11] proposes an algorithm of pattern discovery based on relevant cognitive processes. The system represents music along two dimensions: melody and rhythm. Musical patterns are modelled as a chain of states. The algorithm exploits the main feature of *associative memory*, i.e., the capacity of relating items which show similar properties. Associative memory is represented by hash tables which encode the two different musical parameters. The huge number of patterns that can potentially arise from the algorithm are reduced through a filtering technique, that follows the criteria of selection of the longest and most frequent patterns. However, the system works only at an intra-opus level, since it can only process a single piece at a time, and it is limited to monophonic music.

Conklin and Anagnostopoulou [12] propose an approach that focuses on deeper musical structures called *viewpoints*. Viewpoints model specific typologies of musical features such as melodic contour, duration and intervals. The algorithm can find deeper transformed representation of a pattern, shifting the problem of looking at similarity between two patterns from a surface level into a deeper representational level. The system does not adopt a cognitive approach and again considers only the intra-opus level.

Szeto and Wong [13] tackle the problem of identifying patterns in post-tonal music by modelling a musical work as a network. Every note of a piece is represented by a node, and the relationships between two notes by an edge. Searching for a musical pattern is equivalent to looking for a subgraph of the network. The algorithm also models the perceptual dimension by considering melodic groups of notes as single coherent and continuous line called a *stream*. The system is limited to post-tonal music, and adopts a not very sophisticated strategy to detect similarities between patterns.

Meudic [14] considers similarity in polyphonic contexts. The proposed algorithm uses three musical factors to decide whether or not two patterns are similar. These are pitch, melodic contour and rhythm. The system initially performs a measurement of similarity along these three aspects, and then considers a global similarity measure, which derives from their linear combination. The similarity measure for pitches and melodic contours considers only the musical events falling on the downbeats. Furthermore, the system focuses only on intra-opus analysis.

An interesting approach to pattern discovery is adopted by Lartillot [15], which focuses on analogy and induction. The algorithm of pattern detection copes with approximation rather than repetition, and exploits a powerful system of induction. The system is capable of inducting new patterns based on analogies with older patterns. The algorithm adopts an interesting cognitive approach. It considers the experience of music as a temporal progression, and infers the global musical structure of a piece through induction of hypotheses from local viewpoints. Additionally, the algorithm is capable of inferring patterns of patterns and organ-

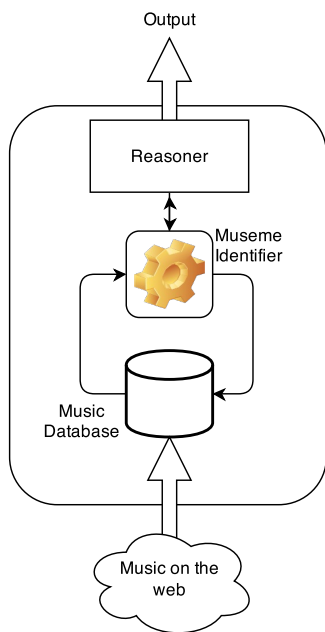


Figure 1. The structure of the GenoMeMeMusic framework.

ising a musical piece in a semantic network, with information distributed throughout the network. This system does not discover similar patterns across different pieces and, moreover, it sometimes does not recognise relevant musical patterns within a piece, due to the inductive cognitive process itself.

Although there are many systems which perform musical pattern discovery, none of them deals with the memetic structure of music. Likewise, none of them analyses the relationships among different musical patterns in order to infer the evolutionary process undergone by music. Indeed, until now the inference of the musical evolutionary process has been carried on a qualitative base by musicologists and music theorists who directly analysed scores.

4. FRAMEWORK

Figure 1 shows the structure of the proposed G3M framework. It gathers music files, in the Music XML standard, from the Internet or other existing sources. Music XML has been selected due to its high expressivity (it can include much more information than other standards, e.g., MIDI) and to the large number of available sources [16]. Music XML is translated in the internal encoding format, described in section 6.1. This encoding has been designed for simplifying the operation that will be performed by the Museme Identifier, namely segmenting music and looking for similarities. The knowledge extracted and organised by the Museme Identifier is then exploited by the Reasoner, which analyses the obtained structures and information and provides the output, i.e., the musical genome. The output is provided under two main forms: networks and meme characteristics. The first focuses on representing information by using relationships between composers and music

pieces, based on musemes shared. The second one is focused on confirming and evaluating the main known properties of musical memes.

The rest of the paper will describe the modules of the G3M framework, in particular from the functional perspective.

5. IN SEARCH OF MUSEMES

The G3M framework substantially differs from any related project on musical pattern discovering in music, since it focuses on musemes and musical evolution. The G3M framework uses a cognitive approach in discovering patterns in musical compositions. Indeed, it considers musical utterances which are maximally relevant for the human brain. These structures are short musical phrases usually from 3 to 5 seconds long, which have fewer than 25 musical events [17]. These reflect the cognitive constraints of human memory.

Indeed, people perceive music in coherent chunks which are stored and processed in Short Term Memory. Some of these chunks, through rehearsal, are then passed to Long Term Memory. This second type of encoding allows the listener to experience motivic connections and relate large hierarchical structures of music while listening to a piece. However, the real-time processing of music is carried by Short Term Memory. This phenomenon implies that the actual musical currency used by the brain is the musical phrase 3 to 5 seconds long as defined by Snyder [17]. For this reason, we propose that musemes, which are bits of musical information that spread from one brain to another, should correspond to this musical structure, which in turn is the most cognitively relevant. It is not surprising that classical composers often adopted these musical structures, instinctively aligning to natural cognitive constraints. Furthermore, musical phrases usually have a character of closure which concludes a small as well as self-contained musical discourse. This can be explained by considering that musical phrases, and thus musemes, are the bits of information directly processed and stored by the brain as a unitary structure.

6. MUSEME IDENTIFIER

This section identifies the strategies adopted by the Museme Identifier in order to encode music, find musemes, manage polyphony and assess similarity between musemes.

6.1 Musical Encoding

Symbolic musical representation is a fundamental aspect of music information retrieval. A good musical representation facilitates the manipulation of musical information, increasing the overall computational efficiency of algorithms which deal with musical segmentation and similarity.

The G3M system uses a basic representation of music which focuses on pitch and duration. This representation is a simplified version of the MIDI encoding. Secondary parameters such as timbre, loudness and articulation are ignored, since they are not exploited by the algorithm and

major challenges of G3M is to decide whether or not two musemes can be regarded as the 'same' pattern. To determine this, the algorithm uses an approach based on cognition, which considers several parameters to judge the similarity of two musemes.

In particular, the algorithm considers the number of tones and the distance in pitch, rhythm and melodic contour between two musical phrases as different parameters to evaluate. Moreover, the algorithm introduces a metric which considers the complexity of the museme itself. The rationale behind this is that the more complex a museme is, the more difficult is to relate two patterns together when they differ along some parameters. All of these metrics are arranged altogether in a linear combination. The resulting score value is used for comparing musemes.

The process of recognising the similarity between two musemes is essential for understanding and explaining the memetic process of music. Indeed, this algorithm, which is part of the Museme Identifier module, is the most critical element of the whole framework.

7. OUTPUT

This section analyses the outputs of the memetic analysis performed by the Reasoner of the G3M framework. These outputs correspond to the genome of music. The section considers both the main properties of memes (i.e., longevity, fecundity and copying-fidelity) as well as the structural organisation of musemes within music pieces considered at the inter-opus level. Furthermore, the Reasoner will exploit time and geographical information encoded within the music pieces, in order to highlight how museme parameters evolve over time and space.

7.1 Meme Properties

In order to prove that music can be regarded as a memetic phenomenon, it is necessary to demonstrate that the extracted patterns show the salient properties of memes.

7.1.1 Longevity

Longevity refers to how long a meme can survive, and can be observed in pieces composed at different times. To assure memetic evolution, memes must survive a sufficient amount of time. Therefore, it is of main importance understanding whether or not the musemes identified by G3M are persistent enough to establish an evolutionary process. The Reasoner measures longevity by calculating the average lifetime, as well as other relevant lifetime-related information, of the musemes in the dataset. However, it is likely that the average lifetime of the musemes could be a meaningless measure, since a power-law distribution is expected. Indeed, we think that just a few musemes are extremely long-lived, whereas the majority of them usually present a shorter lifetime.

7.1.2 Fecundity

Fecundity refers to the rate of replication of a meme. The greater the rate of replication, the greater the possibility of

that meme to spread throughout the meme pool. To measure this parameter, the Reasoner checks the number of occurrences of each identified museme. The measurement considers only one occurrence of a museme per musical piece, whether or not the museme appears more than once within the same piece. The rationale behind this choice is to avoid internal redundancy. The Reasoner extracts the distribution of the number of occurrences of the musemes over the considered dataset. Again, we expect a power-law distribution with a small number of musemes overrepresented within the database.

7.1.3 Copying-fidelity

Copying-fidelity refers to the capacity of producing faithful copies of a meme. The more accurate the copy, the more will remain of the initial pattern after several rounds of replication. The Reasoner measures infidelity by calculating the ratio between the number of mutated occurrences of a museme and the total occurrences of the same museme within the database. Copying-fidelity can be easily derived by subtracting the value of infidelity from one. Then, the system calculates the average fidelity and the standard deviation, and finds the statistical distribution. As for previously discussed properties, a power-law distribution is expected.

7.2 Networks

To visualize the database of musemes as well as to gain analytical insights, the system organises the data in two different complex networks, which provide relevant musical information and which should prove the memetic evolutionary process undergone by music. These networks properly correspond to the musical genome that the research aims to track.

7.2.1 Museme

In the Museme Network musemes are the nodes. Nodes are connected by edges, which correspond to a music piece which two musemes both appear in. The edge is weighted. The greater the number of the pieces two musemes appear simultaneously in, the greater the weight of the link they share. The network organises the musical material depending on the relationships musemes have within pieces of music.

The Museme Network represents a kind of genome of music, since it corresponds to the meme pool of basic musical structures encoded in the human brain. This network can be easily analysed for gaining insights on the closeness of some musemes. We expect a free-scale network, with a small number of components which are hyperconnected, and a huge number of musemes which are connected to few others. Additionally, the network can be generated and studied by considering different time periods, in order to understand how the components, their links and the general parameters which describe the network evolve over time.

7.2.2 Composer

The Composer Network considers composers as nodes and musemes as edges. In particular, a link between two composers is created if they used the same museme in one of their works. The edges are weighted, since the greater the number of common musemes two composers use, the greater the weight of the link that unites them. As a consequence, the Composer Network shifts the focus of the research from the musical materials themselves to the artists who used them. This network highlights the relationships and similarities among composers. It is possible to identify clusters of composers which are aggregated together, since they used similar musical structures. Furthermore, a measure of similarity between composers is also possible by considering the number of the same musemes two composers share. As a consequence, the Composer Network represents a kind of genome of composers based on the musical materials they adopt in their works.

We expect a network with few composers overconnected, who can be regarded as the pillars responsible for the evolutionary process of music. The rationale behind this distribution is that we think of music as a complex memetic system, socially structured and based on imitation and passage of information. All of these aspects inherently imply an aristocratic (i.e. power-law) distribution, where a few hubs act as gigantic connectors.

8. CONCLUSIONS

Despite the increasing availability of musical pieces due to the Internet, very few systems carry out extensive structural analysis of musical works for highlighting their relationships and providing insights into the cultural evolutionary process of music.

In this paper we proposed GenoMeMeMusic, a framework that discovers the musical genome and its evolution, by exploiting the concept of museme. The G3M framework includes two main modules, one of which is devoted to identifying musemes in a large database of compositions, and the other which exploits the knowledge encoded by the Museme Identifier for high-level reasoning. The output of G3M will be in the form of networks, either of composers or musemes, and of meme properties. Future work includes the implementation of the proposed framework and a preliminary analysis on the Essen folksong collection.

9. REFERENCES

- [1] S. Jan, *The memetics of music*. Ashgate, 2007.
- [2] C. Delisi, "The Human Genome Project," *American Scientist*, vol. 76, pp. 488–493, 1988.
- [3] M. Castelluccio, "The Music Genome Project," *Strategic Finance*, vol. 88, pp. 57–58, 2006.
- [4] A. Hawzett, "An empirical investigation of the concept of memes in music using mass data analysis of string quartets," Ph.D. dissertation, University of Huddersfield, 2013.
- [5] D. C. Dennett, *Darwin's dangerous idea: evolution and the meanings of life*. Penguin Books, 1993.
- [6] R. Dawkins, *The Selfish gene*. Oxford University Press, 1976.
- [7] D. C. Dennett, *Consciousness Explained*. Penguin Books, 1995.
- [8] S. Blackmore, *The meme machine*. Oxford University Press, 2000.
- [9] J. Steven, "Meme hunting with the humdrum toolkit: Principles, problems, and prospects," *Computer Music Journal*, vol. 28, no. 4, pp. 68–84, 2004.
- [10] D. Conklin, "Discovery of distinctive patterns in music," *Intelligent Data Analysis*, vol. 14, no. 5, pp. 547–554, 2010.
- [11] O. Lartillot, "Efficient extraction of closed motivic patterns in multi-dimensional symbolic representations of music," in *Proceedings of the 2005 IEEE/WIC/ACM International Conference on web intelligence*, 2005, pp. 229–235.
- [12] D. Conklin and C. Anagnostopoulou, "Representation and discovery of multiple viewpoint patterns," in *Proceedings of the International Computer Music Conference*, 2001, pp. 479–485.
- [13] W. M. Szeto and M. H. Wong, "A graph-theoretical approach for pattern matching in post-tonal music analysis," *Journal of New Music Research*, vol. 35, no. 4, pp. 307–321, 2006.
- [14] B. Meudic, "Musical similarity in a polyphonic context: a model outside time," in *Proceedings of the XIV Colloquium on Musical Informatics (XIV CIM 2003)*, 2003.
- [15] O. Lartillot, "Generalized musical pattern discovery by analogy from local viewpoints," in *Discovery Science*, ser. Lecture Notes in Computer Science, 2002, vol. 2534, pp. 382–389.
- [16] M. Good, "Musicxml for notation and analysis," *The Virtual Score: Representation, Retrieval, Restoration*, pp. 113–124, 2001.
- [17] R. Snyder, *Music and Memory*. MIT press, 2001.
- [18] D. Temperley, *The Cognition of Basic Musical Structures*. MIT Press, 2001.