



HAL
open science

Generalized Musical Pattern Discovery by Analogy from Local Viewpoints.

Olivier Lartillot

► **To cite this version:**

Olivier Lartillot. Generalized Musical Pattern Discovery by Analogy from Local Viewpoints.. Springer-Verlag; Steffen Lange, Ken Satoh, Carl H. Smith. Discovery Science, 5th International Conference, DS 2002, Lübeck, Germany, November 24-26, 2002, Proceedings, pp.382-389, 2002. hal-01161321

HAL Id: hal-01161321

<https://hal.science/hal-01161321>

Submitted on 8 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generalized Musical Pattern Discovery by Analogy from Local Viewpoints

Olivier Lartillot

Ircam - Centre Pompidou, Place Igor-Stravinsky, 75004 Paris, France
Olivier.Lartillot@ircam.fr

Copyright: Springer-Verlag.

Abstract. Musical knowledge discovery, an important issue of digital network processing, is also a crucial question for music. Indeed, music may be considered as a kind of network. A new approach for Musical Pattern Discovery is proposed, which tries to consider musical discourse in a general polyphonic framework. We suggest a new vision of automated pattern analysis that generalizes the multiple viewpoint approach. Sharing the idea that pattern emerges from repetition, analogy-based modeling of music understanding adds the idea of a permanent induction of global hypotheses from local perception. Through a chronological scanning of the score, analogies are inferred between local relationships – namely, notes and intervals – and global structures – namely, patterns – whose paradigms are stored inside an abstract pattern trie. Basic mechanisms for inference of new patterns are described. Such an elastic vision of music enables a generalized understanding of its plastic expression.

1 Music Analysis and Knowledge Discovery

1.1 Music Is a Network

In the realm of digital networks, music takes a pivotal place. Indeed, current and future networks such as Internet will feature a great amount of musical content that will have to be structured pertinently. That is why much research is carried out now in a new domain called Music Information Retrieval (MIR) focused on the automatic processing of music databases.

But music itself, because of its combinational richness, may be considered as a dense and complex network : notes are linked together forming higher structures which themselves communicate one with each other. Musical styles may also be seen as super-networks unifying specific structures inside different pieces as occurrences of an element of common language. A single musical piece may contain a huge quantity of information, implicitly understandable by listening but that was impossible to analyze and represent explicitly. That is why even for traditional music analysis, “a new generation of computational techniques and tools is required to support the extraction and the discovery of useful knowledge”¹.

¹ Call for Papers of this conference.

1.2 Musical Pattern Discovery

For a machine to proceed to musical analysis by itself, without any a priori knowledge, it would first have to be able to automatically discover simple musical patterns. Lots of researches have recently been led in this domain, offering significant results, but without resolving the challenge in a general context yet. In our approach, and in those presented in this paragraph, music expression is limited to scores. For non-written music, a preliminary – and rather intricate – automatic music transcription would be necessary.

Gestalt principles – segmenting musical discourse when some perceptual distance gets maximal values – may be used for broad segmentation of music into large streams and temporal blocks [2]. However, detailed segmentation of music is better ruled by repetition than by Gestalt [3]. In this poster, we will mainly focus on the repetition criterion.

As usual musical patterns are not exactly repeated, an algorithm of pattern detection should cope with approximation. One major axe followed by current MIR research proposes to overcome this difficulty by considering multiple music representations (called *viewpoints*) in parallel [8] [4] [3]. According to this theory, an approximate repetition may be viewed as an exact repetition in one relevant viewpoint. However, this approach cannot explain how a pattern may be considered as similar to a slightly stretched version of it, as in Fig. 1. For this example, repetition will only be found along the contour viewpoint, where only the slope between successive notes is considered (for this pattern: up then down). But this viewpoint is so loose that non-pertinent repetition may also be found. Therefore, information selection is not sufficient for approximate repetition discovery. Repetitions should be actively detected, when some similarity distance exceeds a certain amount of relevance.

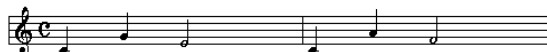


Fig. 1. Two similar patterns

In most approaches, including multiple viewpoints, pattern is considered as sequence of contiguous elements. In such a restricted vision of music, long motives hidden behind dense accompaniments, for instance, would not be retrieved. We prefer considering a broader approach, where pattern are considered as sets of notes. The difficulty of such an approach, evidently, is computational complexity. Considerations of simple segmentations and edit distance are no more accurate.

In many musical pattern detection algorithms, chosen candidate patterns are those that feature a highest score. This score takes into account several independent characteristics – or “heuristics” – such as pattern length (prefer longer patterns), or number of occurrences (prefer most frequently occurring patterns) [3]. This paradigm of pattern selection stems from the idea that only

a limited number of relevant characteristics of musical works have to be chosen. The remaining will be discarded following the constraint of result compactness. This strategy, though alleviating the information retrieval process, arbitrarily infringes a complex and subtle understanding of music. That is why we would prefer another approach, where selection does not come from competition between candidate patterns, but rather from criterion individually applied to each candidate pattern. In a word, a pattern will not be chosen because it is the most perceived pattern, but because it can be perceived. Therefore, to pattern selection, we would prefer pattern *detection*.

2 A Cognitive Approach

2.1 Analogy

We suggest another method, highly developed in contemporary cognitive sciences and probably of high interest here, based on analogy. Indeed, “analogy-making lies at the heart of pattern perception and extrapolation” [5]. Analogy means that the similarity of two entities is hypothesized as soon as one or several resemblances of particular relevant aspects of them are detected. Human understanding heavily relies on such a mechanism since it progressively constructs a representation of phenomenon. The cognitive system has to take some risks, to *induce*, that is, to infer knowledge that is not directly present in the phenomenon, and, then, to check the good fulfillment of this hypothesis. An analogy-based vision of understanding is even more accurate in a musical context. Indeed, music cannot be apprehended as a single object: at each step of music hearing, only a local aspect is presented. Thus the analogy hypothesis of music understanding means that the global music structure is inferred through induction of hypotheses from partial – in particular timely local – point of view.

The temporal characteristic of music implies another important aspect: a chronological order. A temporal oriented approach of musical pattern discovery would imply that a pattern will be more easily detected if it is presented very clearly first, and then hidden in a complex background than the reverse. Moreover, a distortion of pattern impairs more its detection when it occurs in its beginning than in its end.

2.2 Induction

Here the pivotal framework of induction modeling proposed by Holland et al. [6] is of high relevance: induction is expressed along a generalized paradigm, not founded on formal logic, but rather on a semantic network of concepts. New concepts arise from the interaction, competition and collaboration, of old ones. For each concept is associated not a single probabilistic value, but different activation quantities, each relying on different dimensions of relevance, such as past activity, entrenchment of concept inside the network, specificity with current context, etc.

This broad framework also includes an understanding of analogy itself. “Often the interpretation of an input requires the simultaneous activation and integration of multiple schemas”, that is, of different possible analogs. We propose to adapt such a vision of induction of analogies in musical context.

3 The *kanthume* Framework

Our project aims at building a musical pattern detection system based on a cognitive modeling of induction of analogies and of temporal perception of music. In a first approximation, we limit our scope to pitches p , durations d and time onsets t . The score is scanned in a temporal order and, for a same time onset, in a growing order of pitch.

3.1 Local Similarities

Local Viewpoint. As we told previously in paragraph 2.1, an analogy is the inference of an identity relation between two entities knowing some *partial* identity relation between them. Our discussion about the phenomenon of time in music leads to the idea that our experience of music is a temporal progression of partial points of view that consists of the *timely local* perception of it. Thus the analogy hypothesis of music understanding means that the global music structure is inferred through induction of hypotheses from local viewpoint.

First of all, it may be remarked that relations between notes may play an important role in pattern perception. Indeed, transposition of pattern – that is, translation of the whole pattern toward higher or lower pitch – does not change the perception of its inner shaping. Moreover, relations between more than two notes may be decomposed into relations between each possible couple of notes, namely, *intervals* (n_1, n_2) , whose parameters are pitch interval $(p_2 - p_1)$, onset interval $(t_2 - t_1)$ and duration ratio (d_2/d_1) .

Our cognitive approach gets benefit from another core characteristic of temporal perception of music. It seems that the perception of the relations – that we called interval – between each successive note of a pattern is made only possible by the fact that, when we hear a new note, the last one that have just been heard remains active in our mind, and buffered in a so called short term memory. The limited extension of short term memory should be defined both in terms of time (no more than 15 sec., for instance) and size (no more than 7 elements, for instance).

When chronologically scanning the score, for each successive instant t_0 , *current local musical viewpoint* will then consist of:

- *current notes* n with their intrinsic parameters (p, d, t_0) ,
- *current intervals*, that is, relations between current notes, and between current notes and notes of the short term past.

Associative Memory. Each parameter of each element of local viewpoints – pitch, duration and onset of current notes and intervals – may be a characteristic property that, if retrieved later, will recall the considered element. For example, if two distant notes have same pitch, the later note *may* – if some constraints are fulfilled – induce a recall of the former one. Such a remembering mechanism will be possible only if there exist associative memory that links all elements sharing a same characteristic.

In our framework, this may be formalized by considering parallel partitions of the set of notes of the score for each possible parameter (pitch, onset, duration). Here memorizing a local viewpoint would simply mean adding current notes and interval in the equivalence class corresponding to each current value of pitch, duration, onset, etc.

3.2 Abstract Pattern Trie

Pattern has been defined as a set of notes that is repeated (exactly or varied) in the score. All these repetitions, including the original first pattern, are considered as occurrences – or *actual patterns* – of a single *abstract pattern*.

By existence of relation order in each dimension of the score space (namely pitch and time onset), notes inside a set of notes may be ordered through an induced relation order, that would consist of a main temporal relation order – which possesses actual cognitive significance –, composed with a secondary (arbitrary) pitch order – for formal reason. In this way, set of notes may be displayed as strings. As we showed that relations between notes could be decomposed into relations between couples, or intervals, then sets of notes may be displayed as string of intervals between each successive note. This is called the *minimal interval representation*. In the following, intervals will be represented in semitones.

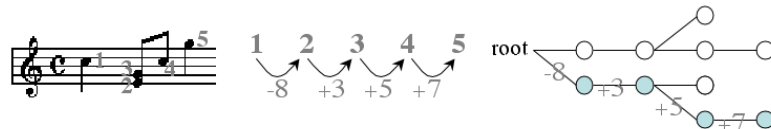


Fig. 2. Polyphonic pattern, its corresponding minimal interval representation, and its corresponding branch (with grey nodes) in one possible APT

Collection of patterns – whether musical or not – may be represented as a trie, where each node is a character (a note here) and where, for each branch, the order set of notes from the root to the leaf, is a pattern. Thanks to this kind of representation, two patterns that have a same prefix will share a same part of branch. This has cognitive ground since patterns cannot be discriminated when only their common prefix is heard. That is why the collection of abstract patterns

– in minimal interval representation – related to a score will be represented inside such a trie called *abstract pattern trie* (APT).

For each node of the APT is associated a value, its *presence*, which indicates the number of actual observed occurrences of the pattern represented by the branch from the root to this node.

Each local viewpoint note, in the score, that terminates one *or several* actual patterns is linked to their corresponding abstract patterns in the APT. And reversely, to each node of APT is associated its corresponding actual patterns in the score.

3.3 Pattern Discovery

The pattern discovery mechanism may be described following three progressive steps (see [7] for details).

Pattern Initiating If, following the associative mechanism described in paragraph 3.1, one of recalled intervals is *sufficiently similar* (in the sense explained in paragraph 3.4) to current interval, a new abstract interval, representing a paradigm of these two similar occurrences, is inferred. As this interval is considered as the beginning of a new pattern, a new node is added to the root of the APT.

Pattern Extending For each note m in the short term memory, its associated abstract patterns are recalled (or we may say still active in memory), and the analogous pattern (that is, the other instances of these abstract patterns) too. Now current interval, between considered short term note m and current note n , is compared to the possible continuations of these analogous patterns. When one continuation is found as sufficiently similar to current interval (m, n), the abstract pattern is extended by this new interval.

Pattern Confirming Other occurrences of an already induced pattern will simply be detected through a progressive scanning of the corresponding branch of the APT: each successive interval of the occurrence is automatically associated to its corresponding node in the APT.

3.4 Interval Comparison

Similitude of current and abstract or previous intervals is expressed along several dimensions: pitch interval, onset interval, etc. In a first approach, a single interval distance is considered, consisting of the sum of pitch distance and onset distance. For intervals to be considered as similar, their distances have to stay below a fixed threshold. Moreover, a long pattern accepts more freely varying continuations because the context is more firmly grounded. That is, the value of this threshold is increased proportionally to the length of abstract pattern. Threshold is also increased if the abstract pattern is linked to a great number of occurrences.

3.5 Pattern of Patterns

Now, patterns may be considered themselves as elementary objects, and patterns of patterns may be discovered in a recursive way. The trouble is, as the definition of pattern used the concept of interval between notes, patterns of patterns needs the rather esoteric concept of interval between patterns. A first approach considers pattern of abstract pattern type. A broader approach should compare all internal parameters within occurrences themselves.

If we consider the example of figure 3, where each box contains a pattern, we can see that a pattern may be either a motive or a chord. Moreover, pattern of patterns are defined here with abstract types (pattern of chord, for instance) and in the same time with internal parameters (same interval between successive chords). Finally, pattern of patterns of patterns is also possible here because of selection of pertinent knowledge (here: repetition of Eb-D).



Fig. 3. Beethoven's 5th Symphony, beginning

3.6 Semantic Network

In this way, on the score is woven a network linking all the notes with abstract concepts. This can be considered as a kind of semantic network, where knowledge is distributed throughout the network, and where implicit knowledge is deduced from the network configuration itself. This implicit knowledge has to be controlled by the system, or redundant and non-pertinent information would be added and "mud" would overwhelm the network. It is necessary therefore to design mechanisms in order to avoid inference of new concepts already subsumed by old ones. That is: Do not induce abstract patterns that already exist. At every step of the pattern discovery algorithm, presented in paragraph 3.3, such mechanism is added.

4 Results and Prospective

The algorithm presented in previous paragraph is being implemented as a library called *kanthume*² of a music representation processing software called *Open Music* [1]. First, a MIDI file of the musical work to be analyzed is loaded. During

² See *kanthume* webpage: <http://www.ircam.fr/equipes/repmus/lartillot>

the analysis, all inferences induced by the machine are displayed. First results seem very encouraging and tend to prove the pertinence of the method. Most motives are detected, even when they are transposed or temporally stretched.

Repetitions of whole patterns are detected here only when each successive interval is first detected. However, sometimes, the first interval of a pattern does not suffice for inducing the presence of this pattern. It would then be preferable to allow the detection of a pattern even after the scanning of several notes of its beginning. This may be envisaged by considering progressive activation of old recalled intervals. Once these activations exceed a certain threshold, inference is made. For this broader approach, the activation of every past note should be updated at every moment in a parallel way. Here a network approach, closer to actual cognitive modeling, would be necessary.

Added to the problematic of conception of cognitive modeling is the question of interface and ergonomics. The result of the analysis has to be displayed graphically in a network of relations, above the score itself. Because of its complexity - that cannot be graphically represented and in fact not catchy for human - this network should not be entirely displayed, but only a part of it. The user should be able to *navigate* inside this network, by choosing temporal objects and hierarchical level of the network.

To current first inquiry of Musical Pattern Discovery will then arise the question of musical form and, why not, of Music Theory Discovery.

5 Acknowledgments

This project is carried out in the context of my PhD, directed by Emmanuel Saint-James (LIP6, Paris VI) and Gérard Assayag (Musical Representations, Ircam).

References

1. Assayag, G., Rueda, C., Laurson, M., Agon, C., Delerue, O.: Computer Assisted Composition at Ircam: From PatchWork to OpenMusic. *Computer Music Journal*, 23(3), 59-72.
2. Bregman, A. S.: *Auditory Scene Analysis*. Cambridge, MA: MIT Press (1990)
3. Cambouropoulos, E.: *Towards a General Computational Theory of Musical Structure*. Ph.D. The University of Edinburgh (1998)
4. Conklin D., Anagnostopoulou, C.: Representation and Discovery of Multiple Viewpoint Patterns. *International Computer Music Conference* (2001)
5. Hofstadter, D. *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. Basic Books (1995)
6. Holland, J. H., Holyoak K. J., Nisbeth, R. E., Thagard, P. R.: *Induction : Processes of Inference, Learning, and Discovery*. The MIT Press (1989)
7. Lartillot, O.: Integrating Pattern Matching into an Analogy-Oriented Pattern Discovery Framework. *International Symposium on Music Information Retrieval* (2002)
8. Smail, A., Wiggins, G. A., Harris, M.: Hierarchical Music Representation for Composition and Analysis. *Computing and the Humanities Journal* (1993)