Olivier Lartillot
Finnish Centre of Excellence in Interdisciplinary Music Research
University of Jyväskylä, Finland
Dagstuhl Seminar "*Knowledge representation for intelligent music processing*"

MOTIVIC PATTERN MINING

This paper presents a concise overview of a research project dedicated to Motivic Pattern Mining, i.e., the automatic discovery of motives within pieces of music through a search for repetitions in score representations. Such tool would enable detailed analyses of databases of scores, extracting the thematic content of pieces of music, and clustering them according to motivic style description inferred from the analysis.

Extensive vs. Intensive Pattern Mining

This objective, for its apparent simplicity, leads to computational problems of particularly high level of difficulty. No solution to this problem has been proposed yet that would prove capable of performing a motivic analysis showing a significant degree of corroboration with musicologists' and musicians' expectations. This is due to the fact that the general problem of pattern mining, as currently studied in computer science in general, has been so far very partially answered. It turns out, in fact, that research in this area are mainly dedicated to the discovery of general trends in the corpus of data, i.e., most recurrent patterns, or at least those whose number of occurrences exceeds a certain threshold. Opposite to this paradigm, which we propose to call *extensive pattern mining*, we suggest another paradigm, called *intensive pattern mining*, rather aimed at exhaustively searching for as many motives as possible, as long as those repetitions fulfil specified minimal requirements suggesting their likely detection. The intensive pattern mining paradigm modifies significantly the computational requirements, demanding therefore an adaptation of the algorithmic framework (Lartillot, in preparation).

Multidimensional Motivic Patterns

One particularity of music is related to the multidimensionality of the underlying parametric space. One common way of tackling this core problem is based on an analysis along all possible dimensions and combination of dimensions (cf. review in Lartillot, 2007). However this does not enable to detect what we propose to call *heterogeneous description of patterns*, where musical dimensions are possibly combined differently in the course of the progressive pattern construction (cf. Figure 1). This allows the taking into account of a larger set of motivic patterns that seems to offer musical and perceptive relevance (Lartillot & Toiviainen, 2007).

Lossless Compression of the Results

The major computational difficulty in pattern mining – and in intensive pattern mining in particular – is due to the fact that a huge number of possible repetitions can be mathematically found, of which very few would be validated by listeners and musicologists. Computer science research have developed methods for filtering a large part of this complexity without any loss of information, through the concept of *closed pattern* (cf. Figure 2). The adaptation of the closed pattern paradigm to the multidimensionality of music required the development of a dedicated methodology, offering in return an interesting taxonomic description of motivic variations (cf. Figure 3) (Lartillot, in press).

Dagstuhl Seminar Proceedings 09051
Knowledge representation for intelligent music processing
http://drops.dagstuhl.de/opus/volltexte/2009/1968

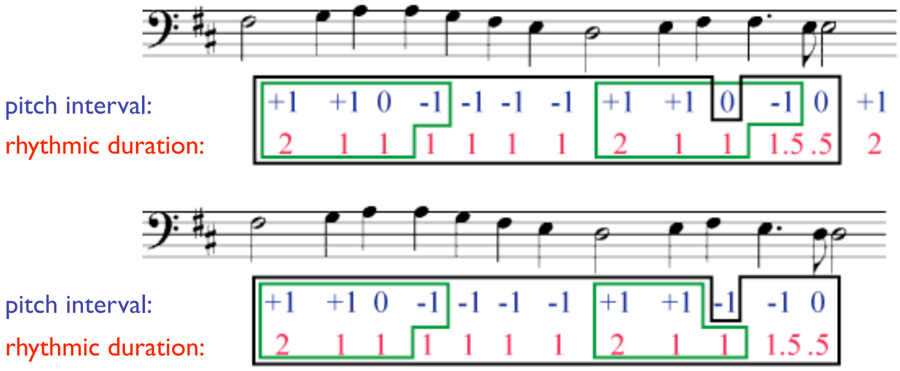

Figure 1. Beginning of Beethoven's Ninth Symphony Ode to Joy. Each whole line is a repetition of a melodico-rhythmic pattern. The only difference between the two lines is the transformation of the unison F#-F# into a descending major second F#-E, and the consequent transposition of the ending of the phrase. As a result, the representation of the motivic pattern (squared in black) includes the series of pitch intervals and rhythmic durations, except one melodic description at the position of melodic transformation.

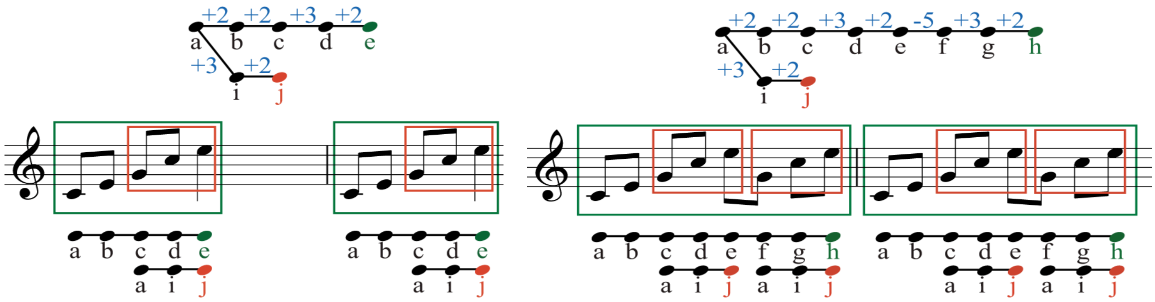

Figure 2. Left: Pattern aij (orange square) is non-closed: it is included into a larger pattern abcde (green square), and has exactly the same number of occurrences than the larger pattern. Such a non-closed pattern should be removed from the final description, as it does not offer any further information. Right: Pattern aij is now closed: it has a higher number of occurrences than the larger pattern abcdefgh. Such a closed pattern should be preserved in the final description, because it can be considered as a pattern of its own.

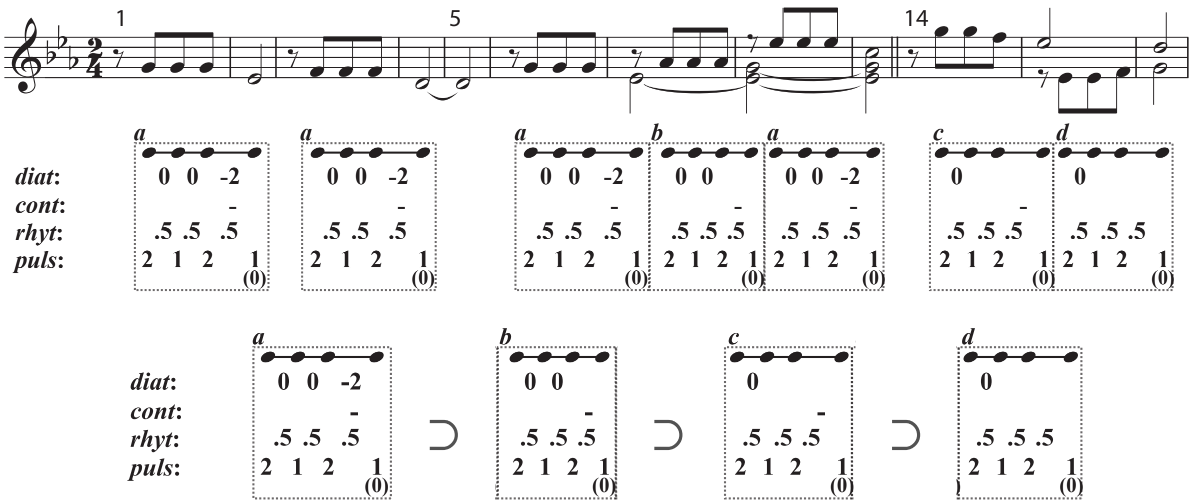

Figure 3. Beginning of Beethoven's Fifth Symphony, reduction. The 4-note motif is progressively transformed through an incremental dilution of the specificity of the original repetition (a), through a removal of the descending diatonic interval (diat = -2) (b), the second pitch repetition (diat = 0) (c), and finally the final descending contour (cont = '-') (d). The rhythmic sequence (rhyt) and the underlying pulsation (puls) remain invariant. The four patterns a, b, c and d are arranged into a decreasing order of specificy. Cf. (Lartillot, 2007) for a more detailed analysis.

Cyclic Patterns

Another combinatorial reduction appears when a given motive is repeated many times successively, such as in Figure 4. We have developed a simple methodology based on cyclic pattern, that enables to avoid such redundant proliferation of patterns (cf. Figure 5). For more details about the model, cf. (Lartillot & Toiviainen, 2007; Lartillot, in project, in preparation).

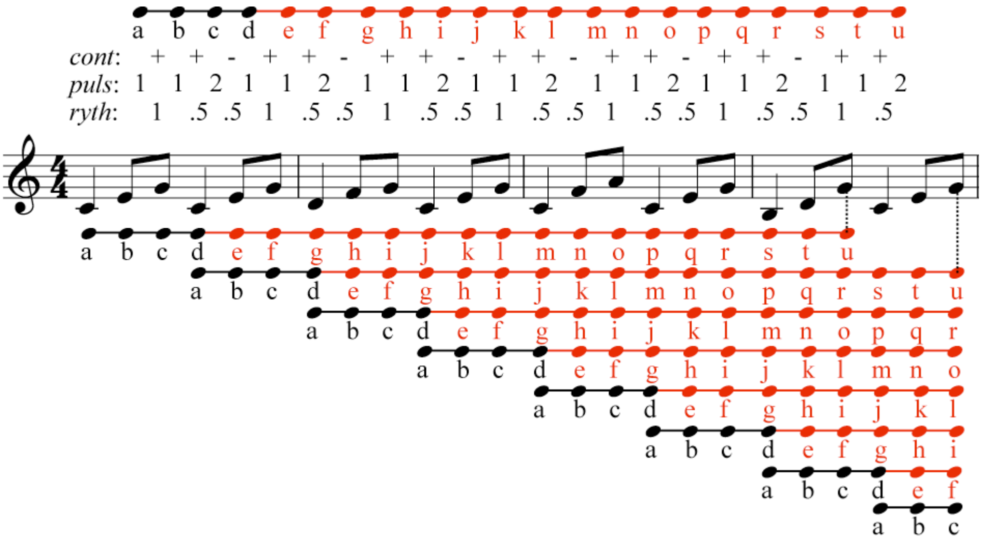

Figure 4. The successive repetition of pattern (abcd) would logically induce a proliferation of further extensions (e, f, g, etc.), leading to a combinatorial explosion harming both the computation and the results simplicity.

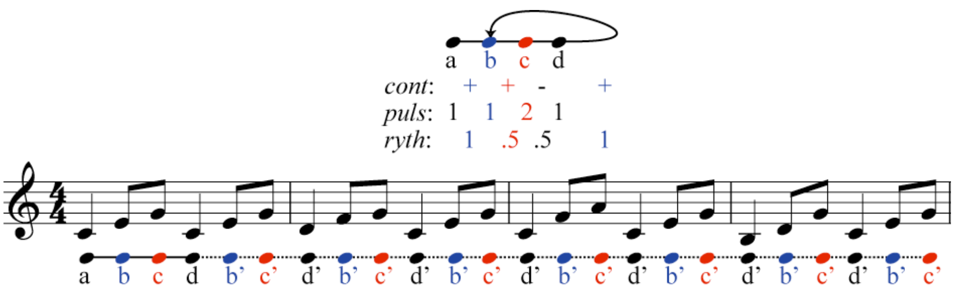

Figure 5. Through an additional transition from the end of the pattern (d) to the beginning (b), a cyclic pattern is obtained. In this way, the successive repetitions of pattern abcd can be understood as mere traversals of a same cycle.

Example of Analyses

This model has been applied to the analysis, among others, of a forteenth-century German Geisslerlied, "*Maria muoter reinû maît*", the beginning of Bach *Invention in D minor* (Lartillot & Toiviainen, 2007), of Brahms *String Quartet* Op.51 No1 (Lartillot, 2008) and to the analysis of Debussy's *Syrinx* (Lartillot, in press, in project).

Besides, the model has been integrated into a more general framework articulating the pattern repetition framework with a study of local discontinuities along the temporal flow of music. This enables to predict segmentation strategies by listeners. A tentative validation of the model has been carried out, applied to the analysis of a transcription of a traditional Tunisian *Istikhbâr* improvisation played at the Nay flute (Lartillot & Ayari, 2008).

Ongoing Research

Current research includes the taking into consideration of ornamentation and polyphony, and the generalisation of the study to the analysis of audio recordings.

The current version of the algorithm is written in Common Lisp and integrated into the *OpenMusic* environment. We plan to publish within one year a core version of the pattern mining algorithm, that could be used for the analysis of sequences of symbols, as a *Matlab* toolbox. A more detailed module would be further released and integrated to *MIDItoolbox* and *MIRtoolbox* in *Matlab*.

References

Olivier Lartillot. (in preparation). *Intensive Pattern Mining in Musicology: Toward Computational Motivic Analysis*, IGI Global.

Olivier Lartillot. (in press). "Taxonomic categorisation of motivic patterns", *Musicae Scientiae*, Discussion Forum 4B.

Olivier Lartillot. (in project). "Automated extraction of motivic patterns and application to the analysis of Debussy's Syrinx", *International Conference of the Society for Mathematics and Computation in Music*, 2007, Berlin, proceedings in project.

Olivier Lartillot and Mondher Ayari. (2008). "Segmenting Arabic modal improvisation: Comparing listeners' responses with computer predictions", *Conference on Interdisciplinary Musicology*, Thessaloniki.

Olivier Lartillot. (2008). "Automated motivic analysis of Brahms' String Quartet based on multi-dimensional closed pattern mining", *International Workshop on Computational Music Analysis*, Paris.

Olivier Lartillot and Petri Toiviainen. (2007). "Motivic matching strategies for automated pattern extraction", *Musicae Scientiae*, Discussion Forum 4A, pp. 281-314.

Olivier Lartillot. (2007). "Motivic Pattern Extraction in Symbolic Domain", in Jialie Shen, John Shepard, Bin Cui, Ling Liu (ed.), *Intelligent Music Information Systems: Tools and Methodologies*, IGI Global.