Olivier Lartillot
University of Jyväskylä, Department of Music
PL 35(A), 40014 University of Jyväskylä, Finland
lartillo@cc.jyu.fi

A Musical Pattern Discovery System Founded on a Modeling of Listening Strategies

Music is a domain of expression that conveys a paramount degree of complexity. The musical surface, composed of a multitude of notes, results from the elaboration of numerous structures of different types and sizes. The composer constructs this structural complexity in a more or less explicit way. The listener, faced by such a complex phenomenon, is able to reconstruct only a limited part of it, mostly in a non-explicit way. One particular aim of music analysis is to objectify such complexity, thus offering to the listener a tool for enriching the appreciation of music (Lartillot and Saint-James, 2004).

The trouble is, traditional musical analysis, although offering a valuable understanding of musical style, does not go into the deepest details of this complexity. Some approaches of 20th-century musicology, such as the thematic analysis by Rudolph Reti (1951), were aimed at a better awareness of complexity. However, their scope was still restricted to a particular aspect of musical structure. For instance, Reti's approach was founded on the hypothesis that a musical work is built on a single motive. And even within such limited scope, the search cannot be undertaken exhaustively, owing to the unreachable combinatory structure of musical works. Even worse, the results of such analyses do not meet a consensus agreement (Cook 1987), which questions the relevance of the underlying methods.

Nicolas Ruwet formalized motivic analysis as a set of partially detailed operations that carry out a top–down hierarchical segmentation of the musical work (Ruwet 1987). However, he never actually followed his model when applying it to concrete examples, but rather he relied implicitly on his own intuitions. In fact, a careful application of this method easily leads to absurd results that invalidate the model (Lartillot 2004). All this points to the necessity of a computational modeling of the discovery processes themselves. Indeed, a computer implementation of the modeling can explicitly show its potential capacities and pertinence and can be validated—or invalidated—according to its operational efficiency. Moreover, non-computational modeling implicitly tends to limit the complexity of algorithms and data structures for practical reasons. For instance, the grouping structure proposed by Lerdahl and Jackendoff (1983), which can successfully be implemented on a computer (Temperley 2001), relies however on the foundations of simple preference rules and the idea of a unique hierarchical system, which, as will be explained below, limits a detailed understanding of musical structure.

The necessary simplifications that are required by the non-computer modeling question the use of modeling itself, according to Reti:

> The thematic phenomena are so manifold and complex that in a sense they evade academic tabulation. Though they can perhaps be described, they can hardly be comprised in an actual ''system.'' They are too intimately connected with the creative process itself.
> (Reti 1951, p. 233)

However, computer modeling can transcend such limitations. Computer systems can escape centralized ''tabulation,'' as Reti calls it, and they can offer real ''creativity.'' Hence, computer modeling may radicalize the creativity of discovery procedures, improving the level of detail of the analyses and leading to the analysis of large databases.

Patterns and Local Groupings

The structural complexity of a piece of music may be decomposed into two dimensions. On the one hand, successive notes may be progressively aggre-

Computer Music Journal, 28:3, pp. 53–67, Fall 2004

gated into local groupings, because of temporal proximity and musical similarities (Lerdahl and Jackendoff 1983). Gestalt theory offers an account of the perception in terms of grouping of the stimuli followings rules of proximity, similarity, and "common fate." These local groupings may in particular segment the temporal dimension into chunks, or segregate the polyphonic flux into streams. Such local groupings are easily delimited if there actually exist strong discontinuities, e.g., rests, longer notes, significant pitch dissimilarities. But one may wonder whether such local groupings may be considered in a systematic and general framework. Indeed, when discontinuities become less pronounced, such segmentations seem to heavily rely on the listener's subjectivity.

On the other hand, different configurations in one or several musical pieces may be identified with one another and may therefore be considered as reproductions of a pattern. This identification seems to be more easily systematized, because it may be founded on explicit rules. The remainder of this article is dedicated to the modeling of such a pattern discovery and recognition process. Moreover, to avoid the difficulties of excessive complexity, the study is limited to patterns consisting of a simple succession of notes (i.e., motives).

Overview of Computational Pattern Discovery

In this section, I suggest that the difficulties faced by computational musical pattern discovery must be overcome with a close mimicking of human listening strategies.

Grouping Instead of Segmenting

The understanding of a musical sequence (where "musical sequence" refers to an entire musical score throughout the article), when considered incrementally, can be formalized in two opposite ways. On the one hand, the sequence can be segmented, in which the entire musical sequence is decomposed into a succession of non-overlapping and contiguous segments (Cambouropoulos 1998). This decomposition induces several significant constraints on the analysis. In particular, the analysis is considered as univocal and global: each note is associated to only one or two segments, and each segment not only results from its intrinsic characteristics but is also determined according to its neighboring segments.

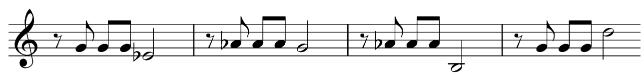
This mechanism of segmentation plays a key role in music. The local groupings, as we have seen previously, stem from a segmentation process. Moreover, this decomposition of musical sequences into phrases—in classical music in particular—is also founded on this idea of segmentation. This decomposition is ruled by a multitude of strategies founded on motivic but also metric and stylistic identifications. It seems, however, that the simple mechanism of motivic identification, which is the topic of this article, does not need to be limited by the constraint of segmentation. More generally, pattern discovery, a second mechanism by which musical sequences can be formalized, can be considered as a simple grouping process: notes are aggregated into groups whose existence is relatively independent from the existence of other groups at the same local context (e.g., two groups of notes may overlap, or they may be separated by other notes, in contrast to strict segmentation).

Similarity, Identity, and Viewpoints

A core property of pattern-discovery methods concerns the underlying basic criteria of identification. Usually, patterns are not exactly repeated but feature numerous kinds of transformations. A pattern-discovery algorithm based on exact identity of pattern occurrences will not be able to recognize transformed patterns. To take into account such characteristics of musical expression, two opposite strategies are mainly considered.

Identity between transformed patterns can be considered in terms of similarity or dissimilarity. Identification can then be inferred once the similarity exceeds a certain threshold (Cope 1991; Meudic and Saint-James, 2004), which may depend on the whole set of similarities (Rolland 1999). However, a general specification of this threshold does not

Figure 1. The first three motives are reproductions of a contour pattern. The four motives taken together represent a pattern that consists of the exact repetition of three identical notes at the same rhythm, followed by a different note at a longer rhythmic value.



seem possible. For instance, in Figure 1, no clear reason seems to justify any identification between the first two patterns that would exclude the third pattern from the identification. Indeed, one could set a threshold of interval size that would exclude the third pattern in this particular case, but it seems that such a fixed threshold could not be applicable to all cases.

On the other hand, identity may be considered along particular parametric "viewpoints" (Conklin and Anagnostopoulou 2001). The three first patterns of Figure 1 can be identified along the contour parameter, that is, the slope of the pitch interval between successive notes. And all four patterns can be identified together with a simple rhythmic identification. Hence, the idea of local relative identity seems to better describe the phenomenon of identification process in music listening than does similarity.

Computational approaches traditionally consider the different viewpoints independently: different analyses are run for each different projection of the original musical sequence along each possible viewpoint (Conklin and Anagnostopoulou 2001; Meredith, Lemström, and Wiggins 2002). However, it seems that the notion of viewpoints can be considered independently for each local aspect of musical expression. Indeed, a musical pattern, or motive, may be progressively constructed following different successive viewpoints. In Figure 1, the pattern that unifies the four motives consists of an exact repetition of three notes followed by a different note with a longer duration; the progressive discovery of the pattern therefore relies on different musical parameters (pitch and rhythm in this instance). This modification of the parametric space between the successive states of a pattern will be called *parametric transition*.

Incremental Construction Instead of Exhaustive Matching

Once the criteria of identification have been defined, another core question concerns the strategy of deployment of such criteria from the rough musical sequence. One method consists of a prior presegmentation of the sequence into a multitude of overlapping segments of different sizes. Patterns are then detected through an exhaustive matching of all these segments (Cope 1991; Rolland 1999). This method is computationally very expensive, and it forces patterns—and in particular their sizes—to remain within a specific domain defined by the presegmentation.

It seems, on the contrary, that patterns can be developed in a potentially infinite combinatorial space: in particular, their sizes are not constrained by any maximal threshold. Patterns should therefore be discovered through a constructive methodology. One common approach is founded on building a suffix tree (Conklin and Anagnostopoulou 2001). This method cannot be applied if identities, such as those proposed in previous paragraph, are adaptively detected in a multiply parameterized space. That is why patterns must be discovered through an incremental discovery algorithm, such as the notorious set-partitioning method, which progressively constructs all possible factors of a string (Crochemore 1981; Cambouropoulos 1998). The model presented in this article generalizes this incremental algorithm on a multiply parameterized space.

The Necessity of a Syntagmatic Approach

Patterns generally contain events that are successive: successive events within a pattern are temporally connected. This connection seems to play a paramount role in musical understanding. In this article, it will be called *syntagmatic relation*, following Saussurian linguistics. The definition of this concept becomes intricate in a polyphonic framework, where there may be simultaneous syntagmatic relations, and relations between chords.

Many current approaches, such as the one developed in this article, avoid tackling the whole complex polyphonic reality of music, and prefer to limit their scope to monodies. Other approaches, on the contrary, directly attempt to tackle polyphony. For instance, patterns can be constructed from successive notes, but also from successive notes that are situated in a particular time—for instance on pulsations, or on the first pulse of bars, and so on (Conklin and Anagnostopoulou 2001). These solutions, by limiting their scope to specific musical configurations, cannot be generalized any further.

Instead of the traditional definition of ''pattern'' as a repetitive sequence of syntagmatic relations between successive notes, one can consider an alternative representation of a pattern in terms of a repetitive translation vector between notes of the same position in several patterns. In this way, polyphonic patterns can be discovered (Meredith, Lemström, and Wiggins 2002). Its generalization (in particular to time-distorted patterns) will surely arouse intricate problems. Moreover, the lack of accounting for syntagmatic relations leads to false discoveries that must be filtered through an additional selective mechanism based on a notion of ''coverage.''

The importance of syntagmatic relations cannot be underestimated. Yet fully taking them into account remains a challenge. Syntagmatic relations can be considered among notes that are sufficiently close within the temporal domain (Rowe 1997; Rolland 1999; Dannenberg 2002). The influence of the pitch register must be considered, too.

Contextual Detection Instead of Global Selection

Most of the computational attempts toward automated pattern discovery that have been developed face the same problem: the set of discovered patterns is very large, and few of them feature actual relevance. That is why each of these approaches includes selective mechanisms that filter the set of possible results. In particular, specific constraints can be applied on the size of the patterns, their minimal number of occurrences, their maximal difference of size, the number of matched works, the global amount of overlapping, and so on. Such heuristics enable a restriction of the combinatorial, but this does not ensure the relevance of the analysis. Also, this global selection can easily discard interesting or even highly relevant patterns that do not meet the selection criterion, for example, a short pattern (such as the four-note motive in Figure 1) or a pattern repeated only two or three times successively.

This article develops another paradigm according to which the patterns discovered by the system should all convey certain relevance. A pattern can be considered as relevant when it matches the listener's intuitions. If all discovered patterns feature a non-negligible value, no further selection is necessary. The resulting set, instead of being reduced, could be browsed according to the user's criteria. For instance, one can look at the patterns related to a particular instant in the musical sequence or a particular pattern class. To fulfill this objective of total relevance of the results, the contextual principles of listening strategies must be considered carefully.

An Adaptive and Multidimensional Motivic Identification

In the model presented in this article, pattern discovery consists of a contextual grouping of syntagmatic relations based on an adaptive and incremental search for repetitions along multiple parameters. First of all, the basic rules of parametric identifications between syntagmatic relations must be detailed.

Melodic Dimensions

Scales

In tonal or modal music, pitches are generally not understood in terms of simple absolute pitches but rather as degrees of scales (Dowling and Harwood 1986). That is why motivic analysis should be founded on a prior tonal or modal representation. However, at the same time, this representation is inferred by the listener directly from the ''rough'' notes. The approach proposed in this article will not be able to consider the issue of scale discovery and will not import modal or tonal knowledge.

Pitch Interval

If no scale is predefined, the pitch of the considered note can be considered with respect to the pitch of the previous note. The chromatic interval between these two values, expressed in semitones, will be called *inter-pitch*. Patterns can be identified through an identification of inter-pitch values. It seems, however, that identification must be exact. Indeed, an identification between patterns featuring similar but unequal inter-pitch values does not generally result from the inter-pitch parameter itself, but rather from other musical parameters that will be considered below, such as contour and rhythm. For instance, when an inter-pitch sequence is repeated much later with a different rhythm, identification seems very unlikely unless the diatonic or chromatic intervals are identical. That is why, in the proposed model, two patterns can be identified along their inter-pitch dimensions only if they feature exactly the same inter-pitch value. This issue needs further investigation, in collaboration with the literature of experimental psychology in particular.

Contour

Alternatively, patterns can be identified through the contour parameter. This identification is possible even when inter-pitch values are very different. On the other hand, identification seems to be possible only for proximate intervals, and it does not generally play any role for distant repetitions owing to the poor qualities of contour concerning discrimination and memorization (Dowling and Harwood 1986). The exact determination of the temporal horizon for contour identification will once again benefit from experimental psychological investigations. For all of these reasons, in the current version of the proposed model, contour is theoretically integrated into the model but cannot practically be considered in the actual experiments.

Rhythmic Dimensions

The rhythmic dimension of motivic identification cannot be underestimated. Indeed, when inter-pitches are too dissimilar to be identified, and when patterns are too far apart in time to be identified via contour, rhythmic values can play a major role.

Meter and Rhythmic Values

The temporal dimension of music is usually perceived in terms of rhythmic values positioned on a pulsation frame determined by am underlying metrical grid. The trouble is, as far as music based on tonal and modal scales, metrical structure is not often explicitly given to the listener; the listener must usually deduce it from the musical sequence itself. The approach proposed in this article does not consider the issue of pulse discovery and does not import metric knowledge; hence, the exact determination of rhythmic values will not be obtained.

Inter-Onset Ratio

If no metric is predefined, rhythm may be considered simply in terms of a temporal distance, expressed in seconds, between successive notes, or *inter-onset*. Unfortunately, contrary to inter-pitch measures, inter-onsets do not seem to be perceived as such by the listener. Even if the pulsation cannot be inferred here, it is possible to get an implicit and adaptive account of the local tempo by considering the inter-onset value between two successive notes—which forms a syntagmatic relation—with respect to the inter-onset value of the previous syntagmatic relation. The ratio between these two numbers, commonly called *inter-onset ratio*, gives a local account of the rhythmic value. This relative definition of rhythm enables an identification of rhythmic augmentations and diminutions, that is, identification between occurrences of a rhythmic pattern at different tempos, or different time scales, as shown in Figure 2.

Actually, the determination by the listener of the inter-onset ratio seems to rely on a quantization of the actual ratio into a simple rational number. In a first approach, our proposed simplification will just consist of a rough decimal quantization of the values of inter-onset ratios. Moreover, regarding inter-

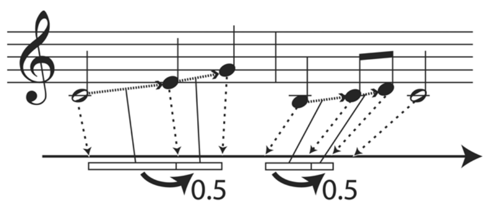


Figure 2. Local assessment of rhythmic value in terms of inter-onset ratio enables an identification of augmented or diminished rhythmic patterns.

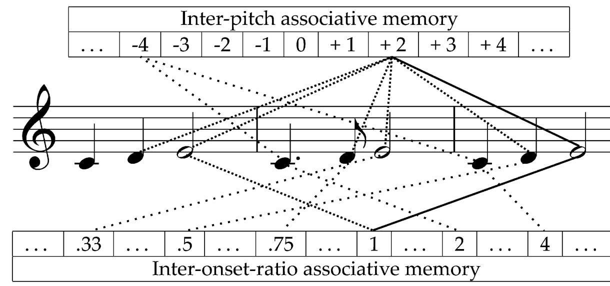


Figure 3. The inter-pitch memory associates with each known inter-pitch value the set of intervals that feature this value; the same occurs for the inter-onset-ratio memory. In this way, a new syntagmatic relation of inter-pitch value +2 can recall old syntagmatic relations of the same inter-pitch value.

pitch measures, it will be supposed that the identification between patterns through the inter-onset ratio parameter must be exact.

The Motivic Network

Music analysis can be considered as a memorization process. Some cognitive studies have modeled memory as a highly interconnected conceptual network (Holland et al. 1989). Each element can be individually retrieved by being activated through specific connections.

The Syntagmatic Relations

First of all, the whole set of elementary notes that constitutes a musical sequence is stored into a memory in a sequential order. All elements can be accessed through a progressive scanning along the temporal axis. The sequential access of the whole memory is enabled by the fact that successive notes are highly connected in memory: when the previous note is activated, the succeeding one can easily be recalled. This linking between successive notes is one of the core properties of what has been called a syntagmatic relation.

As explained previously, the exact determination of syntagmatic relations is far from evident. In a first approach, I will follow what may be called the strict monodic convention, stating that only one note is assigned to each temporal instant, and that syntagmatic relations are only assigned between notes of immediately successive time onsets.

The syntagmatic relation is not solely aimed at connecting successive notes; it becomes an important musical concept that will play a role in pattern discovery. It contains musical parameters that are deduced from the parameters of its elements. In the model considered here, the syntagmatic parameters are inter-pitch, contour, and inter-onset. Moreover, as noted previously, to each succession of two syntagmatic relations will be associated an inter-onset ratio. All of these parameters characterize the context that constitutes this succession of syntagmatic relations.

The Associative Relations

The analysis itself can be considered as a competence of the associative memory. Every structure that is discovered in the score stems from an *associative relation* among several contexts, which is not due to temporal proximity but rather to similarities along certain parameters. At each successive instant of music listening, each new context that is heard constitutes a query. This query induces a recall of similar contexts in the associative memory. Identification does not need to be total: a context can be retrieved if it shares only a partial identification, through one or several specific parameters, with the query.

The Associative Memory

For recall to be possible, to each musical pattern is associated a specific associative memory that assigns to each possible parameter value the set of all elements within the memory where this value holds. As shown in Figure 3, an inter-pitch memory is associated with each inter-pitch parameter; for each known inter-pitch value, the pairs of neigh-

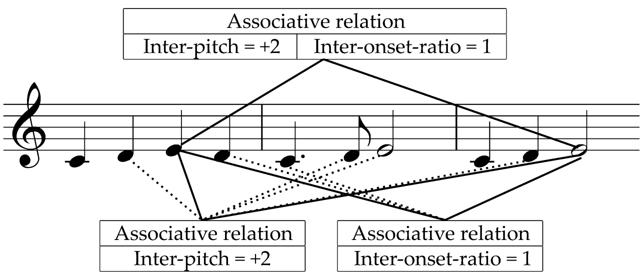


Figure 4. The associative class related to the conjunction of the inter-pitch value +2 and inter-onset-ratio value 1 consists of the intersection of the associative classes respectively related to an inter-pitch value +2 and an inter-onset-ratio value 1.

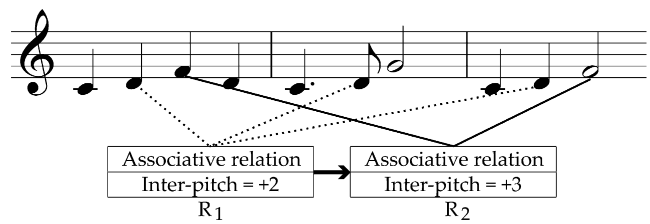


Figure 5. The succession of two associative relations R1 *and* R2 *forms a pattern.*

boring notes that feature this value are associated. The same holds true for the inter-onset-ratio memory.

In this way, each parameter value of a new description can reactivate old contexts of same parameter value through a simple inspection of the corresponding associative memory. It appears then that associative memory features the same property as hash tables, which can therefore be used for their implementation.

The Associative Relation

The associative memories reveal identities between remote contexts. These identities lead to the creation of concepts called associative relations. These concepts represent both the parameter that is identified in all of these contexts (called the *description* of the associative relation), and the set of all these contexts itself (called *associative class*).

These conceptual associative relations can be generalized to identities along conjunctions of parameters. The associative classes of such conjunctions of associative relations are the intersection of the associative classes for each single parameter value, as illustrated in Figure 4.

The Patterns

Syntagmatic relations that follow different occurrences of an associative relation *R1* may also share particular identities that form another associative relation *R2*, as shown in Figure 5. A new associative relation can then be constructed whose description is the succession of the descriptions *R1* and *R2*. This description represents the characteristic succession of parameters that is actualized by all the occurrences, and it can therefore be called a *pattern*. Successively, this pattern can be extended by a new associative relation *R3*, and so on.

Each successive associative relation can be defined in a specific parametric space. This general incremental model thus easily enables a formalization of the mechanism of parametric transition. Moreover, each time a new pattern is discovered, the extensions of its occurrences are memorized in an additional specific associative memory, as will be seen shortly. In this way, new pattern extensions will be discovered through a simple inspection of these hash tables.

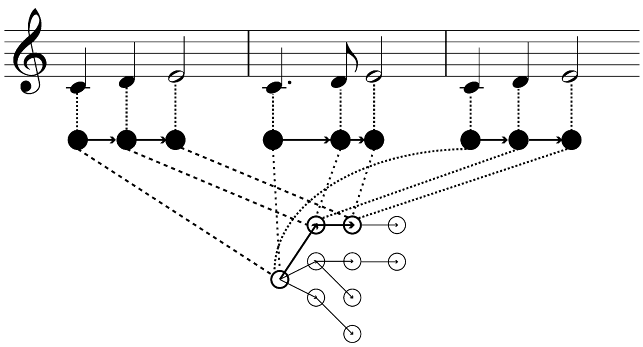
The Pattern Tree

Patterns are progressively discovered interval by interval. Each new extension of a pattern is discovered once an identification is detected between extensions of several occurrences of the original pattern. Even the elementary associative relationships can be considered patterns. For instance, inter-pitch or contour associative relations may be considered as extensions of a particular pattern, called a *note pattern*, which represents the simple concept of the note. Similarly, elementary inter-onset ratio associative relations are extensions of another particular pattern—a syntagmatic pattern—that represents the simple concept of the syntagmatic relation. The syntagmatic pattern may itself be considered as an extension of the note pattern. In this way, the set of all pattern classes constitutes a tree, as illustrated in Figure 6.

The Main Routine of the Program

The main algorithm incrementally considers each successive note n_i of the musical sequence, and it

Figure 6. Three reproductions of a pattern and the corresponding branch in a pattern tree (white circles) whose specific configuration in this particular example is arbitrary. Each pattern occurrence is also represented as a chain of states (in black), each successive state representing in the same time a successive note of the occurrence and a successive state within the branch of the pattern tree.



discovers pattern classes and occurrences of these pattern classes in a recursive way. At the initial stage of the recursion, each note n_i forms an occurrence of the elementary note pattern P_i. The set of all known extensions of this note pattern are the children of the corresponding root node of the pattern tree. Once the next note n_{i+1} is considered, if it can be identified as one of these possible extensions, say P'_j, then a new occurrence of this extended pattern class is created from the succession of the two notes n_i and n_{i+1}. Otherwise, this new note n_{i+1} will be memorized in the associative memory related to the eventual extensions of the note pattern. As can be seen in Figure 7, if this associative memory already features another similar context, a new extension P'_k of the note pattern P_i is discovered. The two notes n_i and n_{i+1} hence form a new occurrence of this new pattern class. These pattern occurrences and pattern classes can then be extended in a recursive way.

The Minimization of the Motivic Network

This model features underlying combinatorial properties that imply the necessary existence of additional heuristics ensuring an optimization of the representation through filtering of the redundant information. We will now discuss each of these additional heuristics in turn.

The Maximization of Specificity

The concept of pattern has been defined as a succession of identities, each successive identity being represented along one or several musical parameters independently of previous or successive identities. Unfortunately, different pattern descriptions can lead to the same set of pattern occurrences, or pattern classes, in terms of successions of specific associative relations, as shown in Figure 8. This redundancy leads to a combinatorial explosion.

Thus, a new heuristic must be introduced: two patterns cannot result in the same pattern class. As the analysis must be as precise as possible, only the most specific pattern should be inferred (shown as the bold branch in the tree of Figure 8). Indeed, a less-specific pattern can be considered as a direct consequence of the more specific one.

The Specificity Relation

The control of the combinatorial efficiency of the pattern-discovery procedures thus relies on the def-

Figure 7. When the extension of an occurrence of pattern P_i *can be identified with the extension of another pattern occurrence through the associative memory of the pattern extension (left), the pattern is extended and forms a new pattern* P'_k *(right).*

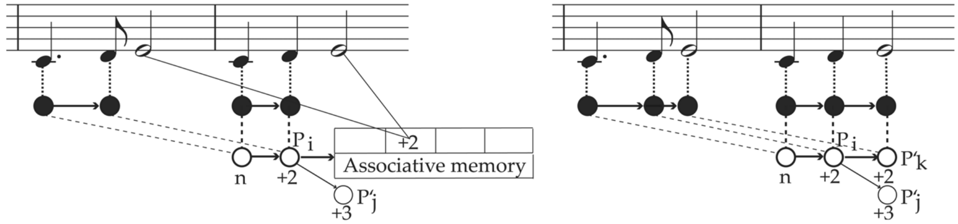
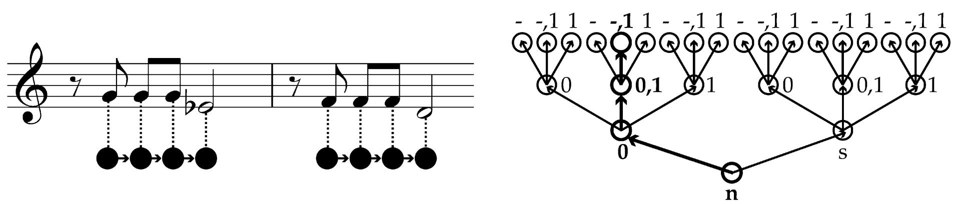


Figure 8. The two motives on the left can be considered as occurrences of 18 different patterns represented as branches of the pattern tree shown at right. The most specific branch is represented by the bold branch, which consists of the succession of a first note (a note pattern n*), a unison interval (inter-pitch = 0), a unison interval of the same rhythm (interpitch = 0, inter-onset-ratio = 1), and a decreasing contour of the same rhythm (inter-pitch = −, inter-onset-ratio = 1).*



inition of a relation of specificity between patterns. A pattern *P1* is considered more specific than a pattern *P2* when the whole description of *P2* is included in the whole description of *P1*. This means that each successive identity within the description of *P2* is defined along no greater number of musical parameters than in the corresponding identity within the description of *P1*. However, this also means that the description of *P1* may contain one or several successive additional identities that are added before the beginning of the description of *P2*, as in illustrated in Figure 9. *P2* can therefore be considered a less-detailed suffix of *P1*. In this way, the pattern *P1* offers a more complete description of its pattern classes through a more detailed and longer description of their contexts.

It should be pointed out that a prefix of a pattern, contrary to a suffix, does not offer a less-specific description of the context, but rather a description of a previous state of the context. The prefix is in fact actually represented as an intermediary state within the branch that represents the pattern in the pattern tree.

The Method of Maximal Description

As patterns are progressively discovered interval by interval, the heuristics of maximization of specificity must be applied to the incremental process. As previously explained, each discovered pattern is in fact an extension of its prefix. When a prefix pattern maximally describes a pattern class, the extension can, on the contrary, describe its pattern class non-maximally, as illustrated in Figure 9.

As mentioned previously, each new note is considered as a possible extension of all pattern occurrences that the previous note concludes. Following the heuristics of maximization of specificity, the set of pattern occurrences must be considered in a decreasing order of specificity. In this way, the most specific pattern is extended first, and all less-specific patterns are extended only if their corresponding pattern classes are not identified with the more specific ones. In addition, the different possible associative relations that extend such patterns must be considered in a decreasing order of specificity, from the total conjunction of all musical parameters—associated to the context of the new note—to the elementary musical parameters.

Thanks to this mechanism of maximization of specificity, patterns can be discovered in an effective way in a multi-parametric musical representation.

The Periodic Patterns

This modeling of listening strategies induces another particular type of overwhelming combinatorial redundancy. It is hypothesized that the control

Figure 9. P2*, although less specific than* P1*, represents a pattern class that is different from* P1 *and cannot therefore be discarded. The extension of* P2 *into* P4 *is still less specific than the extension of* P1 *into* P3*, but it leads to the same pattern class. Therefore,* P4 *should not be inferred.*

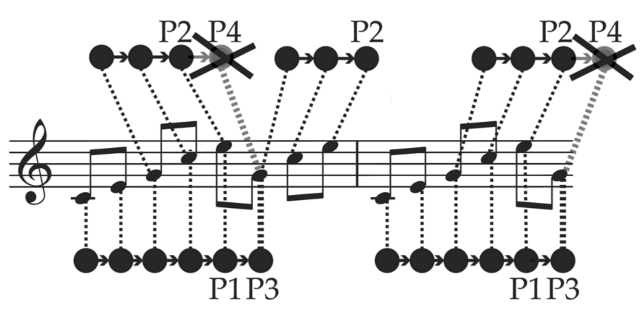
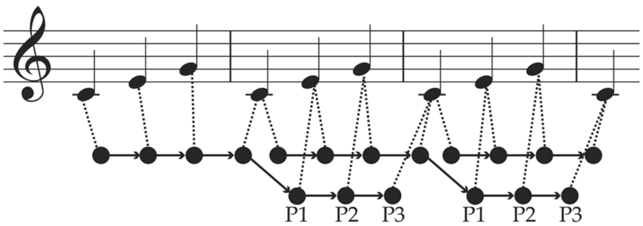


Figure 10. Successive reproductions of a single pattern lead to combinatorial pattern extensions, such as P1, P2, *and* P3.



of such redundancy by the listener is, once again, owing to the presence of specific cognitive strategies.

Combinatorial Cycling

When a pattern *P* is repeated several times successively such that each last note of one pattern occurrence is also the first note of another, then a large number of different patterns can be found. Indeed, the syntagmatic relations that succeed to each complete pattern occurrence can be identified, because they all belong to the beginning of the same pattern class. This equivalence among all such syntagmatic relations should therefore induce an extension of the pattern class, leading to the pattern class *P1*, as shown in Figure 10. This configuration appears once again for the extension of *P1* itself, leading to a new pattern class *P2*, and so on.

For all the structural complexity that appears in this configuration, human listeners are inclined to focus only on the periodic reproduction of the elementary pattern class. We must understand the implicit rules that lead to the decision.

Figure 11. The Non-Overlapping Heuristic prevents the extension of the first two occurrences of O1a *and* O2a, *but it does not prevent the extension of* O1a *with* O3a.

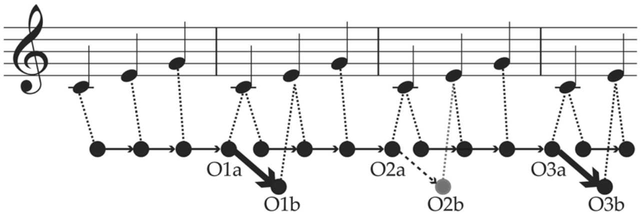
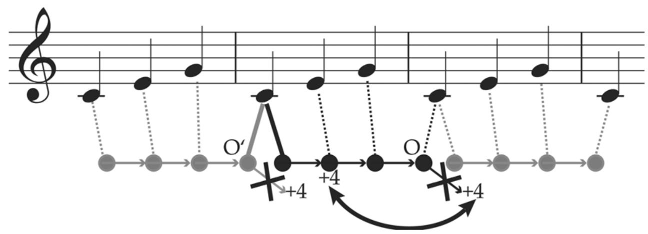


Figure 12. Occurrence O *is cyclic and will not be extended by the beginning of the following occurrence.*



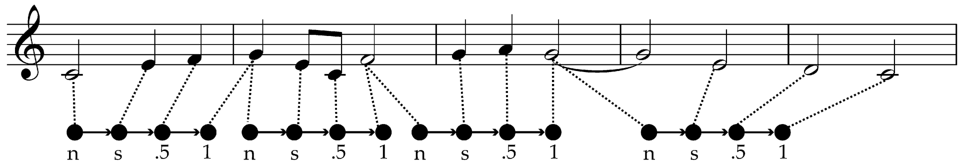
The Non-Overlapping Heuristic

The fact that the pattern class *P* is actually not extended into *P1* can be explained by the existence of a specific constraint that prevents occurrences of a single pattern class from overlapping each other. In Figure 11, the extensions of the occurrences *O1a* and *O2a* into occurrences *O1b* and *O2b* can be blocked thanks to such specific heuristics. Unfortunately, the occurrences *O1a* and *O3a* can on the contrary be extended into occurrences *O1b* and *O3b*, because they do not overlap. The scope of such non-overlapping heuristics is therefore very restricted.

The Cycling Heuristic

When the first note of a pattern occurrence *O* concludes another occurrence *O′* of the same pattern class, as illustrated in Figure 12, the pattern occurrence *O* is considered *cyclic*. In this configuration, an analogy is implicitly drawn between the last and first notes of the pattern and will influence the following discoveries. Suppose that an associative relation is discovered immediately following pattern occurrence *O*, such that its description corresponds—entirely or even only partially—to the description of the first extension of the same pattern from the first note to the second note. Consequently, the associative relation will be considered the beginning of a new occurrence of the same pattern, and it will not be accepted as a candidate for extension of the pattern occurrence *O*. The extension will be rejected both for the recognition of a

Figure 13. The successive reproduction of a single rhythmic pattern in different time scales will not provoke a combinatorial cycling. The initial state of a pattern is the note pattern n, and the second state of a rhythmic pattern is the syntagmatic pattern s.



known pattern extension and for the discovery of a new one.

For a given succession of reproductions of a single pattern class (what may be called a *periodic reproduction* of this pattern class), the first pattern occurrence, on the contrary, is not considered cyclic, because its first note does not conclude a previous pattern occurrence of the same class. Therefore, this first pattern occurrence can be extended. In this way, when a new periodic reproduction of the same pattern class appears after awhile, both periodic reproductions can be considered as occurrences of a larger pattern class whose description is the concatenation of the elementary pattern description.

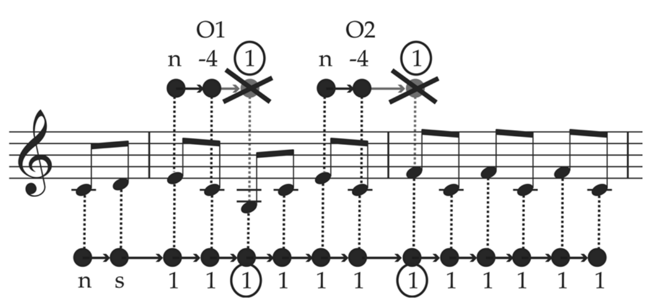
This issue of combinatorial cycling has been considered in one previous approach (Cambouropoulos 1998). Overlapping was considered as a general characteristic globally applied to the pattern classes themselves. It seems, however, that the combinatorial cycling concerns the local configuration of pattern occurrences and does not influence the existence of the pattern classes. Indeed, as we have shown, a pattern class that cannot extend cyclic patterns can nevertheless extend the first pattern occurrence of the periodic reproduction.

Rhythmic parameters must be considered separately. A successive reproduction of a rhythmic pattern will not automatically provoke a combinatorial cycling. Indeed, if each rhythmic pattern is presented in a different time scale (as illustrated in Figure 13), then no combinatorial cycling appears.

Motivic Foreground Against Periodic Background

Finally, the phenomenon of periodic reproduction induces another interesting phenomenon that must also be considered. In Figure 14, two occurrences *O1* and *O2* of a pattern *P* are superimposed on a single phase of a periodic reproduction. Any associative relation that follows these occurrences *O1* and *O2* whose description corresponds—entirely or even partially—to the expected associative relation of extension of the periodic reproduction should not be considered as a candidate extension of the pattern *P*. Otherwise, these pattern occurrences will automatically be extended by the entire remainder of the periodic reproduction. The necessity of rejecting the extension stems from the fact that the periodic reproduction induces a local background on which the pattern occurrence appears as a foreground figure. If the remainder of this occurrence strictly follows the background texture, then this means that the foreground figure has reached its end. This heuristic corresponds to the Gestalt rule of figure against ground.

Figure 14. The two occurrences O1 and O2 of a single melodic interval whose inter-pitch value equals –4 semitones (shown above the score) are included inside the periodic rhythmic reproduction of a single inter-onset-ratio value of 1 (shown under the score). O1 and O2 are foreground figures over the periodic background. As such, following the Gestalt rule of figure against ground, they cannot be extended by a simple rhythmic associative relation whose inter-onset-ratio value equals 1 (circled), because it comes directly from the background.



Current Results

As a consequence of such a mimesis of listening strategies, the relevance of most resulting discoveries is relatively ensured. In this way, global post-filtering mechanisms, which roughly reduce the level of details, can be avoided.

Implementation

These algorithms have been implemented in Common Lisp as a library called OMkanthus within OpenMusic (Assayag et al. 1999), a graphical programming language dedicated to the computation of symbolic representations of music. The analyses are carried out on musical sequences encoded in MIDI format. Primarily dedicated to the real-time control of musical instruments and not concerned with the theoretical musical framework underlying the musical sequence, the MIDI format shows interesting analogies with the musical information actually offered to the listener. Indeed, the sound resulting from the music performance actually contains a temporal disposition of quantized frequencies but does not explicitly feature the underlying theoretical pitches and rhythmic values. The latter must be reconstructed by the listener.

The MIDI representation is graphically represented in OpenMusic as a score in which pitches are reduced to natural and sharp values and the horizontal disposition of notes is proportional to their temporal configurations. In particular, the spatial distance between successive notes is proportional to their temporal interval. Examples of such a score are given in Figure 15.

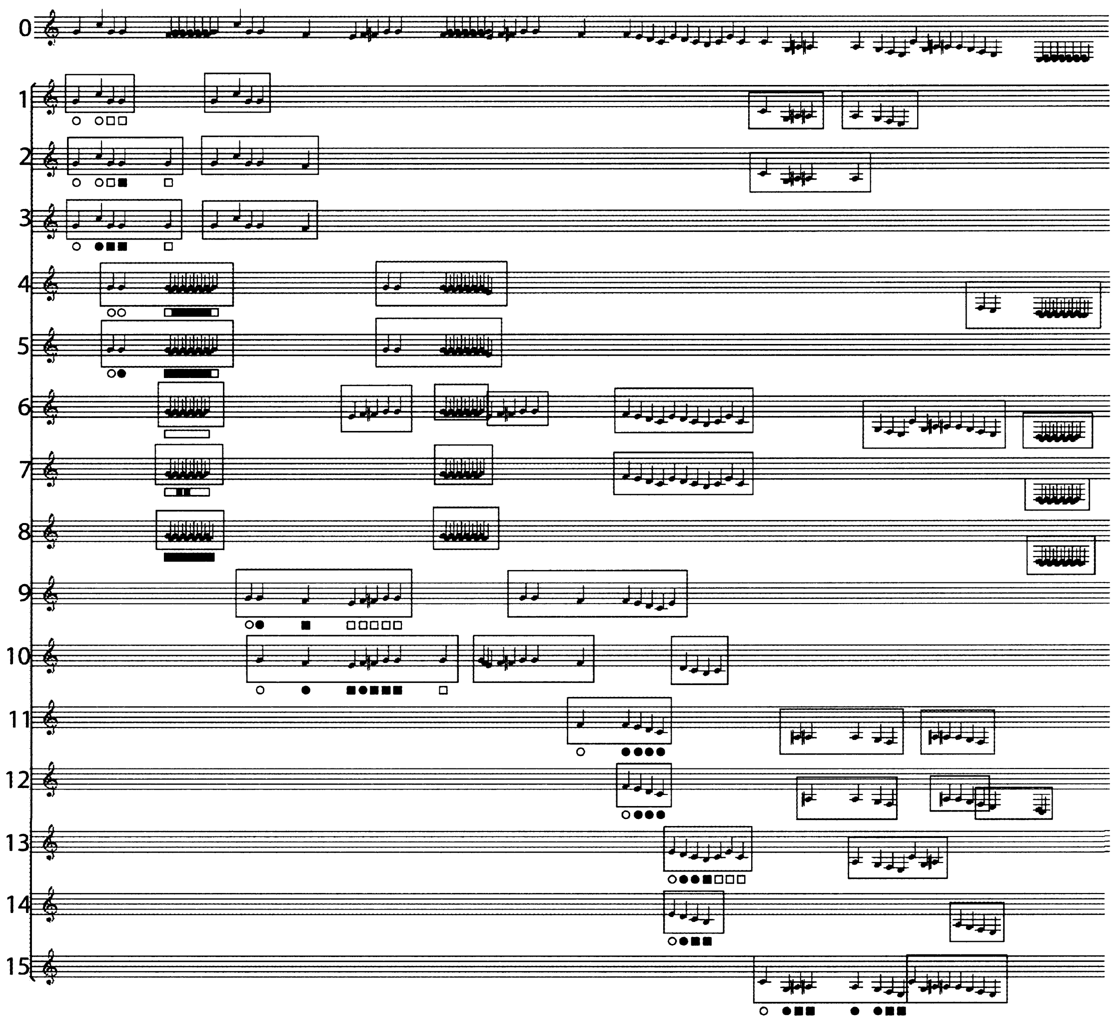
The results of the analysis can be displayed in two different ways. All successive pattern discoveries can be displayed as a list of instructions. This enables one to trace the exact processes actually undertaken by the system. Alternatively, the patterns can be displayed in a score composed of a superposition of several synchronized lines, as illustrated in Figure 15. Each line, representing a particular pattern class, displays all its associated occurrences.

The computational quality of the pattern-discovery system stems from its use of associative memories implemented as hash tables. Each successive syntagmatic relation that follows a certain number of pattern occurrences is memorized in the associative relations of extensions of these patterns. Similarly, each time pattern classes are extended, all syntagmatic relations following all pattern occurrences of these extensions are memorized in new associative relations. Every time a new syntagmatic relation follows a specific pattern occurrence, similar extensions will then be immediately recalled through a simple inspection of the corresponding associative memories. Moreover, the two additional heuristics presented above (the maximization of specificity and factorization of periodicity) prevent the development of combinatorial redundancy and insure a compact representation of the results as well as an efficient process.

Examples

The behavior of the algorithm can be observed through the analysis of the upper voice of Mozart's *Minuetto*, KV 1, displayed in Figure 16. The first line of Figure 15 shows the OpenMusic representation of the upper voice. Each successive remaining line represents a different pattern class in which pattern occurrences are squared. Below the first occurrence of each pattern class, the musical parameters responsible for the successive identities constituting the pattern are indicated. A white square indicates a rhythmic identity, a black circle represents a melodic one, and a black square indicates a melodic-rhythmic identity. A melodic identification requires at least two notes, and a rhythmic identification three notes. That is why the first note of each pattern features no particular characteristic but simply belongs to the note pattern, represented by a white circle. Similarly, the second note of a pattern, if it does not feature a melodic identification, simply belongs to the syntagmatic relation pattern, represented by a white circle as well. Finally, in patterns 4 to 8, the periodicity is represented by a single rectangle that extends through the entire pattern occurrence.

Figure 15. *Analysis of Mozart's* Minuetto, *KV 1.*



This analysis shows interesting results, in particular with rhythmic patterns 1 and 2; melodic-rhythmic pattern 3; and melodic patterns 11, 14, and 15. Pattern 6 simply represents the concept of rhythmic regularity. Patterns 9 and 10 illustrate one important claim of this article: motives can be constructed through a succession of identities stemming from distinct musical parameters. In particular, the second state of pattern 10 is purely melodic, as indicated by the black circle, whereas the last state is uniquely rhythmic, as indicated by the white square. An additional constraint relative to the continuity between successive parameters of the parametric transition enables the filtering of numerous irrelevant patterns. However, the mechanism of parametric transition needs further examination. For instance, patterns 2 and 7 are rhythmic but also feature a specific melodic identity at one or two isolated states. Such isolated characteristics do not seem to be perceived by a listener, as they lack salience.

Patterns 4, 5, and 9 show another problem that must be solved in future work. Each pattern begins with the last occurrence of a periodic reproduction. The previous occurrences are not included in those patterns owing to the cycling heuristic: each previous occurrence cannot be extended by another occurrence of the same periodic sequence unless this

Figure 16. Mozart's Minuetto, *KV 1.*

previous occurrence is actually the first occurrence of a periodic sequence. One strategy that seems to agree with listening intuition may be to initiate each occurrence of those patterns at the first occurrence of the corresponding periodic sequence.

Future Work

The system presented in this article introduces a new approach to musical pattern discovery that is founded on a mimesis of listening strategies. The results corroborate quite well with listeners' intuitions but are restricted to a particularly limited application domain. Taking into account more general aspects of musical construction will require extensive further studies.

The adaptation of this model to the analysis of scores, where theoretical values of pitch and rhythm parameters are explicitly represented, will enable a significant improvement of the results. In particular, diatonic transposition, which is a typical motivic transformation, would easily be identified. The automated reconstruction of the theoretical parameters through pitch spelling and quantization algorithms may also be integrated into the framework.

The concepts of parametric transitions, contour identifications, and cyclic patterns require further development that would benefit from collaboration with experimental psychology. The numerous problems that can still be discovered in the analysis generated by the system will reveal the existence of additional heuristics that must be integrated into the model. In particular, heuristics concerning the capacity of short-term memory or the influence of local segmentation on motivic discovery may be of importance.

Different kinds of transformations, such as inversion and retrograde, should be considered as well. The contribution of this transformation to motivic identification should be carefully studied. It may be hypothesized, for instance, that such transformations cannot be easily recognized by themselves without the contribution of other rhythmic or formal identities.

Moreover, formal characteristics play a pivotal role in the pattern identification. The first occurrence of a pattern already features relevant symmetries that might be noticed by the listener. The second occurrence might arouse the recall of the first one, and more generally of the pattern class, owing to the recognition of such formal symmetries. Thus, formal configurations should be explicitly represented in the pattern class description.

Polyphonic analysis must be progressively integrated into the system. In the monodic convention adopted in this article, the musical sequence is reduced to a simple syntagmatic chain. In general, however, as can be seen in Schenkerian analyses, monodies—and, more generally, polyphonies—can be better considered as syntagmatic graphs that can be covered by multiple syntagmatic chains. Patterns of patterns can also be discovered (Lartillot and Saint-James, in 2004).

Acknowledgments

This study is the result of my Ph.D. research (Lartillot 2004), conducted at l'Institut de Recherche et Coordination Acoustique/Musique (IRCAM) in Paris, supervised by Emmanuel Saint-James (University of Paris 6) and Gérard Assayag (Music Representation Group, IRCAM), reviewed by Jean-Marc Chouvel and Marc Chemillier, and examined by Jean-Pierre Briot, Christian Queinnec, and Nicolas Meeùs. I would like to thank all the reviewers and the members of the music information retrieval (MIR) community that have offered me significant help during the elaboration of these ideas and the writing of the article. Finally, I am grateful to Petri Toiviainen for having kindly granted the wonderful opportunity to continue my investigations at his university.

References

Assayag, G., et al. 1999. ''Computer Assister Composition at IRCAM: From Patchwork to OpenMusic.'' *Computer Music Journal* 23(3):59–72.

Cambouropoulos, E. 1998. *Towards a General Computational Theory of Musical Structure*. Ph.D. Thesis, University of Edinburgh.

Conklin, D., and C. Anagnostopoulou. 2001. ''Representation and Discovery of Multiple Viewpoint Patterns.'' *Proceedings of the 2001 International Computer Music Conference*. San Francisco: International Computer Music Association, pp. 479–485.

Cook, N. 1987. *A Guide to Musical Analysis*. London: Dent.

Cope, D. 1991. *Computer and Musical Style*. Oxford: Oxford University Press.

Crochemore, M. 1981. ''An Optimal Algorithm for Computing the Repetitions in a Word.'' *Information Processing Letters* 12(3):244–250.

Dannenberg, R. 2002. ''Listening to 'Naima': An Automated Structural Analysis of Music From Recorded Audio.'' *Proceedings of the 2002 International Computer Music Conference*. San Francisco: International Computer Music Association, pp. 28–34.

Dowling, W. J., and D. L. Harwood. 1986. *Music Cognition*. London: Academic Press.

Holland, J. H., et al. 1989. *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, Massachusetts: MIT Press.

Lartillot, O., and E. Saint-James. 2004. ''Automating Motivic Analysis through the Application of Perceptual Rules.'' *Music Query: Methods, Strategies, and User Studies* (*Computing in Musicology*, 13). Cambridge, Massachusetts: MIT Press.

Lartillot, O. 2004. *Fondements d'un système d'analyse musicale suivant une modélisation cognitiviste de l'écoute*. Ph.D. Thesis, University of Paris 6.

Lerdahl, F., and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, Massachusetts: MIT Press.

Meredith, D., K. Lemström, and G. Wiggins. 2002. ''Algorithms for Discovering Repeated Patterns in Multidimensional Representations of Polyphonic Music.'' *Journal of New Music Research* 31(4):321–345.

Meudic, B., and E. Saint-James. 2004. ''Automatic Extraction of Approximate Repetitions in Polyphonic MIDI Files Based on Perceptive Criteria.'' *Computer Music Modeling and Retrieval*. Berlin: Springer-Verlag.

Reti, R. 1951. *The Thematic Process in Music*. New-York: Macmillan.

Rolland, P.-Y. 1999. ''Discovering Patterns in Musical Sequences.'' *Journal of New Music Research* 28(4):334–350.

Rowe, R. 1997. *Interactive Music Systems*. Cambridge, Massachusetts: MIT Press.

Ruwet, N. 1987. ''Methods of Analysis in Musicology.'' *Music Analysis* 6:3–36.

Temperley, D. 2001. *The Cognition of Basic Musical Structures*. Cambridge, Massachusetts: MIT Press.