A generative model for the characterization of musical rhythms.

George Sioros[1], Matthew E. P. Davies[1], Carlos Guedes[1,2]

[1]INESC-TEC, Porto, Portugal
[2]New York University Abu Dhabi, Arts & Humanities, Abu Dhabi, United Arab Emirates

Correspondence: George Sioros, georgios.sioros@fe.up.pt

Keywords: rhythm, music analysis, meter, syncopation, transformation, timing.

Abstract

    We present a novel model for the characterization of musical rhythms that is based on the pervasive rhythmic phenomenon of syncopation. Syncopation is felt when the sensation of the regular beat or pulse in the music is momentarily disrupted; the feeling arises from breaking more expected patterns such as pickups (anacrusis) and faster events that introduce and bridge the notes articulated on the beats. Our model begins with a simple pattern that articulates a beat consistent with the metrical expectations of a listener. Any rhythm is then generated from a unique combination of transformations applied on that simple pattern. Each transformation introduces notes in off-beat positions as one of three basic characteristic elements: 1) syncopations, 2) pickup rhythmic figures and 3) faster notes that articulate a subdivision of the beat. The characterization of a pattern is based on an algorithm that discovers and reverses the transformations in a stepwise manner. We formalize the above transformations and present the characterization algorithm, and then demonstrate and discuss the model through the analysis of the main rhythmic pattern of the song "Don't stop 'till you get enough" by Michael Jackson.

1 Introduction

Many musical rhythms elicit a sense of periodicity and regularity in listeners. Simple or complex music, even when it is not repetitive, often evokes the sensation of a regular pulse (Parncutt, 1987, 1994) which is evident when we tap in synchrony with music. These pulses feature accents where some are perceived as stronger than others. These structured, accented pulses form our expectations about the timing of the rhythmic events and are at the foundation of musical meter (Honing, 2012; Jones, 2008; Jones, Moynihan, MacKenzie, & Puente, 2002; McAuley, 2010, p. 168).

Besides meter, the perception of musical rhythm involves another mechanism: the grouping of events, referred to also as serial grouping or "figural coding" (Parncutt, 1994, p. 412). While the strong and weak pulses of meter group non-adjacent events in a periodic way, figural coding groups adjacent events primarily based on their proximity in time. When we listen to a rhythm, strong interactions between these mechanisms and the pattern of physical durations occur in our mind as we interpret the rhythm by inferring a metrical context (Clarke, 1987a, 1987b; Fraisse, 1982; Honing, 2012). Manifestations of the interaction between our metrical expectations and the pattern of durations are the feeling of syncopation (Huron, 2006, p. 295) and anacrusis or "pickup" (Lerdahl & Jackendoff, 1981, p. 500).

According to Huron (2006, p. 295), an event in a relatively weak metrical pulse (off-beat) creates a higher expectation for an event in the following stronger pulse (on-beat). Our expectation can either be confirmed by a following event resulting in the proper binding of the two (London, 2012, p. 107), or be violated when the expected event does not occur, in which case a syncopation is felt (see also Huron & Ommen, 2006). In this sense, an event in a weak position is bonded to the following event in the strong position, or it syncopates when this bond is broken.

In this article, we present a model for the systematic analysis and characterization of musical rhythms based on Huron's (2006, p. 295) metrical expectation principle of weak-strong bonds described above. The model identifies the characteristic relations of music events to a given beat, by codifying their deviations from the on-beat positions as: 1) syncopations, 2) anacrusis or pickups, or 3) the mere articulation of a metrical subdivision.

Sioros and Guedes (2014) generate and analyse syncopation by anticipating events in weaker metrical positions. Such shifts of events were introduced in a qualitative manner by Temperley in his study of syncopation in rock music (1999). Here, we extend both approaches to include two more transformations besides the syncopation shifts: 1) figural shifts that generate anacruses (pickups) and 2) density transformations that insert new events. The three transformations together form a complete model for the characterization of a rhythmic pattern, in a manner similar to a Schenkerian analysis. The model is able to generate any rhythm starting from a metronome-like pattern articulating only beat positions. Each transformation brings a particular element to the rhythm. Thus, characterizing the pattern centers on the discovery of the unique set of transformations that generates it.

Our model is primarily motivated by its analytical applications in computer algorithms. The approach to rhythm analysis through transformations is well established. Several models are based on abstract mathematical transformations (Louboutin & Meredith, 2016; Meredith, 2014; Paiement, Grandvalet, Bengio, & Eck, 2007) or properties particular to a certain class of rhythms (G. Toussaint, 2002; G. T. Toussaint, 2013). Most noticeably, Lewin (2007) formalized generalized algebraic transformations that can be applied to various musical dimensions including the timing of events. Although, the transformations we present here could be expressed as Lewin's generalized transformations, we do not pursue an integration of our model into Lewin's theory. Our approach is different. We formulate

transformational definitions for music phenomena, such as syncopation, based on theories and cognitive principles of rhythm and meter perception. In this way, we aim to create a model that reflects a listener's interpretation of a rhythm in a metrical context and, thus, provide a musically meaningful codification of the patterns

We formalize the transformations as generative processes that introduce new elements to a rhythm but we do not explicitly define their inverse. Instead, our focus is in the characterization method that we developed which effectively reverses the transformations. For instance, while we define a transformation that introduces new syncopation to a pattern, we do not define an inverse de-syncopating transformation. Instead, all instances of syncopation are identified and the respective de-syncopated patterns are generated during the characterization of the pattern.

The characterization provided by the model is unique and unambiguous with respect to a regular beat upon which weak—strong bonds of the musical events are expected. Yet, the model does not exclude the interpretation of the same pattern according to a different beat. For example, a pattern might be understood as syncopated or not, depending on the placement of the beat duration relative to the pattern. Different listeners may infer different metrical contexts to interpret a rhythm (London, 2012), depending on a variety of factors, such as cultural background, music style, or other elements of the music besides the timing of the events. Often, the meter itself is intentionally ambiguous, so that such metrical ambiguity will be inherent in the characterization. Our systematic approach to rhythm analysis does not imply a unique interpretation based on the 'correct' meter of the music. In fact, the proposed model can serve as the basis for comparisons of different interpretations of a single rhythm that are based on different metrical expectations.

The remainder of the paper is structured as follows. In section 2, we present the transformations and the characterization algorithm upon which the rhythm model is based. In section 3, we demonstrate the model in practice using the example of the Michael Jackson song "Don't stop 'till you get enough". Finally, in section 4, we discuss important aspects of the model and of its similarities and differences to other relevant approaches.

2   Rhythm generation model

The rhythm generation model consists of a set of three generative transformations that, given the beat, can generate any rhythmic pattern and a characterization algorithm that identifies the previously applied transformations and reverses them. In subsection 2.1, we formalize the transformations and define the transformation vector, i.e. a consistent way of representing them. Subsequently, in section 2.2, we present the characterization algorithm. For the algorithm to effectively characterize a rhythm, the transformations need to be reversible in a unique and unambiguous way. To this end, we impose two reversibility constraints that we discuss in the final subsection 2.3.

The transformations are derived from the definition of syncopation as a cognitive mechanism related to metrical expectations and the weak-strong bonds between events articulated in weak-strong metrical positions (Huron, 2006, p. 295). Consequently, their formalization first requires a formalized representation of metrical expectations, i.e. a metrical template. We base our metrical templates on the stratification of meter by Yeston (1976) and on the similar definition of meter by Lerdahl and Jackendoff (1983). We use the algorithm devised by Barlow and Lohner (Barlow, 1987; Barlow & Lohner, 1987), which was adapted by Sioros and Guedes (2014) for the purpose of defining syncopation, to automatically generate the

template of any given meter. A metrical template consists of hierarchically grouped pulses as in the example of Figure 1.

A lower and upper threshold for the duration of the metrical subdivisions included in the template is set based on the metrical salience of the subdivisions (Parncutt, 1994). The lower threshold has been estimated in numerous studies to be around 100ms since faster durations are not perceived as metrical subdivisions (London, 2012, p. 29; Parncutt, 1994; Repp, 2006). The upper threshold for the template is set to the most salient subdivision, which we will hereafter call the beat, and is estimated around 1s (Parncutt, 1994). The prominent role given to the beat in our model follows from the weakening of the syncopation feel and of the weak-strong bond for slow metrical levels as discussed in detail by Sioros and Guedes (2014). In the example of Figure 1, the sixteenth-note and the bar levels are omitted as they fall outside of the above thresholds.

The characterization of rhythms presented here is only based on the quantized onsets of the musical events. As each onset in the rhythm belongs to a metrical position, the metrical template reduces a rhythm into a binary string. In the examples of this section we use binary rhythm representations. However, in the more musical example of section 3, the events have other properties besides their timing, such as pitch or lyrics. The transformation and characterization process preserve these properties.

The model consists of three main transformations: 1) the syncopation transformations, which we denote with the letter "S", that place events in weak metrical positions that are not properly bound to following events in strong positions and therefore syncopate (Sioros & Guedes, 2014), 2) the figural transformations, which we denote with "F", that result in the binding of events in weak-strong rhythmic figures of an anacrustic character, and 3) density

transformations, which we denote with "D", that insert events in weak metrical positions surrounded by events in stronger pulses.

Examples of the transformations are shown in Figure 2. Accordingly, syncopation is the result of the anticipation of an event from a strong metrical position to a preceding weaker one. In a similar manner, a weak-strong rhythmic figure is generated by delaying an event from a strong metrical position to a following weaker one, in order for it to approach a later stronger event. Density transformations do not involve the shifting of events. Instead, a new event is inserted between two existing events so that it articulates a faster metrical level.

In the framework of this model, we call a metronome any pattern that articulates only beat positions i.e. that comprises events only at the slowest metrical level included in the metrical template. Departing from a metronome, the transformations can generate any binary rhythmic pattern in a step-wise process, introducing a single new element of the rhythm with each step—either a syncopation, an anacrusis, or a new event. Conversely, given the metrical expectations that derive from a metronome, one can characterize a rhythmic pattern by identifying and reversing those transformations.

2.1   Transformation vectors

Each elementary transformation is characterized by the metrical level difference between the two positions comprising a weak-strong bond. In the case of syncopation, even though the bond is not realized, it is expected. For example, the syncopations of Figure 2 a) on half note level (pulse 8) are produced by the previous quarter (pulse 4), eighth (6) or sixteenth (7) notes. In fact, the difference of the metrical levels has been used as a means to quantify the feeling of syncopation (Longuet-Higgins & Lee, 1984). The anacruses of the figural transformations in Figure 2 b) are characterized by the same level differences; an event in the quarter,

eighth or sixteenth notes (pulses 4, 6 or 7) preceding an event in the half note position (pulse 8). Similarly, the inserted events of the density transformations in the last column of Figure 2 c) are placed a quarter, eighth or sixteenth note (pulses 4, 6 or 7) before the following half note (in pulse 8).

In general, given a metrical template, any of the generative transformations presented here, can be completely and uniquely defined by two components that together constitute the transformation vector: 1) a relatively strong pulse on which the transformation is applied, and 2) the metrical level difference from a preceding weaker pulse that can form a weak-strong bond. A weak-strong bond always consists of consecutive metrical positions, that is, all pulses between them must belong to weaker metrical positions (faster metrical levels) otherwise they would not correspond to the expected proper binding of events (London, 2012, p. 107). If we assign indices and metrical levels to the pulses as in the Figure 2, then a vector can be formed as: {pulse index i, metrical level difference d}. The symbol S, F or D should be used in front of the vector to distinguish between syncopation, figural, and density transformations respectively.

Syncopation

Generating the syncopation described in the vector S {i, d} consists in finding the first pulse k preceding pulse i that belongs to the weaker metrical level $l(k) = l(i)-d$ and shifting (anticipating) the event from pulse i to pulse k (Sioros & Guedes, 2014). In the example in Figure 2 a), the syncopation S {8, 1} is generated by finding the pulse that belongs one metrical level faster than pulse 8 and shifting the event of pulse 8 to that position. Since pulse 8 belongs to the half note metrical level, the event is shifted to the previous quarter note position. In the S {8, 2} example, the same event is shifted to the previous eighth note position.

Figural

Generating the weak-strong rhythmic figure described in the vector *F {i, d}*
consists in finding the first pulse *k* preceding pulse *i* that belongs to the weaker
metrical level *l(k) = l(i)+d,* in the same manner as in the syncopation transformation.
However, instead of anticipating the event of pulse *i,* we delay the event preceding
pulse *k.* The transformation takes place only if the event is shifted to a pulse weaker
than the one it initially belonged to. In the *F {8, 1}* example of Figure 2 b), pulse *k* is
the previous quarter note position (4) and the event preceding *k* is found at the
previous downbeat (0). The event is therefore delayed from pulse 0 to 4. Similarly, in
the *F {8, 2}* example, the event is delayed to the eighth-note position preceding the
second bar. In this way, the figural transformations create a weak-strong rhythmic
figure with the event in the following beat.

Density

The density transformation *D {i, d}* inserts a new event in pulse *k* preceding an
event in pulse *i.* Pulse *k* is determined by the metrical level difference *d* as in the
other two transformations. However, the event is inserted under two conditions: 1)
the event preceding pulse *k* must belong to a stronger pulse *m,* and 2) pulse *k* must
be the strongest pulse between pulses *i* and *m.* These two conditions ensure that the
inserted event will not constitute a simultaneous syncopation or figural
transformation; an essential requirement for the characterization algorithm and the
reversibility of the transformations as discussed in sections 2.2 and 2.3.

Two important remarks should be made. First, the vector *S {i, d}* in syncopation
transformations coincides with the actual shift of an event from pulse *i* to a
preceding weaker position. However, in figural transformations, the shifted event is
not found in the vector but it is obtained indirectly as the previous event from the
weak position determined by the vector. In the example of Figure 2, the solid arrows
representing the vectors and dotted arrows representing the shifts coincide for
syncopation but not for the figural transformations. Regardless, the way the

transformations and vectors are formalized ensures that each vector always describes a unique shift.

The second remark is related to the special case of ternary subdivisions of the beat as in the example in Figure 3. By definition, the transformation vectors refer to the relation of two positions of different metrical strength. While the two ternary subdivisions of a beat belong to the same weak metrical level, Palmer and Krumhansl provide evidence that there is a difference in their metrical strength. They found that the frequency of musical events in the different metrical positions (Palmer & Krumhansl, 1990, fig. 1) coincide with the theoretical strength of the corresponding positions for common time signatures. However, musical events were significantly more probable in the third and sixth eighth-note positions of a 6/8 meter than in the second and fifth positions. In other words, the second of the two ternary subdivisions of the beat was found to be metrically stronger than the first even though they both belong to the same metrical level. Furthermore, their results agree with goodness-of-fit judgments of temporal patterns (Palmer & Krumhansl, 1990, fig. 3) and with theoretical calculations (Barlow, 1987; Barlow & Lohner, 1987).

This finding suggests that a note articulated in the first of the two positions of a ternary subdivision of the beat is normally followed by a note in the subsequent second weak position, as in pattern E, forming a weak-weak rhythmic figure. Therefore, such a weak-weak rhythmic binding of events is of the same nature as the weak-strong binding of events discussed by Huron (2006) and London (2012, p. 107) so that it can be generated by a figural (pattern E, transformation F {4, 0}) or density (pattern F, D {4, 0}) transformation and that in the absence of the second note a syncopation is felt (pattern B, transformation S {4, 0}).

In addition, Povel and Okkerman (1981) argue that a rhythmic pattern as in Figure 3B —although it does not constitute a typical example of syncopation—is

"creating a special rhythmical tension" because of the ambivalence between the metrical accents and the accent caused by the short-long durations. In contrast, a long-short figure (as in Figure 3A) does not create any tension and agrees with the meter. The evidence from Palmer and Krumhansl (1990) and the syncopation transformations presented here suggest that this "special tension" is of the same nature as the more typical examples of syncopation with the difference that it is, perhaps, felt less strongly.

2.2   Characterization algorithm

The characterization of a rhythm is based on identifying and reversing any previously applied transformations, one by one, by shifting or removing the appropriate events. By identifying, we mean determining in a unique way the kind of transformation—whether it describes a syncopation, figural, or density—and the corresponding vector. The process of identifying and reversing a transformation on pulse i of a given pattern is shown in the flow diagram of Figure 4.

The first step in the algorithm (marked 1 in the flow chart) determines the second component of the vector that corresponds to a potential transformation. The weak-strong pair of pulses k and i must comply with the definition of the transformation vector so that all pulses between them must belong to faster metrical levels. If a valid weak event is found, then the algorithm proceeds in determining whether the vector corresponds to a syncopation, figural or density transformation. In the absence of an event in pulse i, a syncopation is identified.

In the opposite case, the algorithm distinguishes between a potential figural and density transformation. It looks for the strong pulse m that precedes the weak event and which would be its original position (marked 2 in the flow chart). If no such valid "origin" for the event is found, then a density transformation is identified.

The last step in all the above cases is the reversing of the identified transformation (marked *S, F* or *D* in the flow chart) and the output of the corresponding vector. This step is restricted by the reversibility constraints described in section 2.3. If reversing is not possible the algorithm outputs *NR* signifying the identified transformation cannot be reversed without first reversing a transformation on another pulse.

In our model, it is possible for a single pulse to undergo more than one transformation. All transformations in the pattern will be identified and reversed by applying the above algorithm recursively to all the pulses until no transformation is found.

2.3   Reversibility constraints

Reversibility is essential and, in many ways, shaped how the transformations are formalized, whereby any elementary transformation should always be reversible in a unique way. This ensures that, given a metrical template, any rhythm can be uniquely represented by a series of transformations. Conversely, it guarantees that a series of transformations can only generate a single pattern when applied to a metronome. Reversing a transformation is essentially a process of characterizing it. For example, de-syncopating is equivalent to characterizing the syncopation that is to be reversed. Having more than one way to reverse a transformation would mean that there is more than one way to characterize it, and thus the analysis of the pattern would be subject to interpretation. Therefore, the reversibility rule makes for a consistent model.

We impose two constraints on the transformations presented above to ensure that the above reversibility requirement is always met:

1) The order of events must always be preserved.
2) Each shift must correspond to a single transformation and a single vector.

In the following, we discuss the implications of the above constraints using appropriate examples.

The first constraint, which requires that for a shift of an event to be possible between two pulses, all in between pulses must be silent (i.e. not carrying an event) is discussed in detail by Sioros and Guedes (2014). In Figure 5, we provide an example of how reversibility is broken if this constraint is not adhered to. If we imagine the transformation S {4, 1} on pattern A of the figure, the event of pulse 4 would be shifted to pulse 2 over the event in pulse 3. However, pattern B can be the result of the S {4, 2} transformation on pattern C. The reversibility constraint requires that pattern C be the result of only one of the two transformations. As the syncopation found in pattern B is de-syncopated to pattern C, the S {4, 1} transformation must be forbidden, else it would break the reversibility constraint.

The second constraint dictates that generating or reversing a transformation should neither generate nor reverse another transformation simultaneously. If such a step were allowed, then the transformations would not correspond to a single vector. It is precisely this constraint which ensures that each transformation introduces a single distinct element in the rhythm. The conditions under which the figural and density transformations can take place are the result of this constraint. We demonstrate the way it is applied and discuss some of its implications in the context of the two examples presented in Figure 6 and Figure 7.

Figure 6 is an example of a figural transformation that would simultaneously generate a syncopation. If the *F* {4, 2} transformation were allowed, the weak-strong bond between the events in pulses 2 and 3 would be broken. In our model, such improper binding can only be the result of a syncopation transformation, which in this example could only be the transformation *S* {2, 1} applied on pattern C. In contrast, a weak-strong bond between two events can arise from either a figural or a density transformation. In pattern B, the bond between the two last events can be

generated through the density transformation D {4, 2} instead of a figural prior to introducing the syncopation S {2, 1}. In this way, each element in the rhythm corresponds to a unique weak-strong bond. When trying to characterize a potential figural transformation, the constraint takes an equivalent form: reversing a figural transformation is not allowed if it simultaneously removes an existing syncopation.

The second example, which is shown in Figure 7, refers to the potential masking of syncopations. In this case, the second constraint dictates that de-syncopating one instance of syncopation should not generate another, or in its reverse form, that generating a syncopation should not de-syncopate another syncopation. Pattern C contains a syncopation (felt on pulse 4) that could be de-syncopated as shown in the right side of the figure. However, this would result in the generation of a syncopation on pulse 2 (pattern E). Instead, we should first reverse the figural transformation (pattern B) and then reverse the syncopation on pulse 4, leading to pattern A. Conversely, the syncopation S {4, 1} cannot be generated on pattern E as it would mask the existing syncopation S {2, 1}.

The last example addressed the issue of applying a series of transformations to a pattern. We will demonstrate the process of characterizing a rhythm that has undergone a series of transformations using the example of Figure 8. In the figure, besides the detailed transformation vectors shown on the binary representation of the patterns, a simplified representation of the vectors is shown on the staffs above the corresponding metrical positions. In this representation, the symbol depicts the kind of transformation (F: figural, S: syncopation, D: density) and the number the metrical level difference (e.g. F {4 ,1} -> F1, or D {4, 2} -> D2). The pulse index is omitted as it is implied by the placement of the vector at the corresponding metrical position.

Starting with pattern A in Figure 8, we scan the pulses from left to right and reverse the transformations following the process detailed in section 2.2:

- Pulses 0 to 3 have not undergone any transformations since they are not preceded by events in weaker metrical positions.

- Pulse 4 carries an event which is preceded by an event in the weaker pulse 3. As the preceding stronger pulse 2 is not empty, a density transformation is identified and reversed.

- The next pulse preceded by a weak event is pulse 10. In this case, the weak event in pulse 9 is preceded by an empty stronger position (8) so that the event is shifted there reversing the figural transformation F {10, 1}.

- Next, a syncopation is found in pulse 12: the silent pulse 12 is preceded by the event in the weaker pulse 10. The syncopating event is shifted to pulse 12 yielding the vector S {12, 1}. This shift is only possible because the figural transformation F {10, 1} has already been reversed in the previous step. Had it been otherwise, reversing the syncopation would violate the second constraint as it would simultaneously generate a new syncopation in pulse 10.

- Continuing in the same fashion to the following pulses, the syncopation S {16, 2} is identified and reversed, followed by the density transformation D {20, 1}.

- Starting the process again from the beginning, pulse 4 is now preceded by the weaker event in pulse 2. This time, the transformation is identified as figural because of the lack of an event on the previous stronger beat (pulse 0).

- Reaching pulse 16, the weak event between the fourth and fifth beat is removed yielding the density transformation D {16, 1}.

- Now all events in the pattern belong to beat positions and the algorithm cannot identify any further transformations. The characterization of the pattern is therefore complete.

In Figure 8, a pattern labelled "Reduced" is shown. This pattern is obtained by only generating the possible density transformations on the metronome without generating any figural or syncopations resulting in the events of the two voices being aligned in their metronomic positions. The reduced pattern is an intermediate step between the initial pattern and the metronome, which preserves non-explicit rhythmic features (e.g. harmony or melodic intervals) while discarding the syncopations and pickups. A further reduction of the density transformations may remove essential non rhythmic features in many patterns. For instance, removing a chord from a weak metrical position could drastically change the harmonic structure of the phrase.

The characterization process can be followed for any pattern until the metronome corresponding to the given metrical template is reached. The order in which the pulses are scanned and the transformations are found and reversed does not affect the final outcome. The recursive nature of the process together with the reversibility constraints ensure that the given pattern can be generated only by the set of transformations found during the characterization.

However, when generating the initial pattern from the metronome, some transformations depend on others and must be performed in the order they were found. In Figure 8, the syncopation $S\{12, 1\}$ must be performed before the figural $F\{10, 1\}$ which depends on the existence of the syncopating note on pulse 10. On the other hand, other transformations, such as the $D\{20, 1\}$, can be applied in any order without affecting the outcome. In general, transformations that share at least one event in their weak-strong pairs form compound transformations and their shifts must be performed in a certain order to ensure the reconstruction of the pattern. Another type of dependence between transformations that do not share a common event is found between $D\{16, 1\}$ and $S\{12, 1\}$. Although each transformation does not depend on the existence of the other, when both are generated, they must be in the

correct order, otherwise the density transformation would not be possible. Because of the reversibility constraints, the characterization algorithm can only reverse a transformation if no other transformation depends on it, thus ensuring that the result of the characterization process will always respect the order of the dependence of the transformations.

All the patterns that originate from the same metronome will lie on a tree-like lattice similar to the syncopation branches and trees presented by Sioros and Guedes (2014). In Figure 9 we present an example of such a "rhythm" tree. The example illustrates in practice two important aspects of the model and how the reversibility constraints result in a unique characterization of the patterns. First, all patterns that can be generated by a metronome will be found on the same tree. Second, certain patterns can be generated through more than one branch. Nevertheless, these branches always share the exact same transformations in a different order.

All the allowed transformations are represented by arrows connecting two patterns. Generating the transformations follows the direction of the arrows; characterizing a pattern follows the opposite direction. As some patterns can be the result of more than one independent transformations, which transformation is reversed first depends on which pulse we chose to characterize first. In any case, a complete characterization of a pattern follows the entire path until the metronome is reached and all transformations are reversed. These properties are general to the model and not specific to the example; they are common for trees constructed for metronomes with any metrical subdivisions (binary, ternary etc.) and for any concatenation of beat durations.

Sioros and Guedes (2014) presented an algorithm for de-syncopating any pattern and obtaining its non-syncopating counterpart, called the *root*. Similarly, the characterization process presented here generates a branch of a tree connecting the pattern to a metronome. The rhythm tree resembles the syncopation tree of Sioros

and Guedes with two important differences: 1) the root of the tree is not merely a non-syncopating rhythm but a metronome that articulates only the beat positions of the given meter, and 2) the branches hold every possible pattern with the same number of beats and not a small collection of patterns with same number of events and similar rhythmic figures.

3   Don't stop till you get enough

In this section, we demonstrate some of the strengths of the model by characterizing and discussing the main rhythmic pattern of the song "Don't stop 'till you get enough" by Michael Jackson. The piece has several interesting properties which make it well suited for demonstrating the strengths of our approach. In Figure 10, the transformations of the original pattern to the metronome are shown, passing through the reduced pattern that contains only density transformations. For clarity, only the simplified transformation vectors are used in the figure, in which the single numerical component is the metrical level difference. The pulse index is implied by the positioning of the vector at the corresponding metrical position (e.g. D {12, 1} -> D1 above the third beat of the melody).

The two bars of the melody share several similarities and some important differences. Examining the melody, we can see three repetitions of a pattern of two sixteenth notes on weak metrical positions (marked with a circle on pulses 6, 22 and 30). The two first repetitions are identical in both their metrical positions and note durations. The third repetition differs only in the duration of the second note. The characterization of the transformations tells us that each repetition has a different character. In the first one, the notes are regarded as an anacrusis of the eighth note position leading to the following beat (transformation F1). The second repetition has the character of a metronome that speeds up as it runs on faster metrical levels as we approach the following beat. The two density transformations D1 and D2 articulate

the two faster metrical levels but no other elements are introduced. The third repetition is regarded different as well, this time as a syncopation (*S*2) to the "one" following on the next downbeat. It is notable that the third repetition is not an anacrusis even though the previous beat position (pulse 28) does not carry an event (in contrast, in the first repetition the same lack of event in the previous beat leads to the characteristic pickup feel). The above differences in the character of each repetition are encoded as the different transformations that generate each repetition.

Examining the short melodic figure (A-C#-F#) that appears twice, once in each bar, we can observe how figural transformations and syncopations have the characteristic effect of emphasizing the corresponding strong positions while density transformations have a less characteristic effect. The first time the melodic figure appears, emphasis is put on the second beat of the bar through the pickup rhythmic figure *F*{8,1}. At the same time, the density transformations resemble a slowing down metronome that runs on slower and slower levels during the second half of the bar (*D*2 -> *D*1 -> No transformation). In the second bar, the syncopation on pulse 28 and the lack of a pickup in the previous beat (pulse 24) take the emphasis towards the third beat of the bar. In other words, replacing the figural transformation with a density transformation on the second beat and introducing a syncopation in the third beat, effectively shifts the stress from one to the other.

The bass line emphasizes each beat in a different way (except the last beat in each bar which is not transformed), giving no clear evidence on where the downbeat is located; a feature characteristic of the "disco" feeling (Danielsen, 2012, pp. 156–157). The melody on the other hand, has a strongly emphasized downbeat more characteristic of funk (Danielsen, 2012, pp. 158–159). The syncopation *S*2 on the downbeat of the first bar is followed by the pickup F1 two beats later, in the middle of the bar, resulting in a clear and strongly felt metrical hierarchy. Beats are emphasized in a complementary way in the bass and melody. When there is a

syncopation in one, there is a figural (or no transformation) in the other, but no syncopation coincides in both rhythmic layers, resulting in a very clear metrical feel.

The same principles could easily be applied to other pieces of music. Conclusions such as the above can be directly and systematically drawn through the transformation vectors without the need to examine the actual rhythmic patterns. The transformation vectors show the characteristic elements of the rhythm, on which beats and metrical positions emphasis and stress is put and how the rhythmic layers relate to each other throughout the duration of the music.

4   Discussion

In this paper, we have presented a model for the analysis and the characterization of rhythms that codifies patterns as transformations of a simple metronome that articulates only beat position. The transformations are categorized as 1) syncopation, 2) figural, which create anacruses, and 3) density, which insert events. The formalization of the transformations was based on the transformation vector, a concept previously introduced by Sioros and Guedes (2014). The characterization of a rhythm is done by reversing the transformations found in a pattern, through an algorithm that identifies them and outputs the corresponding vectors. The algorithm ensures that a single series of transformations describes a unique rhythmic pattern.

The model consists of transformations that take the form of shifts of events. While these shifts are not intended to reflect the composer's or performer's methods and intentions, they serve as a means to characterize and interpret a rhythmic pattern with regard to the metrical expectations of a listener. Each shift in the model introduces a specific musical phenomenon in the rhythm which contributes to its character. For instance, anticipating an event in a weak metrical position introduces a syncopation. In this way, the shifts are used as an effective technique to reach a

meaningful reduction of the rhythm. At the same time, they preserve other properties and qualities of the events besides the timing, such as the pitch, timbre or lyrics.

The transformations were inspired by the syncopation model proposed by Temperley (1999). In his model, a pattern is represented by a music surface that is a transformation of an underlying deep structure. The two are connected by shifts of certain notes, from their metronomic positions in the deep structure to offbeat syncopating positions in the music surface. According to Temperley, transformations such as the syncopation shifts have a local character and a strong cognitive ground in comparison to other Schenkerian approaches to rhythm reduction (Komar, 1971; Schachter, 1980). Temperley draws evidence from music theory, harmonic analysis and rhythm perception to suggest that an unsyncopated deep structure and a syncopated musical surface correspond to actual cognitive representations of rhythm. In our model, the cognitive mechanism of weak-strong bonds described by Huron (2006, p. 295) lies behind such shifts.

On the same grounds, we speculate that the transformations, the metronome and the reduced pattern provide meaningful rhythm representations that correspond to listeners' interpretations of rhythms. Different listeners form different metrical expectations in response to a rhythm that affect and determine their experience of music. In our model, the listener's metrical expectations are modelled in metrical templates that constraint the transformations and determine the characterization of a rhythm. Thus, several characterizations for the same pattern are possible that reflect different listeners' expectations. The automatic construction of templates of section 2 is limited to common time signatures, nevertheless any well-formed meter, including non-isochronous meters (London, 2012), can be expressed in a metrical template. However, a systematic evaluation of the model with respect to rhythm perception is outside the scope of this article and is left for a future study.

Our model also shares similarities with the generative models developed by Temperley (2009, 2010) and his metrical anchoring principle. Anchoring characterizes events in weak metrical positions depending on the existence of events in their surrounding stronger pulses as: 1) both-anchored when events exist in both the previous and the following stronger pulses, 2) post-anchored when only the following pulse contains an event, 3) pre-anchored when only the preceding pulse contains an event and 4) unanchored when neither the receding nor the following stronger pulses contain events. Accordingly, a syncopation transformation generates a pre- or an unanchored event, a figural transformation a post-anchored event and a density transformation a both-anchored event.

However, the two models are not equivalent as they focus on different aspects of rhythm. Temperley (2010, p. 372) views pre- and unanchored notes as modelling the syncopation in patterns, but syncopation itself is as an overall quality of the rhythm related to the likelihood of a rhythm to appear in a certain metrical context. A pattern is felt as more syncopated than another due to a higher number of unanchored notes, but those notes do not necessarily correspond to instances of syncopations. Similarly, post- and both-anchored notes contribute to the overall metrical feel but are not viewed as instances of a rhythmic phenomenon. In contrast, our transformations model and generate specific phenomena that have a local character. Syncopations and anacruses are felt at particular moments and have short durations. The overall metrical and syncopation feel could be attributed to their combinations in a rhythm, but such a mapping would depend on cultural factors or the musical context and style.

For instance, the sixteenth-note found on pulse 30 of "Don't stop 'till you get enough" is unanchored as it is not surrounded by events in the stronger metrical positions. A direct equivalence between anchoring and syncopation would lead to the wrong conclusion that the note on pulse 30 is syncopated. In our model, this is

attributed to the syncopations on the surrounding beats. The note simply *happens* to occur between two syncopating sixteenth notes at pulses 27 and 31 which are pre-anchored.

The set of transformations that correspond to a rhythm carry all its characteristic features with respect to the given beat. On the one hand, the combination of syncopations and pickups describe the timing of the events in relation to certain metrical expectations. On the other hand, the density transformations articulate fast metrical levels as a metronome that varies its speed running at different metrical levels. The reduced pattern represents such a variable metronome. The specific combination of the three elements makes each pattern unique. For instance, in the "Don't stop 'till you get enough" example presented in Section 3, the two different bars of the melody originate from the exact same metronome. Their differences are encoded into the different sets of transformations that generate them. Motifs with similar inter-onset intervals on the musical surface, such as the repeated sixteenth notes, become distinct as their individual feel and character is brought to the foreground through the characterization process.

The encoding in transformation vectors gives the model and the representation of transformations great flexibility. On the staffs of Figure 8 and Figure 10, the transformation vectors are simplified by omitting the pulse index and showing only the metrical level difference. The pulse is implied by their placement above the corresponding metrical position. This simplification of the vector reveals an important aspect of its nature. After the characterization of a pattern is complete, the transformation vectors can be separated from the metrical template. Their metrical position can be referenced using appropriate references for each music material such as the beat number together with phrase boundaries or simply placing them in a time line as in Figure 8 and Figure 10.

A similar encoding based on syncopation transformations alone was initially developed by Sioros and Guedes (2014). Our model is based on the same concept of transformation, extending it to include the density and figural transforms, and thus resulting in a complete model for binary or quantized musical rhythms. The introduction of the two transformations has a strong impact on the syncopation transformations as well which is evident in the example in Figure 7. This earlier syncopation model allowed for consecutive syncopations to mask or hide one another, which posed no problem considering the focus on generative applications. However, in our complete model, the focus is shifted towards a more meaningful and accurate characterization of rhythms, and such "hidden" syncopations are avoided using a combination of figural and syncopation transformations.

While this paper introduces the model and its general aspects as a method for the analysis of musical rhythms and not any of its specific applications, we now present some of potential uses and future developments. The automatic and systematic encoding of each rhythmic pattern into a unique set of transformations can be of great value in musicological analysis. A systematic and automatic comparison and classification of music excerpts as well as measures of similarity or distance between patterns can be based on the common transformation vectors the excerpts share. As opposed to the more abstract mathematical properties used in the measures compared by Toussaint (2004), the transformation vectors describe music elements such as syncopations and pickups, rather than distance and statistical properties of binary rhythmic representations. Their comparison should depend on the intended applications and appropriate weights should be used according to the music material in question. On the other hand, the vectors and tree structures are not designed to describe an optimal transformational path between two patterns in the fashion of the edit distance measure proposed by Post and Toussaint (2011). As we discuss in section 2.3, the reversibility constraints impose a certain order for the

transformations which could force the path between rhythms to pass through unnecessary intermediary patterns, but with the advantage of providing a unique characterization for each pattern.

Prototypical transformations characteristic of a style or genre can be modelled and used in the classification of rhythms according to whether they share the same branch. Even the relation between different styles can be explored by examining the transformation path between style prototypes.

The rhythmic variations that are generated as a by-product of the characterization of a pattern can assist the automatic analysis and comparison of other properties of the events. For instance, in the "Don't stop till you get enough" example, we can directly observe in the metronome and reduced pattern that the harmony and melodic structure of the two bars is identical. Their differences are the result of a different rhythmic interpretation of the melody. Such analyses and comparisons can be facilitated and automated for any multi-layered rhythm. The notes of the various rhythmic layers are aligned in the generated metronomes and reduced patterns, so that their relations are simplified. A systematic analysis of these simplified patterns instead of the complex and often idiosyncratic original rhythms is easier and more effective.

In conclusion, the model presented here effectively combines powerful features and has a great potential in music analysis. It codifies rhythms in a systematic and flexible way, while at the same time provides with a musically and perceptually meaningful and detailed characterization, and the generative nature of the model has added value beyond the scope of rhythm analysis.

5   References

Barlow, C. (1987). Corrections for Clarence Barlow's Article: Two Essays on Theory. *Computer Music Journal*, *11*(4), 10.

Barlow, C., & Lohner, H. (1987). Two essays on theory. *Computer Music Journal*, *11*(1), 44–60.

Clarke, E. F. (1987a). Categorical Rhythm Perception: an Ecological Perspective. In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music* (pp. 19–33). Stockholm: Royal Swedish Academy of Music.

Clarke, E. F. (1987b). Levels of structure in the organization of musical time. *Contemporary Music Review*, *2*(1), 211–238. https://doi.org/10.1080/07494468708567059

Danielsen, A. (2012). The Sound of Crossover: Micro-rhythm and Sonic Pleasure in Michael Jackson's "Don't Stop "Til You Get Enough"." *Popular Music and Society*, *35*(2), 151–168. https://doi.org/10.1080/03007766.2011.616298

Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The Psychology of Music* (1st ed., pp. 149–180). New York: Academic Press.

Honing, H. (2012). Structure and Interpretation of Rhythm in Music. In D. Deutsch (Ed.), *The Psychology of Music* (3rd ed., pp. 367–404). Academic Press, Elsevier.

Huron, D. (2006). *Sweet anticipation: music and the psychology of expectation*. Cambridge, Massachusetts / London, England: The MIT Press.

Huron, D., & Ommen, A. (2006). An Empirical Study of Syncopation in American Popular Music, 1890?1939. *Music Theory Spectrum*, *28*(2), 211–231. https://doi.org/10.1525/mts.2006.28.2.211

Jones, M. R. (2008). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The oxford handbook of music psychology* (pp. 81–92). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199298457.013.0008

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319. https://doi.org/10.1111/j.0956-7976.2002.00458.x

Komar, A. . J. (1971). *Theory of Suspension*. Princeton, NJ.

Lerdahl, F., & Jackendoff, R. (1981). On the Theory of Grouping and Meter. *The Musical Quarterly*, *67*(4), 479–506.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: The MIT Press.

Lewin, D. (2007). *Generalized Musical Intervals and Transformations*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195317138.001.0001

London, J. (2012). *Hearing in Time* (2nd ed.). New York: Oxford University Press.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception: An Interdisciplinary Journal*, *1*(4), 424–441. https://doi.org/10.2307/40285271

Louboutin, C., & Meredith, D. (2016). Using general-purpose compression algorithms for music analysis. *Journal of New Music Research*, *45*(1), 1–16. https://doi.org/10.1080/09298215.2015.1133656

McAuley, J. D. (2010). Tempo and Rhythm. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 165–199). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-6114-3

Meredith, D. (2014). Compression-based geometric pattern discovery in music. *4th International Workshop on Cognitive Information Processing - Proceedings of CIP*

2014. https://doi.org/10.1109/CIP.2014.6844503

Paiement, J.-F., Grandvalet, Y., Bengio, S., & Eck, D. (2007). A generative model for rhythms. *NIPS'2007 Music Brain Cognition Workshop*, 1–8. https://doi.org/10.1145/1390156.1390249

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology. Human Perception and Performance, 16*(4), 728–41.

Parncutt, R. (1987). The perception of pulse in musical rhythm. In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music* (pp. 127–138). Stockholm: Royal Swedish Academy of Music.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11*(4), 409–464.

Post, O., & Toussaint, G. (2011). The Edit Distance as a Measure of Perceived Rhythmic Similarity. *Empirical Musicology Review, 6*(3), 164–179.

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics, 30*(6), 565–572. https://doi.org/10.3758/BF03202011

Repp, B. H. (2006). Rate Limits of Sensorimotor Synchronization. *Advances in Cognitive Psychology, 2*(2), 163–181. https://doi.org/10.2478/v10053-008-0053-9

Schachter, C. (1980). Rhythm and linear analysis: durational reduction. In F. Salzer (Ed.), *The Music Forum 4*. New York.

Sioros, G., & Guedes, C. (2014). Syncopation as Transformation. In M. Aramaki, O. Derrien, R. Kronland-Martinet, & S. Ystad (Eds.), *Sound, Music, and Motion* (Vol. 8905, pp. 635–658). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-12976-1

Temperley, D. (1999). Syncopation in rock: a perceptual perspective. *Popular Music, 18*(1), 19–40. https://doi.org/10.1017/S0261143000008710

Temperley, D. (2009). A Unified Probabilistic Model for Polyphonic Music Analysis. *Journal of New Music Research, 38*(1), 3–18. https://doi.org/10.1080/09298210902928495

Temperley, D. (2010). Modeling Common-Practice Rhythm. *Music Perception, 27*(5), 355–376.

Toussaint, G. (2002). A Mathematical Analysis of African, Brazilian and Cuban Clave Rhythms. In R. Sarhangi (Ed.), *Bridges: Mathematical Connections in Art, Music, and Science* (pp. 157–168). Towson, Maryland: Bridges Conference.

Toussaint, G. T. (2004). A Comparison of Rhythmic Similarity Measures.

Toussaint, G. T. (2013). *The Geometry of Musical Rhythm: What Makes a "Good" Rhythm Good?* Chapman and Hall/CRC.

Yeston, M. (1976). *The Stratification of Musical Rhythm*. New Haven, CT: Yale University Press.

6   Figures
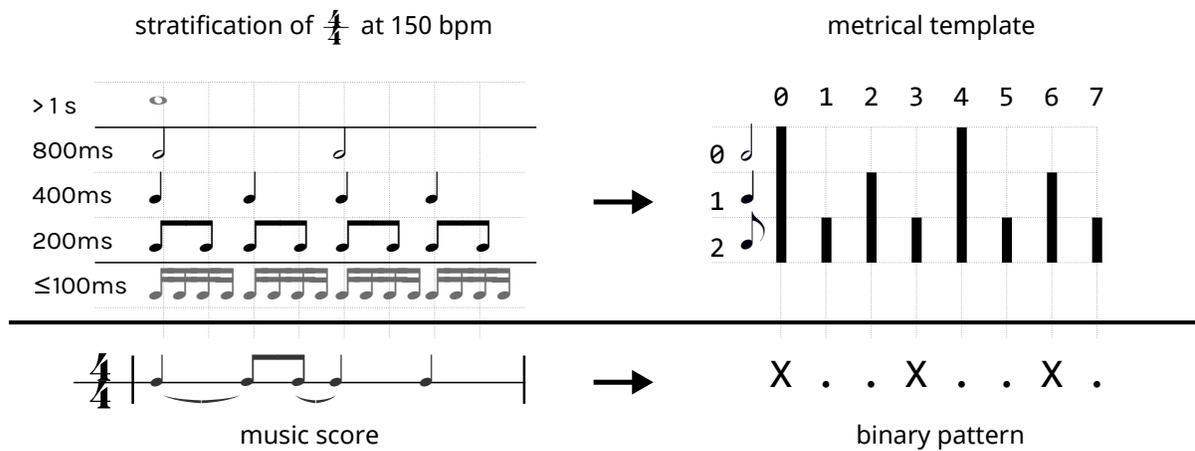
Figure 1. Example of metrical templates and binary representation of rhythms. Left: A rhythmic pattern in

4/4 at 150bpm. Above the pattern, the meter is stratified to its metrical subdivisions. Right: The

binary representation of the pattern (X = event, . = silent pulse). Above it, the metrical template is

shown according to the stratification of the meter on the left. Metrical levels slower than 1s or faster

than 100ms are ignored as described in detail by Sioros and Guedes (2014) and London (2012, Chapter

2).
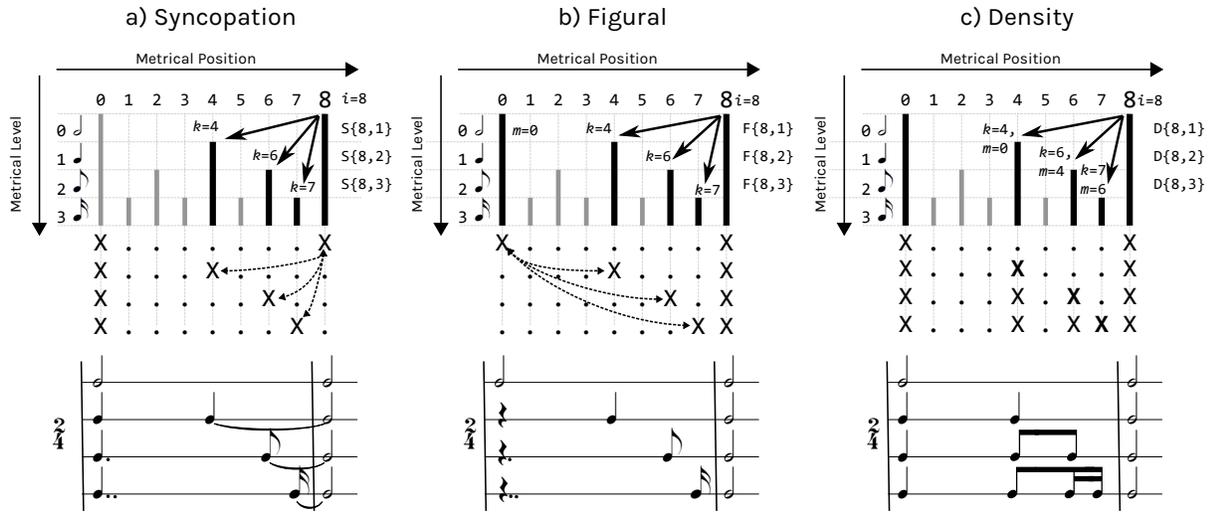
Figure 2. Definition of Syncopation, Figural, and Density transformations. Each transformation corresponds to a vector (solid arrows) consisting of the metrical position on which the transformation is applied (pulse 8) and the metrical level difference to preceding weaker pulse. The syncopation and figural vectors point to the pulse an event is shifted to, while the density vectors point to the pulse on which a new event is inserted. The dotted arrows represent the shift of an event.
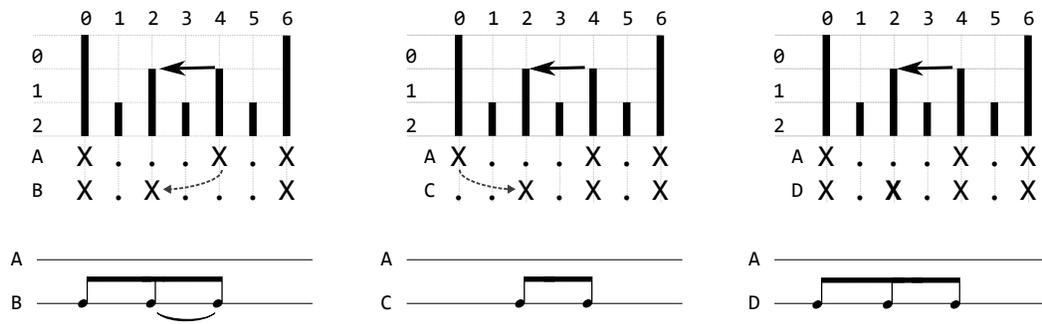
Figure 3. Example of vectors with 0 metrical level difference.

characerize and reverse a transformation on pulse *i*
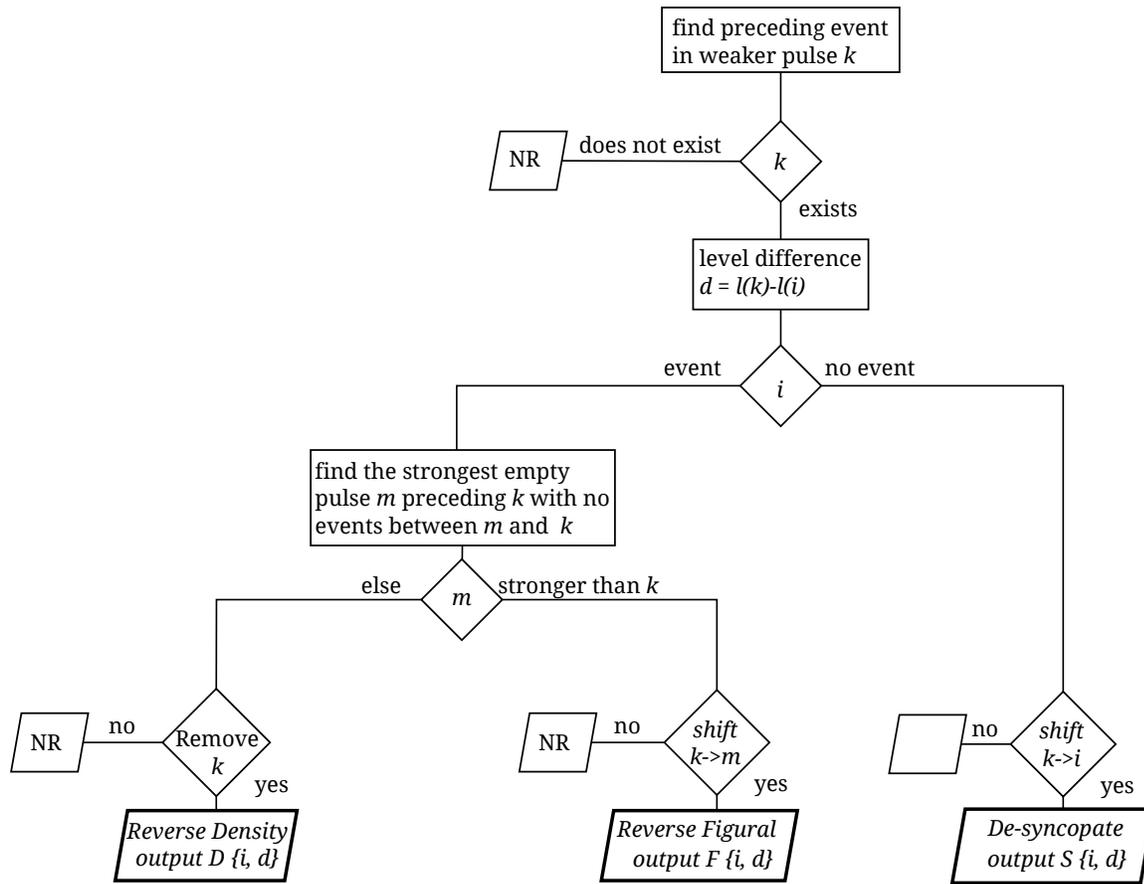
Figure 4. Flow diagram of the process of characterizing and reversing a single transformation previously

applied on pulse *i*. The processes and decisions made are subject to the reversibility constraints

described in the section 2.3. The NR outputs indicate that no transformation is reversed because

either no transformation was found (step 1) or a transformation on another pulse must be reversed

first (step 2).
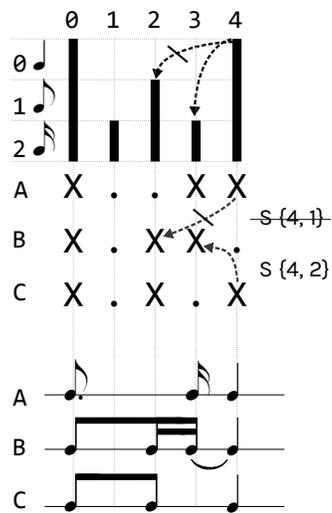
Figure 5: Example of the first reversibility constraint. The strikethrough transformation S {4,1} is forbidden since it does not respect the initial order of events. Pattern B is generated through the S {4,2} transformation.
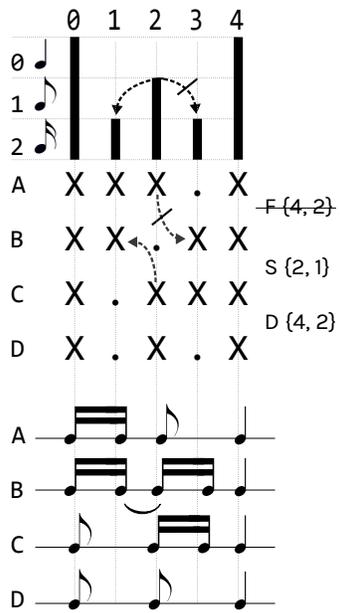
Figure 6. The strikethrough transformation _F_ {4, 2} violates the second constraint. It simultaneously

generates the weak-strong rhythmic figure between pulses 3 and 4 and the syncopation _S_ {2, 1}

between pulses 1 and 2.
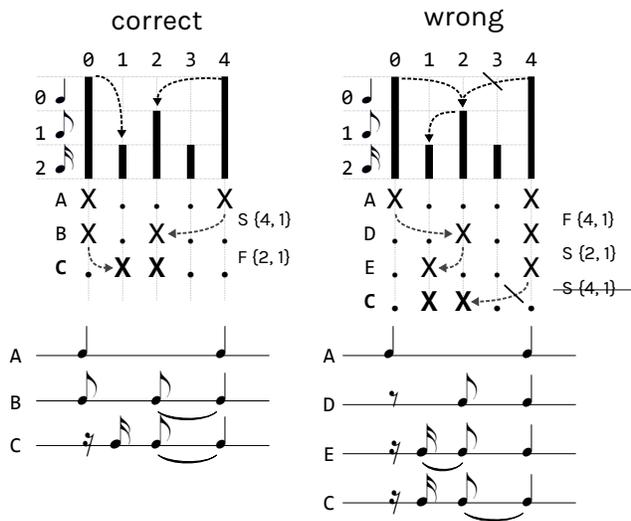
Figure 7. Example of the second reversibility constraint. The bottom pattern is the result of the left-side (correct) transformations rather than the right-side ones (wrong). Introducing the syncopation S {4, 1} on pattern E "masks" the previous syncopation S {2, 1} and therefore is not allowed.
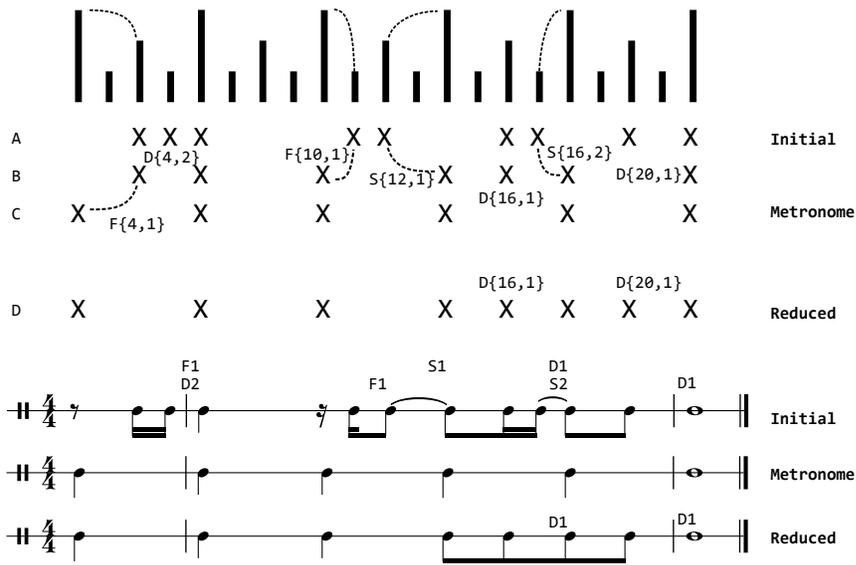
Figure 8: An example of a pattern of five beats, starting with an anacrusis, is characterized by the transformations: *D*{4,2} *F*{10,1} *S*{12,1} *S*{16,2} *D*{20,1} *F*{4,1} *D*{16,1}. Only the events (X) are shown in the binary representation; the dots (.) corresponding to the silent pulses are omitted for clarity. On the Initial and Reduced staffs, a simplified representation of the vectors, which does not contain the pulse index but only the kind of transformation (*F*: figural, *S*: syncopation, *D*: density) and the metrical level difference (e.g. *F*{4 ,1} -> *F*1), is shown above the corresponding metrical positions.
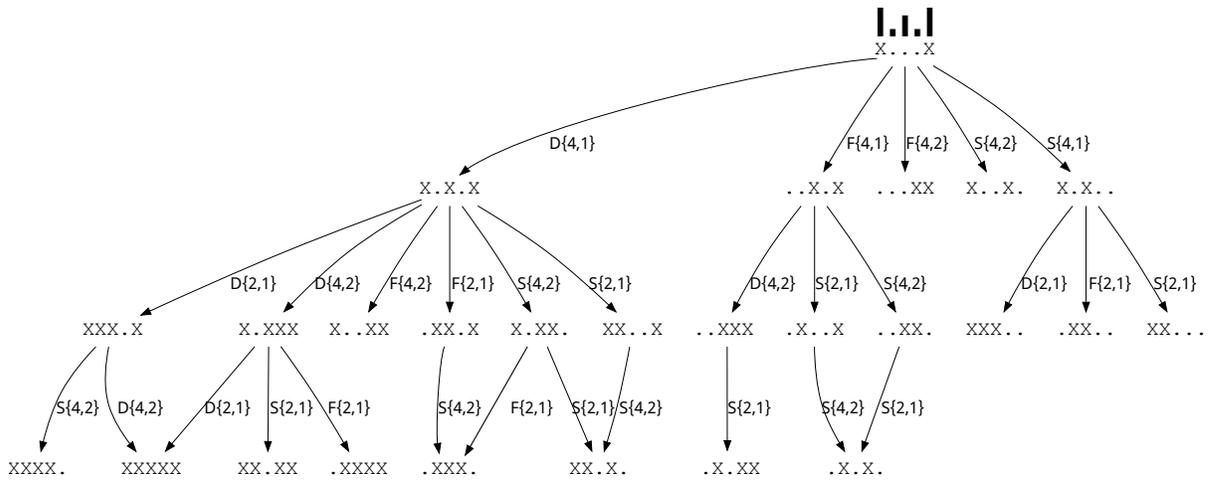
Figure 9: An example of a tree structure consisting of all the patterns generated by a metronome of a single beat duration corresponding to the metrical template shown with bold lines above the root of the tree.

Figure 10. Analysis of the main rhythmic pattern (bass line and voice) of "Don't stop till you get enough" by Michael Jackson. On the staffs, a simplified representation of the vectors is shown where the pulse index is omitted as it is implied by the positioning of the vector at the corresponding metrical position (e.g. *F* {8 ,2} -> *F*2 at the third beat of the bass).