Evaluating melodies by the complexity of polyrhythm

Andranik Tangian

Evaluating melodies by the complexity of polyrhythm

Andranik Tangian
FernUniversität Hagen, D-58084 Hagen
Andranik.Tangian@FernUni-Hagen.De

Abstract

A melody is considered as polyphonic, that is, composed by several latent voices. The rhythm structure of the latent voices constitutes the polyrhythm, whose complexity is estimated. It is supposed that 'good' melodies have some 'optimal' complexity which is not too high and not to low. The model is intended for computer composition and music education.

Keywords: Melody, rhythm, polyphony, polyrhythm, complexity.

1 Introduction

Inventing a good melody is generally considered to be the most creative and the least rationalizable task of music composition. If one asks "What is a good melody?" there is very little to answer. A good melody is easier to remember, or at least easier to recognize. It must be simple and natural, and at the same time not primitive and not resembling other melodies.

Composition students are often given hints like: A good melody is long, has at least three types of durations, and contains both large and small intervals. Even such a general hint is not really true. The opening motive of Mozart's *Eine kleine Nachtmusik* is short and contains only large intervals. The choir melody from Beethoven's 9th Symphony has only quarter notes. The remarkable theme of Bach's E minor fugue from Book I of *Das wohltemperierte Klavier* is only two measures long and contains exclusively sixteenth notes (Figure 1 top).

Several publications on modeling melodies by a number of rules date back to the origins of cybernetics but the enthusiasm was soon exhausted; see Zaripov (1971, 1983) for review and original results. The lack of knowledge about what makes melodies good explains why the contemporary formal music prefers to avoid them at all. Since there are no criteria for their evaluation, there is no rule to sort out bad ones, and no algorithm to construct good ones.

My current work in progress attempts to partially fill in this gap with a model for evaluating melodies. The given paper outlines the idea of the approach developed.

While analyzing melodies, one is confronted with rhythm, pitch, and eventually with underlying harmony. Among these three factors, rhythm seems to be most important. Rimski-Korsakov (1844–1908) often repeated: "Rhythm is already music". He did not say anything similar about pitch or harmony, although he was a famous colorist.

However only the apparent rhythm of a melody, that is, the sequence of durations, is insufficient to make any serious judgement about the melody quality. In fact, a melody can have a poor rhythm but still be good (the choir from Beethoven's 9th Symphony, Bach's theme cited).

The *latent polyrhythm structure* of a melody, or simply its *polyrhythm*, seems to be more adequate to the task. To reveal it, the melody is regarded as polyphonic, that is, as constituted by *latent voices*; see Figure 1 bottom. Latent voices arise due to leaps, so that a melody with disjunct voice-leading is perceived as several voices with conjunct voice-leading. Since each latent voice has its own rhythm, we get an interaction of *latent rhythm layers*, constituting the polyrhythm of the polyphonic melody.

What is actually being done is similar to

Figure 2a

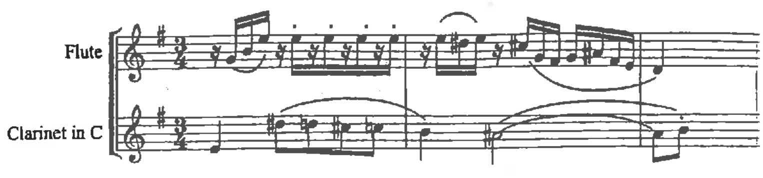


Figure 1: The theme of J.-S. Bach's E minor fugue from Book I of *Das wohltemperierte Klavier* and its latent voices

arranging the rhythm of a melody for a drum set with each drum playing a particular rhythm layer. Since the polyrhythm is recognized by melody leaps, certain pitch information is also taken into consideration. The harmonic factor is beyond the scope of the model.

The second step of analysis (most developed) is aimed at estimating the complexity of rhythm layers and of the total complexity of the polyrhythm as the sum of the layer's complexities. Below the corresponding model is described in more detail.

The polyrhythm of a good melody is assumed to have some 'optimal' complexity, neither too high, nor too low. It reflects the fact that a good melody must be enough simple but not primitive. Finding the lower and upper thresholds for the complexity indices of good melodies is a subject of psychological experiments which are not yet performed.

2 Pitch and polyrhythm

To explain why our melody analysis with no reference to pitch and harmony nevertheless makes sense, let us outline the evolution of musical pitch, also with regard to harmony.

Recall that early music traditions are mainly monophonic and pre-modal, that is, with no definite scale (Alekseev 1976). The pitch served for enhancing the intonation, e.g. in singing declamation, or for more dramatic effects like register changes. Such imprecise *intonation pitch* and *register pitch* are inherent in all early music traditions (Alekseev 1986).

The *modal pitch* with several degrees is typical for more advanced musical cultures. In early Western music, modes were used in a rather free way and could replace one another. Since music was still monophonic and tones were not assumed to sound simultaneously, the accuracy of pitch was not very critical.

Harmony has been 'invented' in Europe in the 15–16th centuries together with innovations in scales and tuning systems facilitating polyphony (Honneger 1996, p. 448). Harmony is often opposed to melody because it adds a vertical aspect to horizontal development. However, it is not that simple. Harmony has horizontal effects because successive and even distant tones create harmonic sensations. Melody against harmony was the point of a famous discussion between Rousseau and Rameau (Holopov 1976, p. 523). Rousseau subordinated harmony to melody, having compared the melody with the drawing in painting and the harmony with the color. Rameau argued that harmony, like a road, indicates the way and thereby gives birth to melody.

Which consequences had the invention of harmony for melodic thinking? Owing to certain harmonic advantages, major and minor scales gained superiority over other modes. Melodies became tonal (as opposed to modal and atonal) with particular gravity tones and a specific function of every scale degree. The pitch classes became fixed and at the same time fine tuned by backing harmonies, e.g. C in chord $D7$ is intonated lower than in chord C.

Harmonic deviations gave a means to control tonal tension thus enabling constructing long melodies. Finally, harmony makes a melody most beautiful and thus becomes its attribute (monophonic music with no harmony is less concerned with the 'beauty' of melody but rather with mood, intonation, contrasts, timbre, rhythm, and certainly words). Therefore, harmonic pitch with underlying harmonic relations is the most comprehensible form of musical pitch.

One can see that the register and intonation pitch have successively evolved into *modal*, *tonal*, and *harmonic pitch*, corresponding to the context where the pitch has been considered. The pitch, having become more accurate, was charged with more functional relations.

The historical perspective proves that *the register pitch is the primary form of pitch and, consequently, must be taken into account in the first turn.* This is exactly what is done while evaluating melodies by polyrhythms. Since the polyrhythm is revealed by leaps in the melody, the discrimination between large and small intervals is also implemented in the model.

Recognizing polyrhythm with respect to register pitch closely relates to perceptual streaming (Bregman and McAdams 1979, Bregamn 1990). Unlike perceptual classification (e.g. pitch recognition), streaming operates with permanently updated separating thresholds. For instance, the progression of tones c_1, c_2, d_1, $d_2, \ldots, c_3$, c_4 is streamed into two voices, low $c_1, d_1, \ldots, c_3$ and high $c_2, d_2, \ldots, c_4$. Note that the low voice remains low voice even at c_3 which is much higher than the first note c_2 of the high voice. Streaming uses only grouping and simplicity principles and does not require any learning. Therefore, it relates to 'naive' perception rather than to knowledge-based 'intelligent' perception.

Figure 2 top displays a monophonic melody, containing exclusively eighth durations. Nevertheless it is rhythmically perceived in a more complex way. The effect of rhythmic variety emerges due to perceptual streaming of tones into latent voices with their own rhythms. The streaming is explicitly shown in R. Brunetti's score (Figure 2 bottom) which emphasizes the rhythm complexity of the polyphonic melody. Tapping the Brunetti's rhythms on three drums resembles the original theme, proving that certain melodic information is retained. Similarly, a melody remains recognizable even if it is hummed with inaccurate intonation.

Thus, evaluating melodies by reducing them to polyrhythms deals with most fundamental melodic features, rhythm (as recognized by Rimski-Korsakov) and primary register pitch (as discovered by Alekseev). On the other hand, it is based on 'naive' perceptual streaming. It makes our incomplete analysis consistent with perception priorities, as if we simulate the reaction of a musically inexperienced listener.

3 Formal analysis of rhythm

Applying the ideas of Kolmogorow (1965) to rhythm, define the complexity of a rhythm to be the amount of data storage required for the algorithm of the rhythm generation (Tangian 1993, 1994, 1998). For instance, a repetitive rhythm requires storing its generative pattern and the number of repeats. If the rhythm of a melody uses several patterns, all of them must be stored, implying an increase in the rhythm complexity. In certain cases one rhythm pattern can be derived from another, which reduces the total complexity.

Generative rhythm patterns are not necessarily linked to measures and bar lines. Therefore, revealing them needs a special techniques which is implemented in the rules described below.

3.1 Accentuation

Rule 1 (Durations) *The only characteristic of a time event is the duration of time interval between its onset and the onset of the next time event. This inter-onset time interval is said to be the duration associated with the event. If the given time event is the last in the sequence, the associated duration is assumed to be not fixed.*

Rule 2 (Accentuation Distinguishability) *In order to distinguish accentuated events*

Figure 2: The *Student's Theme* by A. Tangian and its polyphonic arrangement by R. Brunetti's

in a sequence of time events, at least two types of durations are necessary.

Rule 3 (Accentuated Durations) *The duration associated with an event is said to be strongly accentuated if*

(a) it is longer than its closest neighbors, that is, it follows a shorter duration and the next one is also shorter;

(b) it follows an equal duration and the next one is shorter.

The duration is said to be weakly accentuated if

(c) it follows a shorter duration and precedes an equal duration which is not strongly accentuated, that is, the second next is not shorter.

A time event is said to be accentuated (strongly or weakly) if the duration associated with the event is accentuated (strongly or weakly, respectively).

The idea of the third rule is that a longer duration next to a shorter one is accentuated. The most evident Item (*a*) concerns a situation when a longer duration is between two shorter ones. If a duration is between shorter and equal one then it is usually accentuated (Items *b* and *c*), yet in order to avoid simultaneous accents at two equal successive durations between two shorter ones, we assume no accent at the first duration (Item *c*). No accentuation emerges when durations are successively getting shorter or longer.

In order to provide unambiguous segmentation when several accentuated durations emerge in a short phrase, we distinguish between strong and weak accents, with the priority of strong accents. We suppose that a change from a longer duration to a shorter one is immediately recognized, resulting in an accent. Yet after a change from a shorter duration to a longer duration one can expect some further increase in duration, resulting in a weaker sensation of accent; this is the case of weak accentuation.

Figure 3: Accentuation by timing cues

To illustrate the above rules, consider the sequence of time events shown in Figure 3. The first crotchet marked by symbol ">" is weakly accentuated by virtue of Rule 3c, since it is the duration between a shorter duration and an equal one. The last crotchet which is also marked by ">" is strongly accentuated by virtue of Rule 3b, since it is the duration between an equal and a shorter one. Note that although the second crotchet precedes an eighth, it is not accentuated. Indeed, by virtue of Rule 1 the duration of the event associated with this eighth is crotchet, and no shorter duration is adjacent to it. By virtue of Rule 1 the last note of the sequence can be considered both as accentuated, or not.

To show the accentuation in notation, bar lines are put before accentuated events as shown in Figure 3. In the above example, the segmentation with respect to the accentuation determines the 3/4 time of the given phrase.

3.2 Rhythm segmentation

Accentuation as defined above is not sufficient for rhythm segmentation. Indeed, consider a periodic sequence of time events segmented with respect to the accentuation defined in two different ways as shown in Figures 4a–b. To prove the perceptual ambiguity of segmentation of these events, we have performed the following audio experiment: The given sequence of time events has been recorded and played back in a loop, having been amplified gradually from zero level. A series of audio tests has shown that listeners recognize the two segmentations with almost equal probability.

Thus to recognize a rhythm segmentation we need some other cues in addition to the accentuation. For that purpose we introduce rules of classification and elaboration of rhythm patterns. By a *rhythm pattern* we understand any segment of a given sequence of durations.

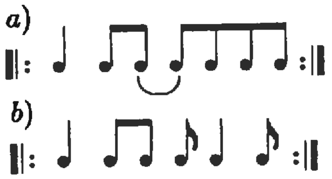


Figure 4: Rhythm segmentation by timing cues

Figure 5: Syllable as an indecomposable unit

However, the most important is the case when rhythm patterns are segmented with respect to the accentuation. Thus we obtain the following definition.

Rule 4 (Phrases and Syllables) *A rhythmic phrase is defined to be a sequence of durations which follows an accentuated duration and ends at an accentuated duration. A rhythmic phrase with the only accentuated duration is said to be a rhythmic syllable* (Katuar, 1926).

Consequently, a rhythmic syllable is a simplest rhythmic phrase. Any rhythmic phrase is formed by adding syllables to each other.

Note that by virtue of Rule 4 a syllable is determined by the durations which precede an accentuated event. The accentuated duration itself is not included into the syllable. The accentuated event just marks the end of the syllable, and the associated accentuated duration may be not fixed (cf. with Rule 1).

We suppose that rhythmic syllables are perceived as indecomposable time units. In order to prove it we have performed the following audio experiment: The rhythmic syllable shown in Figure 5 with two fixed absolute durations 0.1*sec* has been repeatedly reproduced under variable delays divisible by 0.1*sec*, e.g. 0.8, 1.0, 1.2, 0.8, ... *sec*. If the rhythmic syllable was perceived as a composed structure, the common time unit (0.1*sec*) would result in a sensation of constant tempo with changes of rest durations. (Tempo determination with respect to the common time unit was proposed by Messiaen (1944)). However, in our audio experiment listeners have recognized tempo deviations rather than rhythm changes. This proves that the syllable is perceived as an entirety rather than as composed of smaller units and that the end duration is not important for identifying equal syllables.

Such a way of recognizing tempo by time intervals between the entries of similar rhythm patterns meets the principle of correlativity of perception. In fact, in our experiment we have shown that similar rhythm patterns are used as reference indivisible units for tempo tracking. Besides, we have shown that the tempo is a percept of another level than the rhythm.

3.3 Operations on rhythm patterns

In order to classify rhythmic phrases and recognize generative rhythm patterns, we define a reflexive transitive binary relation E, *"is the elaboration of"* on the set of rhythm patterns X. Recall that a binary relation E on X is reflexive if xEx for all $x \in X$ (a rhythm pattern is the elaboration of itself), and transitive if xEy and yEz implies xEz for all $x, y, z \in X$ (a successive elaboration of a rhythm pattern is its elaboration).

Rule 5 (Elaboration) *Rhythm pattern A is the elaboration of rhythm pattern B if A preserves the pulse train of B, that is, if A results from a subdivision of durations of B by inserting additional time events* (Mont-Reynaud & Goldstein, 1985).

Figure 6 illustrates the idea of rhythm elaboration with an example of subdivisions of a crotchet duration (recall that by virtue of Rule 1 the crotchet duration, in order to be determined, should be followed by a next tone onset which is not shown in the figure).

The idea of elaboration can be explained in correlation terms. Represent the rhythm patterns in Figure 6 by 0 and 1 within the accuracy of a sixteenth. Then the top pattern which we denote by T is written down as follows

$$T = \{t_1 \dots t_4\} = \{1000\} ,$$

and the bottom pattern which we denote by B is written down as

$$B = \{b_1 \dots b_4\} = \{1111\} .$$

It is easy to see that pattern B is the elaboration of pattern T if and only if $B \supset T$. This means that B contains all ones of T. Since the number of ones in T is equal to the autocorrelation

$$R_{T,T} = \sum_{i=1}^{4} t_i \cdot t_i$$

(which in the given case is equal to 1), and the number of coinciding ones in B and T equals to the correlation

$$R_{B,T} = \sum_{i=1}^{4} b_i \cdot t_i$$

(which in the given case is equal to 1), we obtain that B is the elaboration of T if and only if

$$R_{B,T} = R_{T,T} .$$

Since the correlation is usually understood as a measure of similarity, the last equation means that the pattern B, being the elaboration of pattern T, is similar to pattern T.

For the patterns of equal duration which are not the elaboration of each other (as in the second line of Figure 6), the correlation is less than autocorrelation. For example, putting

$$L = \{l_1 \dots l_4\} = \{1100\}$$

and

$$M = \{m_1 \dots m_4\} = \{1010\} ,$$

we obtain

$$\begin{aligned} R_{L,M} &= \sum_{i=1}^{4} l_i \cdot m_i = 1 \\ &< 2 \\ &= R_{L,L} = \sum_{i=1}^{4} l_i \cdot l_i \\ &= R_{M,M} = \sum_{i=1}^{4} m_i \cdot m_i . \end{aligned}$$

Now we define the junction of syllables.

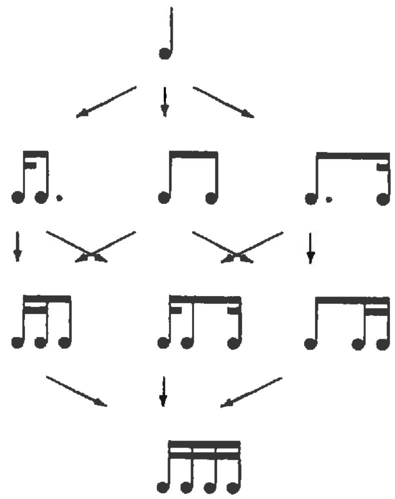


Figure 6: The elaboration of a crotchet rhythm pattern

Rule 6 (Sum and Junction of Syllables) *The sum of two successive rhythm patterns is defined to be the rhythm pattern constituted by the time events of these patterns which are put one after another.*

The junction of two successive rhythmic syllables is defined to be a rhythmic syllable which is the elaboration of their sum.

Note that the sum of two syllables is more than the two syllables in succession. Besides the two syllables themselves, the sum contains the *link*—the accentuated duration after the first syllable. According to the remark following Rule 4, this duration is undefined if the first syllable is considered separately, since instead of the whole duration we consider just an accent. In the sum of syllables, this accent turns to be a duration, linking the two syllables. Therefore, there can be many different sums of the same two syllables, depending on the link duration.

Also note that the sum of two rhythmic syllables is a rhythmic phrase, whereas their junction is a rhythmic syllable. This means that the sum of two syllables can have two accents, at the ends of each syllable, whereas in their junction the internal accent is suppressed by dividing the associated duration into shorter ones

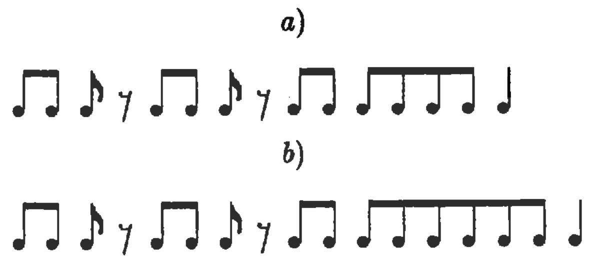


Figure 7: Two different junctions of the same rhythmic syllables

which are no longer accentuated. This implies that the junction has a new rhythm quality, a through tension towards its end.

Figure 7a displays two identical rhythmic syllables and their junction. The total duration of the third syllable is the same as the sum of the two syllables. This results in the symmetry of the whole passage, providing its structure to be $1 + 1 + 2$.

Consider another junction of the two syllables, for instance, obtained by adding two quavers to the third syllable as shown in Figure 7b. This implies that we link the first two syllables not by a crotchet duration but by a half-note duration, that is, we consider the elaboration of another sum of the syllables.

The connection between the three syllables in Figure 7b is less evident than in Figure 7a. Indeed, in Figure 7a one can see not only the two syllables, but already their sum which is elaborated next. In a sense, the elaboration is already "prepared" for easy perception. On the contrary, in Figure 7b the sum of the two syllables is different from the sum which is elaborated. In Figure 7b the intermediate phase between the two syllables and their junction is missed, breaking the successiveness in their perception.

We could provide the effect of such a successiveness in Figure 7b, making the rest between the first two syllables longer, up to a half-note duration. The duration of the second rest is not so important. Even if we change the duration of the second rest in Figure 7a, the third rhythmic syllable is still perceived as the elaboration of the sum of the first two.

From our standpoint, we can explain the simplicity of rhythm construction $1 + 1 + 2 + 4 + \dots$. Such a structure contains a rhythm pattern, the elaboration of the preceding segment, then the elaboration of two preceding segments, and so on. Therefore, the origin of such a structure is quite simple, adding junctions of all preceding segments. This results in perceiving such rhythms with ease; moreover, the perception is "prepared" to recognize the elaboration since the sum is already exhibited.

3.4 Time and rhythm complexity

Thus we have introduced the rules of representation of a given sequence of time events in terms of generative syllables. Constructing such representations, one can reveal origins of a given rhythm with conclusions concerning its time.

Note that rhythm patterns of equal total duration constitute an *ordered directed set* with respect to the elaboration, where every two elements have a common superior—their common *root*. An example of such an order is shown in Figure 6 with a common root pattern at the top and its successive elaborations indicated by arrows.

The patterns of the same total duration which are not elaborations of each other (like in the second line of Figure 6) are of particular interest. If a rhythm contains such patterns then this rhythm has no embedded levels of the pulse train and can be represented as a succession of irreducible units whose pulse train becomes predominant.

The idea of a pulse train generated by indecomposable rhythm patterns can be applied to rhythmic syllables. Since each syllable has the only accent, the accents of syllables determine a pulse train with a certain rhythm. We use this rhythm to determine the time of a given sequence of time events.

Rule 7 (Determination of Time) *If a sequence of time events is representable in terms of elaboration of certain rhythmic syllables (phrases), then the time of the given sequence is determined by the duration ratio of their roots.*

In other words, one has to find a stable preimage (with respect to the elaboration) of generative patterns.

Roughly speaking, the time is defined to be the rhythm of roots of generative syllables.

Besides time determination, rhythmic patterns which are irreducible to each other can be used for estimating the complexity of rhythm. Indeed, their number corresponds to the number of generative patterns required to generate the given sequence.

For example, consider the rhythm in Figure 8 which is constituted by two rhythm patterns of equal duration. One can see that the crotchet duration is the root for the two rhythm groups beamed but no rhythm group is the elaboration of another. This means that the pulse train of crotchets is supported by no pulse train of quavers or some other shorter durations.

Such a rhythm can be considered as less redundant and therefore as more complex. The *complexity of a rhythm* can be identified with the branching index of the graph of the rhythm patterns used, that is, by the maximal number of irreducible to each other rhythm configurations of the same level. For instance, the rhythm in Figure 8 is generated by two patterns of equal duration which are not reducible to each other; consequently, its complexity index is equal to 2.

Such an understanding of rhythm complexity meets the ideas of Messiaen (1944) who has characterized the variety of rhythm by the number of non-commensurable patterns used.

Thus finding irreducible (with respect to elaboration) patterns has two applications: time recognition and estimation of rhythm complexity.

3.5 Example of estimation

Consider the snare drum part from *Bolero* by M. Ravel (Figure 9). Since we use time data only (Rule 1), our method cannot be applied to a rhythm which is based on pitch and dynamic accentuation. Since the chosen rhythm contains two types of durations, by virtue of

Figure 8: A rhythm with complexity 2

Rule 2 it is an appropriate object for our analysis. Let us trace the procedure of structurizing this rhythm step by step.

1. Consider Duration 0. The following one is shorter, consequently, by virtue of Rule 3b it is strongly accentuated. Since it is the first event in the sequence, we recognize the first syllable S as constituted by Duration 0 only. To write down the syllables, we shall use the denotations from Section 3.3, with the only difference that a digit will correspond not to the duration of sixteenth but to the duration of sixteenth triplet. Thus,

   $S = \{100\}$, that is,  .

   Thus up to the current moment our rhythm is represented by the only syllable

   $$S\ .$$

2. Consider Duration 1. It is preceded by a longer duration and succeeded by an equal one. By Rule 3 it is not accentuated. By Rule 4 we don't recognize the end of a syllable at Duration 1.

   Since Durations 2 and 3 are not preceded or succeeded by shorter ones, by virtue of Rule 3 they are not accentuated. Since they are not accentuated, by Rule 4 we don't recognize the end of syllable at these durations.

3. Since Duration 4 is between two shorter durations, by virtue of Rule 3a it is strongly accentuated. By Rule 4 we recognize the end of syllable which we denote

   $S_1 = \{111\ 100\}$, that is, 

   .

   Now we compare syllable S_1 with the earlier recognized, verifying:

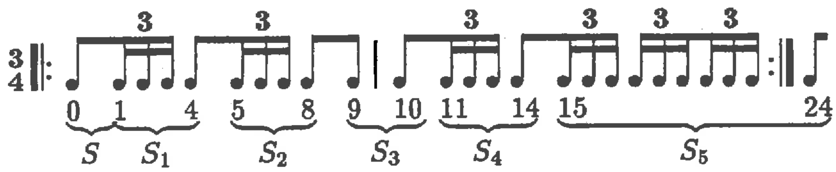


Figure 9: Determination of time by recognizing rhythmic syllables

(a) whether the given syllable is the elaboration of another one;

(b) whether any other syllable is the elaboration of the given one;

(c) whether the given syllable is the junction of other syllables;

(d) whether any other syllable is the junction of the given syllable with another one.

One can see that syllable S_1 is not the elaboration of any other syllable, no syllable is the elaboration of S_1, but S_1 is the junction of two syllables S. Therefore, up to the current moment our rhythm is represented as

$$S \quad S_1 \ ,$$

or

$$S \quad E(2S) \ ,$$

where $E(2S) = E(S+S)$ denotes the elaboration of the sum $S+S$ (that is, the junction of two syllables S).

4. Similarly to Item 2, there is no accentuation at Durations 5–7 and we don't recognize the end of syllable.

5. Similarly to Duration 4 analyzed in Item 3, there is an accentuation at Duration 8, with the only difference that Duration 8 is *weakly* accentuated. By Rule 4 we recognize the end of syllable which we denote

$S_2 = \{111\ 100\}$, that is,  .

Note that syllable S_2 is equal to S_1. Consequently, everything said about syllable S_1 relates also to S_2. Therefore, up to the current moment our rhythm can be represented in the following two ways

$$\begin{array}{lll} S & E(2S) & E(2S) \ ; \\ S & S_1 & S_1 \ . \end{array}$$

6. One can see that Duration 9 is not accentuated, and therefore no syllable ends at Duration 9.

7. By Rule 3b Duration 10 is accentuated, and we recognize syllable

$S_3 = \{100\ 100\}$, that is,  .

Answering the questions (a)–(d) enumerated in Item 3, we recognize that S_1 and S_2 are the elaborations of S_3; besides, S_3 is the sum of two syllables S. Thus we obtain the following equivalent representations of our rhythm:

$$\begin{array}{llll} S & E(2S) & E(2S) & 2S \ ; \\ S & E(S_3) & E(S_3) & S_3 \ . \end{array}$$

8. Since Durations 11–13 are not accentuated, no syllable ends at these durations.

9. Since by Rule 3a Duration 14 is accentuated, we recognize syllable

$S_4 = \{111\ 100\}$, that is,  .

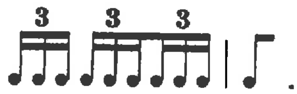
Having answered the questions (a)–(d) enumerated in Item 3, we obtain the following representations of the rhythm:

$$\begin{array}{lllll} S & E(2S) & E(2S) & 2S & E(2S) \ ; \\ S & E(S_3) & E(S_3) & S_3 & E(S_3) \ . \end{array}$$

10. Since Durations 15–23 are not accentuated, no syllable ends at these durations.

11. By virtue of Rule 3a Duration 24 (or Duration 0, taking into account the repeat sign) is accentuated. Consequently, we recognize syllable

$$S_5 = \{111\ 111\ 111\}, \text{ that is,}$$



Having answered the questions (a)–(d) enumerated in Item 3, we obtain that

$$S_5 = E(S_1{+}S_3) = E(S_2{+}S_3) = E(S_3{+}S_3).$$

Hence, we get the following two representations of our rhythm:

$$S \,\|{:}\ E(2S)\ E(2S)\ 2S\ E(2S)\ \ E(4S)\ {:}\| \ ;$$

$$S \,\|{:}\ E(S_3)\ E(S_3)\ S_3\ E(S_3)\ E(2S_3)\ {:}\| \ ,$$

or

$$S \,\|{:}\ S_1\ S_1\ S_3\ S_1\ E(S_1 + S_3)\ {:}\| \ . \qquad (1)$$

If we consider strong accents only, ignoring weak accents, then syllables S_2 and S_3 join into syllable

$$S_{2+3} = \{111\ 100\ 100\ 100\}, \text{ that is,}$$

Since rhythmic syllable S_5 is the junction of syllables S_2 and S_3, we obtain even more simple representation of the rhythm as follows

$$S \ \|{:}\ \ S_1\ \ S_{2+3}\ \ S_1\ \ E(S_{2+3})\ \ {:}\| \ . \qquad (2)$$

With regard to the repetitions of the given rhythm, syllable S can be interpreted as the end of syllable S_5. Finally, we obtain the representation of the given rhythm as generated by phrase S_1, S_{2+3}. Since S_{2+3} is two times longer than S_1, by virtue of Rule 7 we interpret our rhythm as having triple time: 3/4, or 3/8, etc. The choice of denominator (unit of counting) is a question of convention.

Note that there is a risk to interpret the period in (1) as consisting of three equal groups, that is, instead of "correct" segmentation

$$S \,\|{:}\ [S_1\ S_1\ S_3]\ [S_1\ E(S_1 + S_3)]\ {:}\| \ ,$$

on can accept the "wrong" segmentation

$$S \,\|{:}\ [S_1\ S_1]\ [S_3\ S_1]\ [E(S_1 + S_3)]\ {:}\| \ .$$

This corresponds to recognizing the time of the rhythm as 2/4. However, the representation (2) which is obtained by ignoring local accents leaves no doubts in the triple time basis. Thus distinguishing between strong and weak accents is rather useful.

Since S_1 is elaboration of S_3, we obtain

$$S \|{:}\,[E(S_3)\,E(S_3)\,S_3]\,[E(S_3)\,E(E(S_3) + S_3)]\,{:}\|$$

Assume that the complexity $C\{S_3\}$ is two bytes (= two durations), and that calling the algorithms of repeat, of sum, and of elaboration require 1 byte each, we obtain the complexity of the rhythm inside the repeat signs as shown in Table 1.

Table 1: Complexity of the *Bolero* rhythm

Operation	Complexity
Elaboration of S_3 by 2 durations	1+2
Repeat of $E(S_3)$	1
Coding of S_3	2
Repeat of $E(S_3)$	1
Sum of $E(S_3)$ and S_3	1
Elaboration of $E(S_3) + S_3$ by 2 durations	1+2
Total complexity	11

4 Summary

1. It is supposed that a polyphonic melody can be evaluated by estimating the complexity of its polyrhythm.
2. A basic model for estimating the complexity of (poly)rhythm is proposed.
3. Psychological experiments on establishing upper and lower complexity thresholds characterizing good melodies are planned together with improvements the model.
4. The method is intended for several applications, including computer composition.

References

ALEKSEEV E.Y. (1976): *Problems of the mode's formation*, Sovetskii Kompozitor, Moscow (Russian).

ALEKSEEV E.Y. (1986): *Intonating in early traditional music: the pitch aspect*, Sovetskii Kompozitor, Moscow (Russian).

BREGMAN A.S. (1990): *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, Massachusetts: M.I.T. Press.

HOLOPOV Y.N. (1976): "Melody", *Music Encyclopedia, Vol. 3*, Soviet Encyclopedia, Moscow, 511–529 (Russian).

HONNEGER M. (ED.) (1996): *Connaissance de la musique,* Bordas, Paris. (1st Ed. 1966.)

KATUAR G. (1926): *Muzykalnaya Forma. 1. Ritm.* Moscow. (Russian).

KOLMOGOROV A.N. (1965): Three Approaches to Defining the Notion "Quantity of Information". *Problemy Peredatchi Informatsii,* 1(1), 3–11. Reprinted in: Kolmogorov A.N. *Theory of Information and Theory of Algorithms.* Moscow: Nauka, 1987, 213–223. (Russian).

LERDAHL F. & JACKENDOFF R. (1983): *A Generative Theory of Tonal Music.* Cambridge, Massachusetts: M.I.T. Press.

MCADAMS S., & BREGMAN A. (1979): Hearing musical streams. *Computer Music Journal,* 3(4), 26–43, 60, 63. Peprinted in: C. ROADS & J. STRAWN (EDS.) (1985) *Foundations of computer music,* MIT Press, Cambridge Mass., 658–698.

MESSIAEN O. (1944): *Technique de mon langage musical. Vol. 1.* Paris: Leduc.

MONT-REYNAUD B. & GOLDSTEIN M. (1985) On Finding Rhythmic Patterns in Musical Lines. *Proceedings of the International Computer Music Conference'1985.* San Francisco: Computer Music Association, 391–397.

TANGUIANE A.S. (1993): *Artificial Perception and Music Recognition.* Berlin: Springer-Verlag (Lecture Notes in Artificial Intelligence No. 746).

TANGUIANE A.S. (1994): A Principle of Correlativity of Perception and Its Applications to Music Recognition. *Music Perception,* 11 (4), 465–502.

TANGIAN A.S. (1998): An operational definition of rhythm and tempo. *General Psychology,* Vol. ** (Temporal Dynamics and Cognitive Processes), 55–89.

ZARIPOV R.H. (1971): *Cybernetics and music,* Nauka, Moscow (Russian).

ZARIPOV R.H. (1983): *Computer generation of variants for modeling human creativity,* Nauka, Moscow (Russian).