Complexity measures of musical rhythms

Ilya Shmulevich and Dirk-Jan Povel

The notion of complexity has generally been studied in the context of information theory and is closely connected with concepts, such as randomness, information, regularity, and coding (Calude, 1994). Classical information theory as well as notions of randomness, based on Shannon's concept of entropy (Shannon, 1948), relies on a priori knowledge of a probability distribution. In that respect, it does not allow one to speak of a particular object or outcome as being random or complex.

In general, an object's complexity reflects the amount of information embedded in it. The representation of the object's information is achieved via coding. When a human being enters the equation, however, care must be taken in interpreting the notion of complexity, which necessarily becomes subjective. Moreover, depending on the context, only certain types of codes may be perceptually significant and hence coding efficiency or complexity must be considered within such constraints (Chater, 1996). This is well known, for example, in the field of visual perception (Leeuwenberg, 1971). In order to obtain complexity measurements from subjects, Pressing (n.d.) suggests equating complexity with difficulty of learning, which in turn could be expressed by recognition or production.

In this work, we consider the complexity of temporal patterns. Our aim is to construct a measure of complexity that to a high degree corresponds with a human's subjective notion of complexity. Pressing (n.d.) discusses three notions of complexity. The first is termed *hierarchical complexity*, which refers to structure on several levels simultaneously. The receiver is then able to perceive structure on one or more levels, inducing an appropriate complexity judgement. Leyton (1986) proposed a general approach to hierarchical structure in perception. The divisible nature of Western rhythms lends itself to hierarchical subdivision and reveals regularity of time organization on several levels simultaneously (Lerdahl and Jackendoff, 1983). Another type of complexity is referred to as *dynamic complexity*. This notion refers to the degree of stationarity or change with respect to time, in the sensory input. A highly nonstationary stimulus would tend to be perceived as being complex. In music, rhythms tend to be stationary or periodic in that events or groups of events are repeated in time. Finally, the third type of complexity, called *generative complexity*, refers to tendency towards the most economical description (Hochberg and McAlister, 1953; Chater, 1996).

We examine three new measures of rhythm complexity. The first measure is based on the work of Tanguiane (1993), and uses the idea that a rhythmic pattern can be described in terms of (elaborations of) more simple patterns, simultaneously at different levels. The patterns that are not elaborations of any other pattern are called

root patterns. The complexity of the rhythmic pattern is defined by taking the maximum number of root patterns, over all possible structural levels, required to generate the rhythmic pattern in question. A one-to-one correspondence can be established between the set of root patterns and the set of minimal true vectors of a monotone Boolean function. This can be used to determine the maximum complexity under this measure. Figure 1 illustrates the idea of elaboration used by this measure. As can be seen, the quarter note is elaborated into patterns consisting of two notes on the second row, which in turn are further elaborated into patterns of three notes on the third row, which are all finally elaborated into a pattern of four notes on the last row. Thus, all patterns that are linked by a line, either directly or indirectly, to a higher pattern form elaborations of that higher pattern.
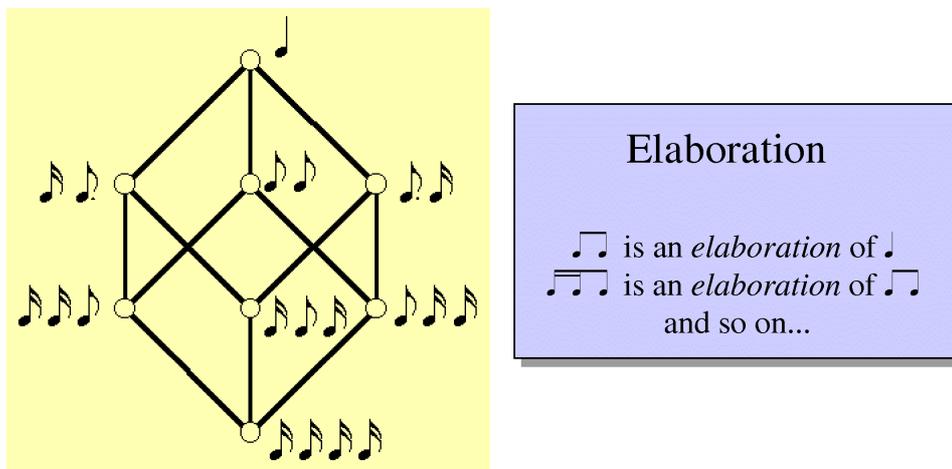

Figure 1. Example of elaborations of a quarter note

The second measure is based on the complexity measure for finite sequences proposed by Lempel and Ziv (1976), which is related to the number of steps in a self-delimiting production process by which such a sequence is presumed to be generated. Essentially, this complexity measure captures the number of "new" substrings found as the sequence evolves from left to right (as is the case in music). As soon as a new substring is found, the complexity increases by 1. An example of this is shown in Figure 2. The measure in essence takes into account repetitions of patterns on all structural levels. It should be pointed out, however, that the LZ complexity in general is not well suited for very short sequences and thus an assumption of cyclical rhythms is very useful. The measure is intended to capture the multi-level redundancy embedded in the rhythmic pattern without regard to any perceptual mechanisms involved in coding it. Thus, the measure does not take into account the possibility that some of the information embedded in the sequence may not be perceptually relevant to a human listener. Therefore, it can be used as a reference point for other measures that do incorporate perceptual constraints in that they should exhibit greater correspondence to subjective judgements of complexity than the LZ-Measure.

111011111001111011111001 • • •

1*110*1111*100*111101*1111001 • • •

➡ **5**

Figure 2. Example of the measure based on Lempel-Ziv complexity. The first step shows the binary coding of the rhythm. In the second step, each star (*) illustrates the point at which a new substring is found, moving from left to right. The complexity is the total number of stars

The third measure, proposed here, is rooted in the theoretical framework of perception of temporal patterns discussed in Povel and Essens (1985). A basic notion of that model is that a listener attempts to establish an internal clock (beat) that segments the rhythm into equal intervals. While there is no physiological explanation of the internal clock nor reason for its existence, structural regularity of auditory stimuli induce in the listener a certain response, be it tapping of a foot or merely recognition of the rhythmic organization. Presumably, this temporal segmentation serves to reduce the coding complexity of the stimulus, which would be consistent with the Gestalt simplicity principle, implying that sensory input is coded in the simplest possible way (Chater, 1996).

The induction of the clock is determined by the distribution of accents in the sequence. As several possible clocks will fit with any given rhythm, it is assumed that the clock which best fits the distribution of accents in the rhythm is the one actually induced. This clock is referred to as the best clock. Furthermore, the ease with which the best clock is induced depends on how well it fits the distribution of accents. After the selection of the best clock, the rhythm is represented by coding the segments produced by this clock.

Discussing the complexity of rhythms, the authors state that a "… given temporal pattern will be … judged complex when either no internal clock is induced or, where it is induced, when the coding of the pattern is relatively complex" (Povel and Essens, 1985). In light of that, the proposed measure of complexity should be a combination of the induction strength of the best clock on the one hand and the efficiency of coding of the rhythm on the other.

The first part of the proposed measure thus pertains to the induction strength of the best clock, which is captured by the C-score (Povel and Essens, 1985). The C-score is computed by taking into account a weighted combination of the number of clock ticks that coincide with unaccented events and with silence:

$$C = W \cdot s + u$$

in which s stands for the number of clock ticks coinciding with silence and u for the number of unaccented events. The lower the score, the higher the induction strength of the clock; hence higher scores correspond to higher complexity.

|	.	.	.	|	Empty (E)
|	.	|	.	|	Equally subdivided (S_2)
|	.	.	|	|	Unequally subdivided (U)
.	|	.	.	|	Starting with silence (N)

Figure 3. Four types of segments

The second part of the proposed measure pertains to the efficiency of the code. In determining coding complexity, we distinguish between four types of possible segments as shown in Figure 3: an empty segment (E), an equally subdivided segment (S_k, where k indicates the number of equal subdivisions), an unequally subdivided segment (U), and finally a segment which begins with silence (N).

To compute the coding complexity, a different weight is associated with each type of segment. Weights $d_1,\ldots,d_4$ correspond respectively to the four types of segments distinguished above. Finally, a weight d_5 is used in order to account for repetitions of segments. Specifically, if a segment is different from the segment following it, d_5 is added to the sum of all weights accumulated so far. The rationale behind this is that two different consecutive segments are likely to increase complexity. Now, the formula for the total coding complexity is:

$$D = \sum_{i=1}^{n} c_i + m \cdot d_5$$

where $c_i \in \{d_1,\ldots,d_4\}$ is the weight of the ith segment, n is the number of segments, and m is the number of consecutive segment pairs containing different segments. Finally, the proposed measure is defined as the weighted combination of the induction strength of the clock and the total coding complexity:

$$P = \lambda \cdot C + (1 - \lambda) \cdot D$$

where C is the induction strength of the best clock and D is the total coding complexity obtained by segmenting the rhythm with that clock.

Table 1. Comparison of Measures: Correlations

	T	LZ	P	Judged
T	1.00	0.12	0.32	0.13
LZ		1.00	0.35	0.26
PS			1.00	0.83
Judged				1.00

All parameters were determined by utilizing the results of an experiment reported in Essens (1995). In Experiment 3 of that work, twenty subjects were asked to make complexity judgements on 24 rhythmic patterns, on a scale of 1 to 5. All parameters were optimized so as to increase the correlation between the average

judged complexity collected in that experiment and the proposed measure. To achieve this, simplex search as well as quasi-Newton search methods were used.

Table 1 shows a comparison of the three measures. The table contains correlations between the measures and the judged complexities. As can be seen, the proposed measure correlated strongest with the judged complexities, since it incorporates perceptual information and is based on an empirically tested model of rhythm perception.

References

Calude, C. (1994). *Information and randomness: an algorithmic perspective*. Berlin: Springer-Verlag.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review, 103,* 566-581.

Coyle, E. J., and Shmulevich, I. (1998). A system for machine Rrecognition of music patterns. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*: Seattle, WA, 3597-3600.

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics*, *57 (4),* 519-532.

Hochberg, J, and McAlister, E. (1953). A quantitative approach to figural "Goodness". *Journal of Experimental Psychology, 46,* 361-364.

Leeuwenberg, E. L. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, *84 (3),* 307-349.

Lempel, A., and Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory.* IT-22 (1), 75-81.

Lerdahl, F, and Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Leyton, M. (1986). Generative systems of analyzers. In: Rosenfeld A. (Ed.) *Human and machine vision II*. Orlando: Academic Press, 149-189.

Povel, D.J, and Essens, P. J. (1985). Perception of temporal patterns. *Music Perception, 2,* 411-441.

Pressing, J. (n.d.). Cognitive complexity and the structure of musical patterns [WWW document]. URL http://psy.uq.edu.au/CogPsych/Noetica/OpenForumIssue8/Pressing.html

Shannon, C.E. (1948). The mathematical theory of communication. *Bell System Technical Journal, 27*.

Shmulevich, I., and Povel, D.J. (1998). Rhythm complexity measures for Mmusic pattern recognition. In *Proceedings of the Workshop on Multimedia Signal Processing 1998*: IEEE Signal Processing Society, 167-172.

Tanguiane, A. S. (1993). *Artificial perception and music recognition.* Berlin: Springer-Verlag.