The *Local Boundary Detection Model (LBDM)* and its Application in the Study of Expressive Timing

Emilios Cambouropoulos

Austrian Research Institute for Artificial Intelligence
Schottengasse 3, A-1010 Vienna, Austria

emilios@ai.univie.ac.at

Abstract

In this paper two main topics are addressed. Firstly, the *Local Boundary Detection Model (LBDM)* is described; this computational model enables the detection of local boundaries in a melodic surface and can be used for musical segmentation. The proposed model is tested against the punctuation rule system developed by Friberg et al. (1998) at KTH, Stockholm. Secondly, the expressive timing deviations found in a number of expert piano performances are examineded in relation to the local boundaries discovered by *LBDM*. As a result of a set of preliminary experiments, it is suggested that the assumption of final-note lengthening of a melodic gesture is not always valid and that, in some cases, the end of a melodic group is marked by lengthening the second-to-last note (or, seeing it from a different viewpoint, by delaying the last note).

1 Introduction

Expressive performance of a musical work relies to a large extent on the underlying musical structure. From traditional music performance theories to contemporary computational models of musical expression a strong link between musical structure and expression is assumed. For instance, in a series of experiments, Widmer (1996) has shown that learning rules of melodic expression by computational means is significantly improved when structural aspects of the music, such as motivic and phrase structure, are taken into account.

It is commonly hypothesised that the ending of a musical group, such as a melodic phrase, is marked by a slowing down of tempo, i.e. relative lengthening of the last notes (Todd 1985). For musical groups at the lowest level, i.e. small melodic gestures of just a few notes, it is commonly assumed that the final note IOI is lengthened and a small micropause inserted (Friberg et al. 1998)

In this paper two main topics will be addressed. Firstly, a computational model that enables the detection of local melodic boundaries will be described. This model is simpler and more general than other models based on a limited set of rules (e.g. Lerdahl and Jackendoff 1983) and can be applied both to quantised score and non-quantised performance data. The proposed model will be tested against the punctuation rule system developed by Friberg et al. (1998). Secondly, the expressive timing deviations found in a number of expert piano performances will be studied in relation to the local boundaries discovered by *LBDM*. As a result of these preliminary experiments it will be suggested that the above assumption of final-note lengthening is not always valid and that in some cases the end of a melodic group may be marked by lengthening the second-to-last note or, seeing it from a different viewpoint, delaying the last note.

2 Finding Local Boundaries in a Melodic Surface

2.1 The *Local Boundary Detection Model* (Refined version)

The *Local Boundary Detection Model (LBDM)* calculates boundary strength values for each interval of a melodic surface according to the strength of local discontinuities; peaks in the resulting sequence of boundary strengths are taken to be potential local boundaries.

Other models (Tenney and Polanski, 1980; Lerdahl and Jackendoff, 1983) that determine melodic boundaries are based on formalisations of the Gestalt principles of similarity and proximity that amount to finding larger pitch/time/dynamic intervals in between smaller ones. Such models, however, detect no boundary, for instance, in the following rhythmic sequence: ♩ ♩ ♩ ♪ ♩ ♩ ♩ even though a listener easily hears a possible point of segmentation.

It has been suggested that a more general approach should account for *any* change in interval magnitudes. An early version of the *LBDM* model (Cambouropoulos 1996, 1997) has shown the potential of this approach (an empirical study in support of this model is presented in Battel and Fimbianti, 1998). It has been demonstrated that both T&P and L&J's models are special cases of the proposed model. The refined version of *LBDM* presented in this paper is more advanced than the earlier version in that it can be applied on any melodic sequence (not only on quantised score-like representations) and it also takes into account the *degree of change* between time/pitch intervals (this way boundaries can be determined at various hierarchic levels).

The proposed refined version of the *Local Boundary Detection Model (LBDM)* is based on two rules: the *Change rule* and the *Proximity rule*. The *Change rule* is more elementary than any of the Gestalt principles as it can be applied to a minimum of two entities (i.e. two entities can be judged to be different by a certain degree) whereas the *Proximity rule* requires at least three entities (i.e. two entities are closer or more similar than two other entities).

Change Rule (CR): Boundary strengths proportional to the degree of change between two consecutive intervals are introduced on either of the two intervals (if both intervals are identical no boundary is suggested).

Proximity Rule (PR): If two consecutive intervals are different, the boundary introduced on the larger interval is proportionally stronger.

The *Change Rule* can be implemented by a degree-of-change function - see suggestion in the description of the *LBDM* algorithm (Figure 1). The *Proximity Rule* can be implemented simply by multiplying the degree-of-change value with the absolute value of each pitch/time/dynamic interval. This way, not only relatively greater neighbouring intervals get proportionally higher values but also greater intervals get higher values in absolute terms - i.e. if in two cases the degree of change is equal, such as sixteenth/eighth and quarter/half note durations, the boundary value on the (longer) half note will be overall greater than the corresponding eighth note.

In the description of the algorithm in Figure 1 only the *pitch, IOI* and *rest* parametric profiles of a melody are mentioned as these have been used in the experiments reported in this paper. It is possible, however, to construct profiles for dynamic intervals (e.g. velocity differences) or for harmonic intervals (distances between successive chords) and any other parameter relevant for the description of melodies. Such distances can also be asymmetric – for instance the dynamic interval between ***p*** - ***f*** should be greater that between ***f*** - ***p***.

The current implementation of *LBDM* is not expected to find all local boundaries correctly as it does not include harmonic profiles of melodies and also it does not take into account melodic similarity which is paramount for establishing important groups of notes (see Cambouropoulos 1998 for a possible complementary module that finds salient repetitions in a melodic sequence). The proposed model, however, identifies successfully a large number of local boundaries as will be shown in the limited evaluation test below.

The *LBDM* algorithm (refined version)
A melodic sequence is converted into a number of independent parametric interval profiles P_k for the parameters: *pitch* (pitch intervals), *ioi* (interonset intervals) and *rest* (rests - calculated as the interval between current onset with previous offset). Pitch intervals can be measured in semitones, and time intervals (for IOIs and rests) in milliseconds or quantised numerical duration values. Upper thresholds for the maximum allowed intervals should be set, such as the whole-note duration for IOIs and rests and the octave for pitch intervals; intervals that exceed the threshold are truncated to the maximum value. A parametric profile P_k is represented as a sequence of *n* intervals of size x_i: $P_k = [x_1, x_2, \ldots x_n]$ where: $k \in \{pitch, ioi, rest\}, x_i \geq 0$ and $i \in \{1,2,...n\}$ The degree of change *r* between two successive interval values x_i and x_{i+1} is given by: $r_{i,i+1} = \frac{\lvert x_i - x_{i+1} \rvert}{x_i + x_{i+1}}$ iff $x_i + x_{i+1} \neq 0$ and $x_i, x_{i+1} \geq 0$ $r_{i,i+1} = 0$ iff $x_i = x_{i+1} = 0$ (N.B. It may be useful to add a small value to the size of all the intervals, such as 1 semitone to pitch intervals, so as to avoid irregularities introduced by intervals of size 0). The strength of the boundary s_i for interval x_i is affected by both the degree of change to the preceding and following intervals, and is given by the function: $s_i = x_i \cdot (r_{i-1,i} + r_{i,i+1})$ For each parameter *k*, sequence $S_k = [s_1, s_2, \ldots s_n]$ is calculated, and normalised in the range: *[0, 1]*. The overall local boundary strength profile for a given melody is a weighted average of the individual strength sequences S_k (weights used in current experiments: w_{pitch}=0.25, w_{ioi}=0.50, w_{rest}=0.25). Local peaks in this overall strength sequence indicate local boundaries.

Figure 1

2.2 Evaluation of *LBDM*

Friberg et al (1998) developed a set of punctuation rules that mark low-level structural boundaries in a melody. A 'comma' is inserted in such marked points realised in a performance in terms of a micropause combined with a small lengthening of the IOI duration. This rule system was adjusted and tested against the punctuation data provided by an expert performer for a set of 52 melodies (the musician marked manually on the melodic scores his preferred punctuation positions).

As it seemed rather plausible that the punctuation marks of a melody were closely related to local boundary positions, an attempt was made to test the *LBDM* algorithm against the punctuation rule system on the same melodic data. The results obtained are depicted in Table 1. Overall the *LBDM* performance was comparable to the performance of the punctuation rule system. The results were worse for the default settings of the *LBDM* algorithm but the performance of the algorithm improved when groups of *one* note were allowed, the cutoff threshold for the important peaks was adjusted and the weights for the different parametric profiles altered (best values: Found=74% and Extra=49%). It should be noted that the punctuation rule system was developed so as to fit the data of the human analysis so it may be the case that if the two systems are tested against other new melodic data the results may be different. It is still interesting to find that the *Local Boundary Detection Model* performs rather well even for the default settings.

Performer	Punctuation Rule System				*LBDM* (default setting)			
N_{Perf}	N_{Rule}	N_{Same}	Found %	Extra %	N_{Lbdm}	N_{Same}	Found %	Extra %
498	501	345	69	31	567	312	63	45

Table 1 Comparison of the *Punctuation Rule System* and the *Local Boundary Detection Model*. N_{Perf}, N_{Rule} and N_{Lbdm} are the numbers of punctuation points/boundaries indicated by the expert performer, the rule system and *LBDM* respectively; N_{Same} is the number of marked positions that are same for the performer and each of the automated systems; *Found* is the percentage of the performer's marks found by the rule system and *LBDM*; *Extra* is the percentage of marks inserted by the rule system and *LBDM* which are not marked by the performer.

It is clear that the *LBDM* is not a complete model of grouping in itself; extensions of the current model (e.g. harmonic component) and also complementary models for establishing musical groups via melodic similarity are necessary. The grouping suggested by the composer, by means of slurs, breath marks and so on, should also be taken into account as this occasionally comes into conflict with the local boundaries (using slurs in *LBDM* has been avoided as this, in some sense, defeats the point of grouping analysis since slurs indicate one possible grouping structure in advance – i.e. the composer's preference). Integrated sophisticated models for musical segmentation are still in their infancy.

3 Expressive Timing Deviations at Local Boundaries

In computational systems of musical performance it is often hypothesised that the last note of a melodic group is lengthened and a micropause inserted. For instance, this assumption is adopted in the Punctuation Rule System described briefly in the previous section. The boundaries between 'smaller melodic units, typically consisting of 1-7 notes' are marked in the performance 'by means of a micropause combined with a small lengthening of the interonset duration' (Friberg et al 1998, p.272). How general is this assumption? Is it always true or are there special cases where different strategies are employed by a performer in order to 'mark' the end of a melodic gesture?

In the following sections two experiments will be described that examine the relation between local boundaries and corresponding interonset intervals.

3.1 Experiment on 7 piano Sonatas by Mozart

In the first experiment the *LBDM* was applied on the melodic parts of 7 complete Mozart piano sonatas (KV279-283, KV332, KV333) performed by a well-known Viennese pianist on a computer-monitored Boesendorfer SE290 concert piano. This melodic dataset consists of approximately 21000 notes. For each soprano part both the *LBDM* and timing deviation (ratio of nominal ove performed IOIs) curves are computed.

Initially, a boundary strength threshold was selected that split the *LBDM* values into *strong* (roughly 25% of all the notes) and *weak* values. The timing deviation values were divided into three categories: *longer*, *shorter* and *same* (±3% from ratio 1). Then, the percent of the three types of IOI deviations in relation to the two boundary strengths were calculated (see Table 2, first bold section). It is clear that there are almost twice as many lengthened strong boundaries than shortened ones, whereas weak boundary values are lengthened and shortened equally. This is a rough first estimate that shows that notes with stronger *LBDM* boundary values tend to be lengthened in terms of the IOI.

In the last three columns of Table 2 (second bold section) the percentages of lengthened or shortened notes on boundaries (i.e. local peaks in the *LBDM* curve) are depicted. In this case again it is seen that the percentage of notes on boundaries (i.e. notes that end melodic groups) that are lengthened is roughly double than those that are shortened.

LBDM Strengths	weak			strong			boundary peaks		
IOI Deviations	shorter	same	longer	shorter	same	longer	shorter	same	longer
Percent of Notes	39%	24%	37%	28%	24%	48%	29%	24%	47%

Table 2

From the above small experiment it is shown that the IOIs of notes on boundaries tend to be lengthened. However, these results are much weaker than the aforementioned assumption that all notes on boundaries are lengthened. Actually only roughly half of the notes on boundaries are lengthened as the other half remain the same or are shortened. What might be the reasons for such a discrepancy? Three possible reasons are given below:

a. The boundaries detected by *LBDM* are partly wrong.

b. Deviation curves are not tottaly reliable in terms of which notes are lengthened/shortened.

c. The assumption that all the notes at the end of melodic groups are lengthened is wrong.

The limitations of the *LBDM* have already been briefly addressed in the previous section. Regarding the IOI deviation curves, these are quite reliable as they are computed in relation to a local average tempo so that slowing down or speeding up is taken into account for calculating the IOI deviations (ratio of notated over performed IOIs); however, occassional mistakes cannot be ruled out. Finally, in relation to the assumption of final-note lengthening we have seen that there is some support for this assumption but this is not very strong and not very reliable under the specific circumstances of experimentation.

For these reasons a further controlled experiment was designed. The primary aim of the design was to eliminate the main problem of uncertainty in boundary detection; this was achieved by using a short ‘simple’ musical piece for which the *LBDM* would give boundaries that clearly agreed with the structure of the melody from the viewpoint of musical experts (see figure 2). The same piece should be performed by many expert pianists so as to generate a large enough data set.

3.2 Experiment on 22 performances of Chopin’s *Etude Op10, No3*.

In this experiment the *LBDM* was applied on the first 20 bars of Chopin’s *Etude Op10, No3.* This piece has a rather clear low-level grouping structure which is determined by rather long notes in between short ones (approximatelly 17 boundaries). The *LBDM* detects correctly all the important boundaries with only one exception (boundary between bars 15 and 16). This piece was performed by 22 different expert Viennese pianists on a computer-monitored Boesendorfer SE290 concert piano. This dataset consists of approximatelly 2200 notes. For each of the 22 performances an IOI deviation curve was computed (for the melodic part).

In Table 3, the percent of the three types of IOI deviations in relation to two classes of boundary strengths are calculated as explained in the previous subsection (the number of strong boundaries roughly 25% of all notes). It is clear that there are more lengthened strong boundaries than shortened ones, whereas weak boundary values are lengthened and shortened almost equally. This rough estimate shows that notes with stronger *LBDM* boundary values tend to be lengthened in terms of the IOI.

In the last three columns of Table 3 the percentages of lengthened or shortened notes on actual boundaries (i.e. local peaks in the *LBDM* curve) are depicted - for practical reasons relating to the current implementation of the expression/boundary matching algorithm, only the 11 strongest boundaries are used. In this case again it is seen that the percentage of notes on boundaries (i.e. notes that end melodic groups) that are lengthened is roughly double than those that are shortened.

LBDM Strengths	weak			strong			boundary peaks		
IOI Deviations	shorter	same	longer	shorter	same	longer	shorter	same	longer
Percent of Notes	46%	13%	41%	31%	18%	50%	21%	31%	48%

Table 3

In any case, the overall number notes on boundaries that are lengthened is less than 50%. This is not in good agreement with the strong assumption of end-note lengthening. Is it possible that the end of a melodic group is emphasised by some different expressive devise? Is it perhaps the case that the last note is simply delayed, i.e. the second-to-last note lengthened?

The second-to-last note lengthening assumption was tested for the 11 strong boundaries (i.e. 242 instances for the 22 performances). The following very interesting results emerged - see Table 4. Almost all of the notes before the last note of the melodic group are lengthened! The relative lengthening of the second-to-last and last notes was also examined (this eliminates any problems with the deviation curves in terms of computing note lengthening/shortening as to a local tempo average) and it was found that 89% of the cases the second-to-last note was relatively more lengthened than the last note (see Figure 2). ? Ref. WIDMER 2001 new?

LBDM Strengths	boundary peaks		
IOI Deviations	shorter	same	longer
Percent of Notes	3%	5%	92%

Table 4

When examined more closely, these strong boundaries correspond to boundaries inserted after a long note that is preceded by short notes (e.g. ♪♪♪♩). One possible hypothesis regarding the lengthening of the penultimate (short) note is the following: When a note IOI is long in relation to its surrounding notes, further lengthening should be quite significant in absolute terms for it to be perceptible whereas a much smaller lengthening of a preceding short note (delay of long note) is more effective.

These results give some first evidence that the commonly hypothesised principle of last-note IOI lengthening need not be always valid, and that the end of a melodic group may be emphasised by other means such as delaying it (lengthening the previous note).

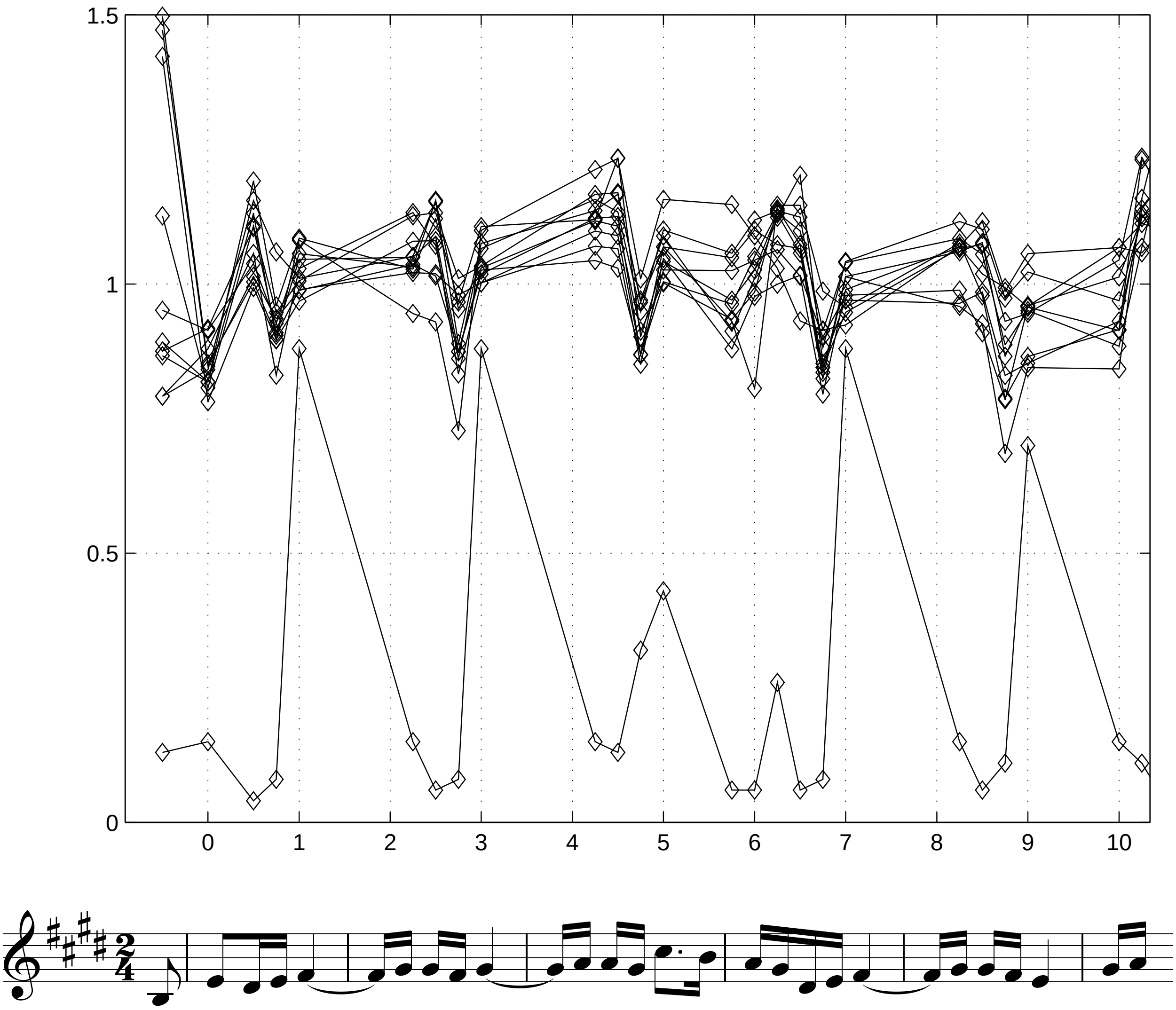

Figure 2 The first few bars of Chopin's *Etude Op10, No3*. The lower curve in the graph depicts the *LBDM* boundary strength profile (peaks in this curve indicate potential local boundaries). The cluster of curves that oscillate around the value 1 on the vertical axis depicts a subset of 10 interonset interval deviation curves (out of the 22). The lower the points of these curve the slower the tempo is. Note that tempo curve points that occur just before the peaks of the *LBDM* curve have a much lower value, i.e. they are played relativel slower than the surrounding notes.

Conclusions

In this paper, firstly, the *Local Boundary Detection Model (LBDM)* was described. This computational model is capable of detecting local boundaries in a melodic surface based on two Gestalt-related principles: identity/change and proximity. The proposed model was tested against the punctuation rule system developed by Friberg et al. (1998) on the same dataset of 52 melodies. It was shown that overall the *LBDM* performance was comparable to the performance of the punctuation rule system (possible advantages of the *LBDM* is its generality and simplicity).

Secondly, the expressive timing deviations found in a number of expert performances of piano works by Mozart and Chopin were studied in relation to the local boundaries discovered by *LBDM*. More specifically the common hypothesis that the final note of a melodic group or phrase is lengthened was examined. It was shown that this hypothesis is not always valid; there are strong indications that often the end of a melodic gesture is marked by lengthening the second-to-last note rather than the last note - or, seeing it from a different viewpoint, by delaying the last note.

Acknowledgements

This research is part of the project Y99-INF, sponsored by the Austrian Federal Ministry of Education, Science, and Culture in the form of a START Research Prize. The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry of Education, Science, and Culture. The 52 melodies were generously provided by Anders Friberg of the KTH Music Acoustics Group in Stockholm, the Mozart piano sonatas by the Bösendorfer company and the Chopin data by Werner Goebl of the Music Group at OEFAI in Vienna. Thanks also to Gerhard Widmer for many interesting comments and suggestions on this work.

References

Battel, G.U. and Fimbianti, R. (1998) Aesthetic Quality of Statistic Average Music Performance in Different Expressive Intentions. In *Proceedings of the XII Colloquium of Musical Informatics*, Gorizia, Italy.

Cambouropoulos, E. (1998) Musical Parallelism and Melodic Segmentation. In *Proceedings of the XII Colloquium of Musical Informatics*, Gorizia, Italy.

Cambouropoulos, E. (1997) Musical Rhythm: Inferring Accentuation and Metrical Structure from Grouping Structure. *In Music, Gestalt and Computing* - *Studies in Systematic and Cognitive Musicology*, M. Leman (ed.), Springer-Verlag, Berlin.

Cambouropoulos, E. (1996) A Formal Theory for the Discovery of Local Boundaries in a Melodic Surface. In Proceedings of the *III Journees d' Informatique Musicale*, Caen, France.

Friberg, A., Bresin, R., Frydén, L. and Sunberg, J. (1998) Musical Punctuation on the Microlevel: Automatic Indentification and Performance of Small Melodic Units. *Journal of New Music research* 27(3):271-292

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

Tenney, J. and Polansky L. (1980) Temporal Gestalt Perception in Music. *Journal of Music Theory*, 24: 205-241.

Todd, N.P. McA. (1985) A Model of Expressive Timing in Tonal Music. *Music Perception* 3:33-58.

Widmer, G. (1996). Learning Expressive Performance: The Structure-Level Approach. *Journal of New Music Research* 25(2): 179-205.