ISMIR2007 Tutorial

T4: Techniques for Implementing the Generative Theory of Tonal Music

by Keiji Hirata, Satoshi Tojo and Masatoshi Hamanaka

Abstract

This tutorial on Techniques for Implementing GTTM will summarize the entire body of work related to computational approaches to the GTTM and report it in a comprehensive way to MIR researchers and computational musicologists. If one wants to realize MIR based on musical semantics, the techniques used for implementing GTTM can provide a powerful tool. Furthermore, the tutorial will put a special focus on perspectives for future deployments, and discussion will be encouraged with experts in the audience.

2007, Sept. 23

Contents

GTTM -- an overview

Satoshi Tojo
JAIST

Generative Linguistic Theory

- Noam Chomsky
- To create an indefinitely large number of sentences.
- A literal translation of some aspect of linguistic theory into music terms, e.g., part of speech, deep structures, transformations, or semantics, has foundered.
- Structural Description is the desideratum.

Well-formedness and Preference rules

- Well-formedness rules specify the possible structural descriptions
- Preference rules designate out of the possible structural descriptions whose that correspond to experienced listeners' hearings of any particular piece.

The Four Analyses

1. Grouping Analysis
2. Metrical Analysis
3. Time-span Tree
4. Prolongational Tree

Grouping Analysis

Grouping Analysis (1)

Well-formedness rules

- GWFR1 Any contiguous sequence of pitch-events, drum beats, or the like can constitute a group, and only contiguous sequences can constitute a group.
- GWFR2 A piece constitutes a group.
- GWFR3 A group may contain smaller groups.
- GWFR4 If a group G1 contains part of a group G2, it must contain all of G2.
- GWFR5 If a group G1 contains a smaller group G2, then G1 must be exhaustively partitioned into smaller groups.

Grouping Analysis (2)

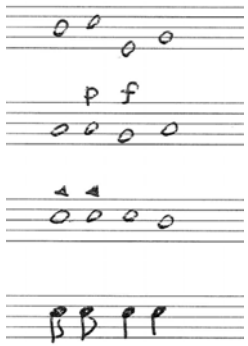
- GPR2 (proximity) Consider a sequence of four notes. The transition of the mid two notes may be heard as a group boundary if
  a. the interval of time from the end of the 2nd note to the beginning of the 3rd is greater than others.

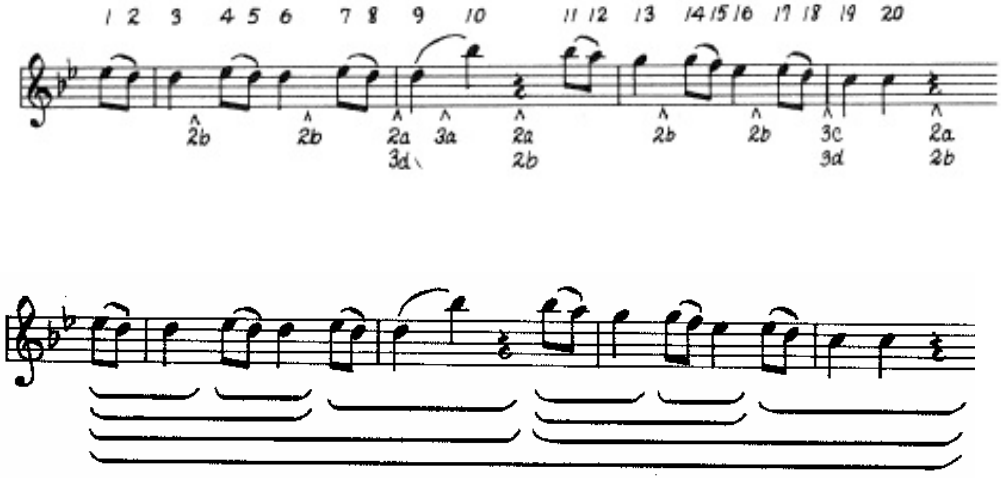
  b. the interval of time between the attack points of the 2nd and the 3rd is greater than others.

Grouping Analysis (3)

- GPR3 (change) Consider a sequence of four notes, n2-n3 may be heard as a group boundary, if
  a. register
  b. dynamics
  c. articulation
  d. length

  changes.





W. A. Mozart: Symphony No.40 in G minor, K.550

Grouping Analysis (4)

- GPR5 (Symmetry) Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length.
- GPR6 (Parallelism) Where two or more segments of music can be construed as parallel, they preferably form parallel parts of groups.

Grouping Analysis (5)

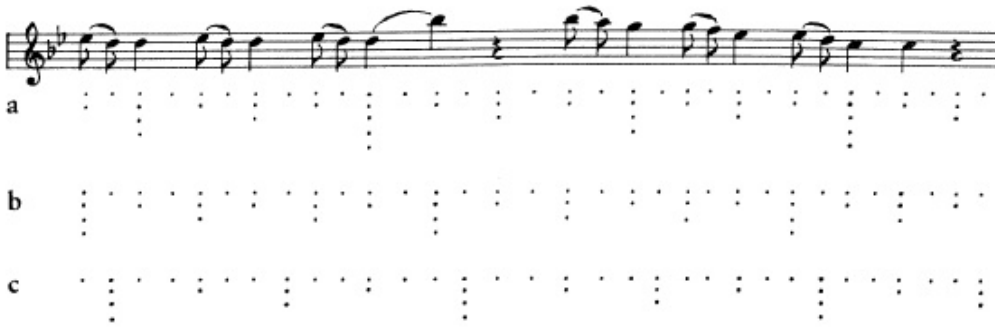
- GPR1 Strongly avoid groups containing a single event.
- GPR4 Where the effects picked out by GPRs2 and 3 are relatively pronounced, a larger-level group boundary may be placed.
- GPR7 Prefer a grouping structure that results in more stable time-span and/or prolongational reductions.

Metrical Analysis

Metrical Analysis (1)

Well –formedness rules (Preliminary version)

- MWFR1 – Every attack point must be associated with a beat at the smallest level of metrical structure.
- MWFR2 – Every beat at a given level must also be a beat at all smaller levels.
- MWFR3 – At each metrical level, strong beats are spaced either two or three beats apart.
- MWFR4 – Each metrical level must consist of equally spaced beats.

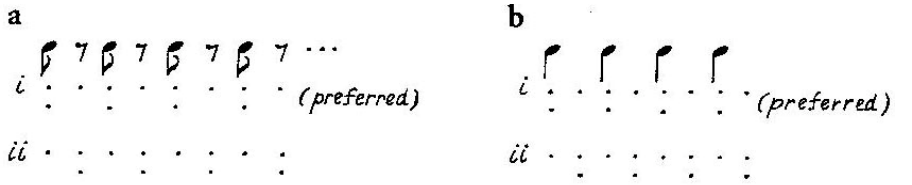


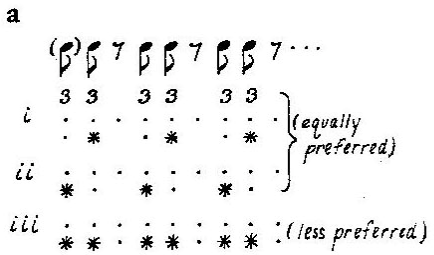
Metrical Analysis (2)

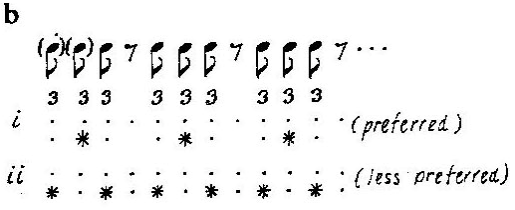
- MPR1 – (parallelism) Where two or more groups or parts of groups can be construed as parallel, they preferably receive parallel metrical structure.
- MPR2 – (strong beat early) Weakly prefer a metrical structure in which the stronger beat in a group appears relatively early in the group.

Metrical Analysis (3)

- MPR3 – (event) The inception of pitch-events, instead of rests or continuations of other pitch-events, are strong beats.
- Violation of MPR3, marked by asterisks, should be minimized.
- Syncopation, to meet the requirement of metrical regularity.

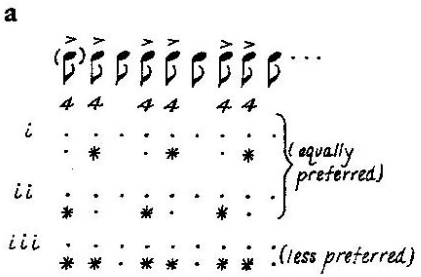


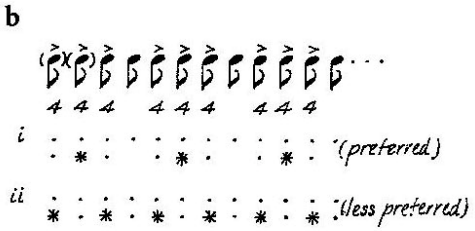




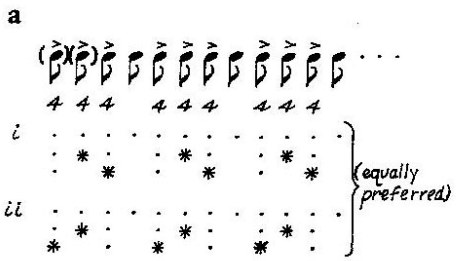
Metrical Analysis (4)

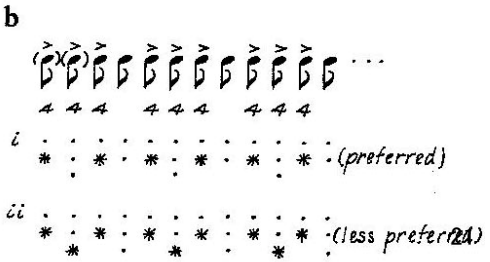
- MPR4 (stress -- accent) Prefer a metrical structure in which stressed beats are strong beats.





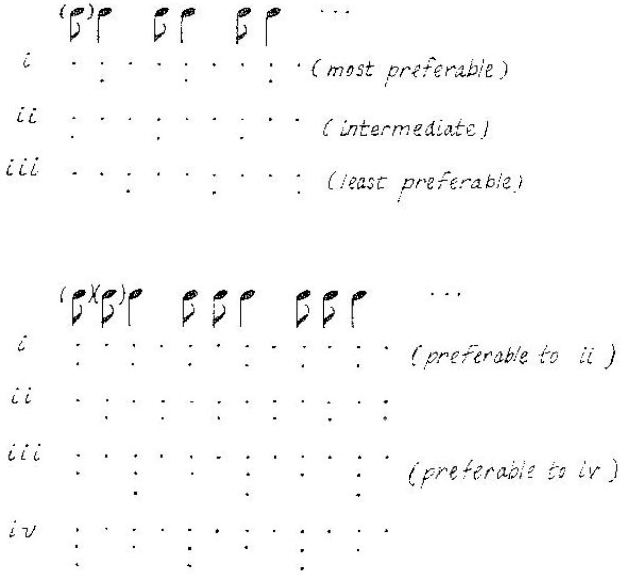
4.21

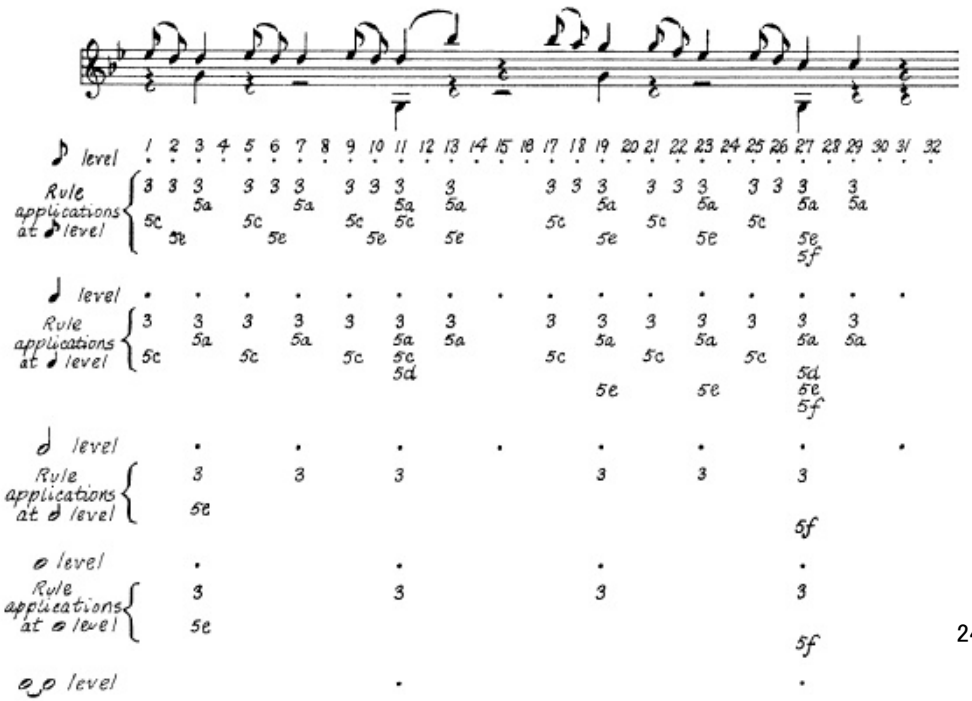




Metrical Analysis (5)

- MPR5 (Length) – Prefer a metrical structure in which relatively strong beats occur at the inception of notes of relatively long duration. (preliminary version)





Metrical Analysis (6)

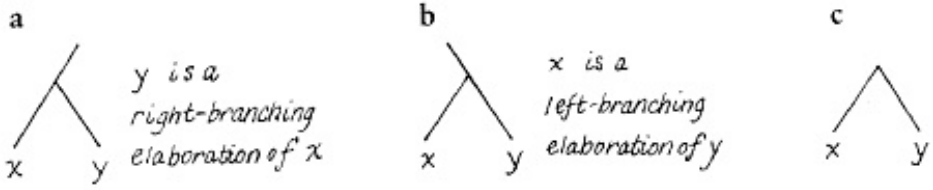
- MPR6 (Bass) Prefer a metrically strong bass.
- MPR7 (Cadence) Strongly prefer a metrical structure in which cadences are metrically stable; that is, strongly avoid violations of local preference rules within cadences.
- MPR8 (Suspension) Strongly prefer a metrical structure in which a suspension is on a stronger beat than its resolution.
- MPR9 (Time-span interaction) Prefer a metrical analysis that minimizes conflict in the time-span reduction.
- MPR10 Prefer duple meter to triple.

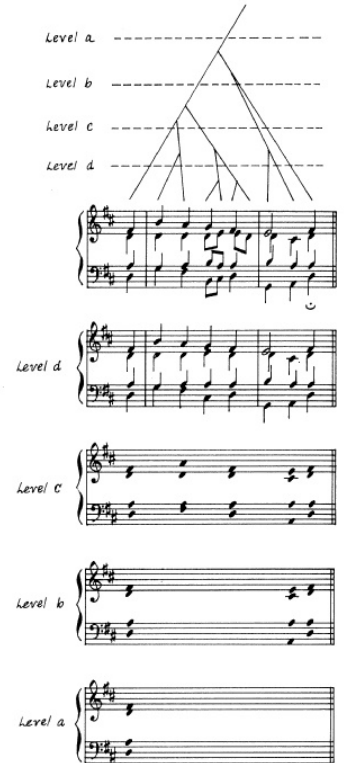
Time-span Reduction

Reduction Hypothesis

- Reduction (step-by-step simplification): The listener attempts to organize all the pitch events of a piece into a single coherent structure, such that they are heard in a hierarchy of relative importance.
- Pitch events are heard in a strict hierarchy
- Structurally less important events are not heard simply as insertions, but in a specified relationship to surrounding more important events.

The Tree Notation





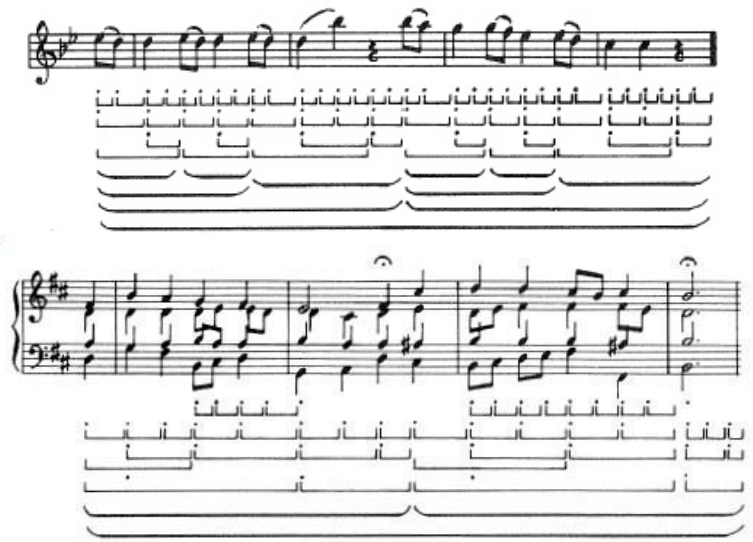
J. S. Bach: `O Haupt voll Blut Und Wunden' from St. Matthew Passion BWV244.

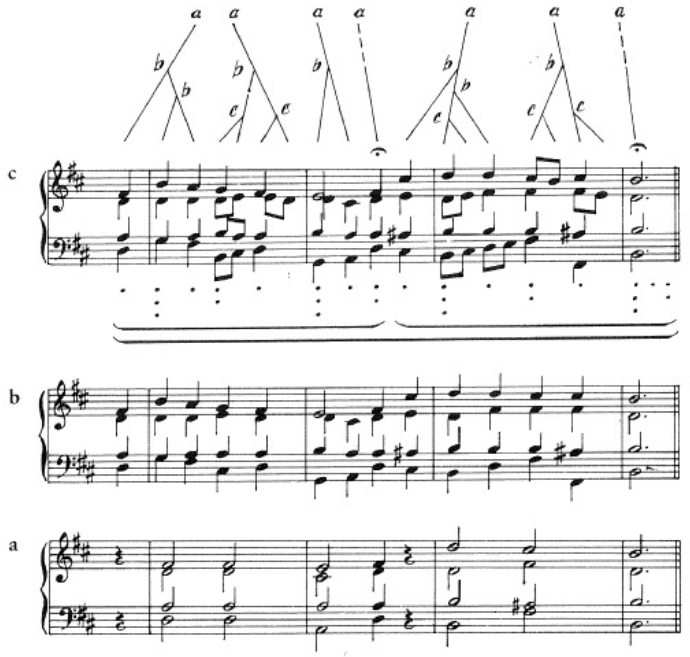
What is Stability?

- Rhythmic stability (grouping + metrical)
- Pitch stability

Two reduction theories
- time-span reduction
- prolongational reduction
…prolongs across group boundaries.
they may not be congruent.

Time-Span Reduction – preview





Time-span Tree and Structural Accents

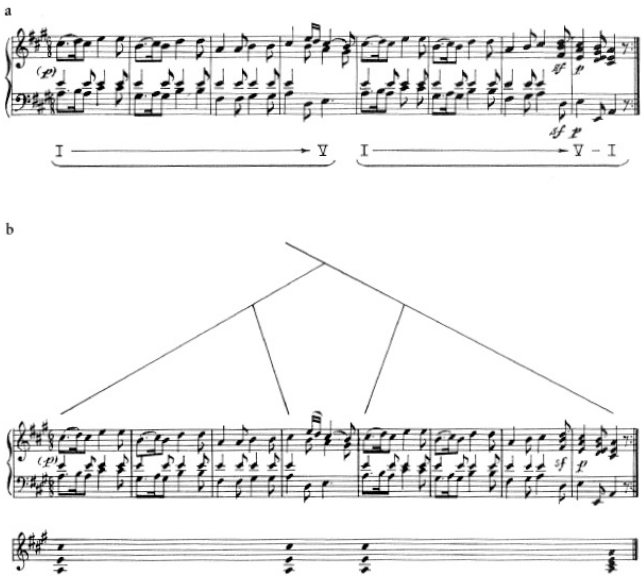
- [b] and [c]: like a ball thrown and caught, the overarching elements of a phrase are its structural beginning and its cadence.
- The [b] and [c] of a phrase must emerge as its structurally most important events in the time-span reduction.
- A phrase can be characterized roughly as the smallest level of grouping in which there is a [b] and a [c].

Cadential Retention

- The half cadence or the deceptive cadence might not emerge simply on grounds of pitch stability as structurally important at the phrase level.
- The full cadence and the deceptive cadence possess two members, joined together as a unit.

↓

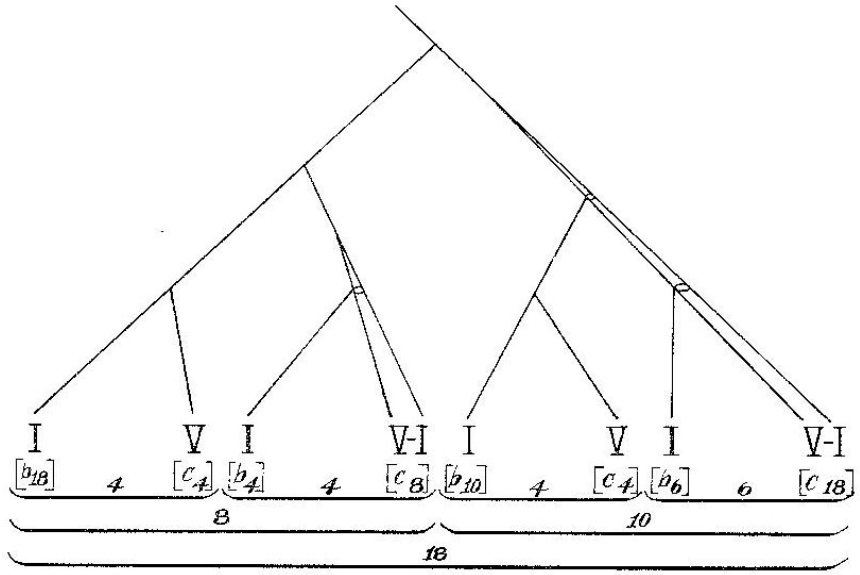
- The two-membered cadence should be counted as one event. → Egg-like shape attaching.



Our hearing: half cadence and full cadence.

The absurd result: I emerges because its pitch structure is more stable and the V in the full cadence disappears.

W. A. Mozart: Piano Sonata No.11 in A major, K.331

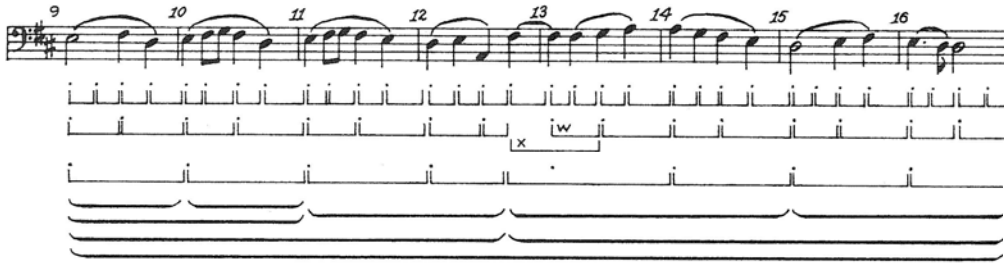


Background Structure

- Intermediate V is not the structural dominant. The structural dominant (the most important V in a passage) is the V at the full cadence that resolves the passage as a whole.

Time-span

- A time-span is an interval of time beginning at a beat of the metrical structure and extending up to, but not including, another beat.
- A group is a time-span.
- If a group boundary intervenes, the beats determine the augmented time-span.



L. van Beethoven: The fourth movement from Symphony No.9 in d minor, op125.

In measure 13, the first beat of the half-note level determines two time-spans, marked w and x. The regular time-span is w and the augmented one is x.

Time-span reduction (1)

- TSRWFR1 – There is a 'head' event.
- TSRWFR2 – The lowest tree is an event.
- TSRWFR3 – Ordinary reduction: heads are chosen among sub-trees' heads. Plus, fusion, transformation, and cadential retention. (cf. next page)
- TSRWFR4 – A two-element cadence: the penult < the final.

Time-span reduction (2)

- Fusion – Arpeggiation of a chord, etc.
- Transformation – The hypothetical chord is inserted.
- Cadential retention – a sequence of events forming a cadence to serve as the head.

Time-span reduction (3)

- TSRPR1 – Prefer strong metrical position for the head.
- TSRPR2 – Prefer relatively consonant or closely related to the local tonic for the head.
- TSRPR3 – Weakly prefer a higher melodic pitch or a lower bass pitch.
- TSRPR4 – Parallelism
- TSRPR5 – Prefer metrical stability: unstable chord in strong metrical position vs. more stable chord in weak metrical position.

Time-span reduction (4)

In choosing the head of a time-span, prefer a choice that results in:

- TSRPR6a – more stable linear connection with events in adjacent time-spans.
- TSRPR6b –more stable harmonic connection with events in adjacent time-spans.

→ More stable choice of prolongational reduction.

Time-span reduction (5)

- Cadenced group:
  - Structural beginning [b]
  - Structural ending (cadence) [c]

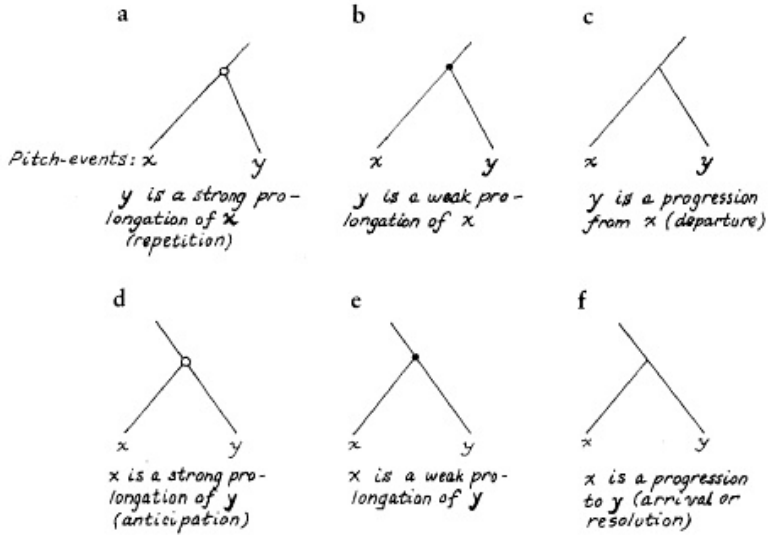
If there is G (group) larger than T (time-span),

- TSRPR7 (preliminary)– Of the possible choices for head of T, prefer an event or pair of events that forms a cadence as [c] of G.
- TSRPR8 – Choose an event relatively close to the beginning of T as [b] of G.
- TSRPR9 – [c] > [b].

Prolongational Reduction

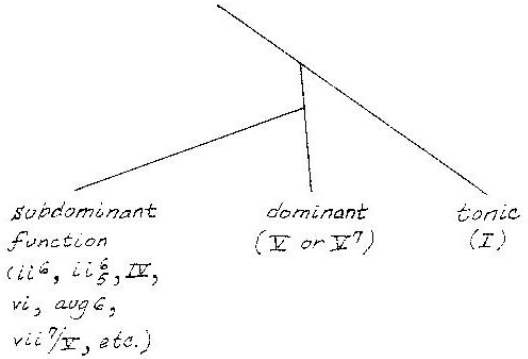
Prolongational reduction -- preview

- Tension – right branching
- Relaxation – left branching
- Progression – no consonant
- Weak prolongation – same root (filled-in circle)
- Strong prolongation – same bass, melody, root (open circle)



Basic form

- Corresponding to Schenkerian Ursatz
- Antecedent – Consequent
- Cadential preparation – events leading up to the cadence – subdominant-dominant-tonic.

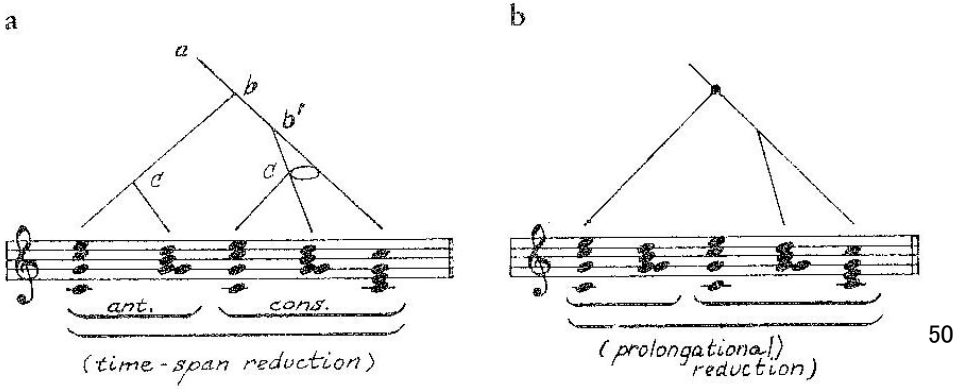


Prolongational reduction (1)

- PRWFR1 – There is a prolongational head
- PRWFR2 – (elaboration)
  - Progression
  - Weak prolongation
  - Strong prolongation
- PRWFR3 – Either head or elaboration
- PRWFR4 – No crossing branches

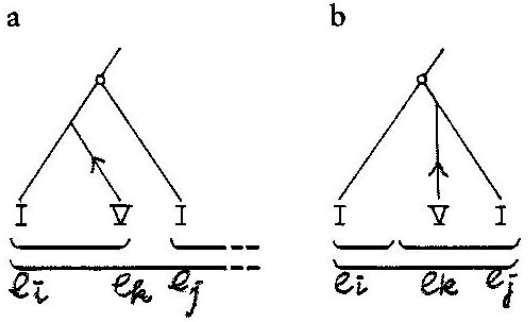
Prolongational reduction (2)

- PRPR1 – The most important event in the next lower level is time-span important.



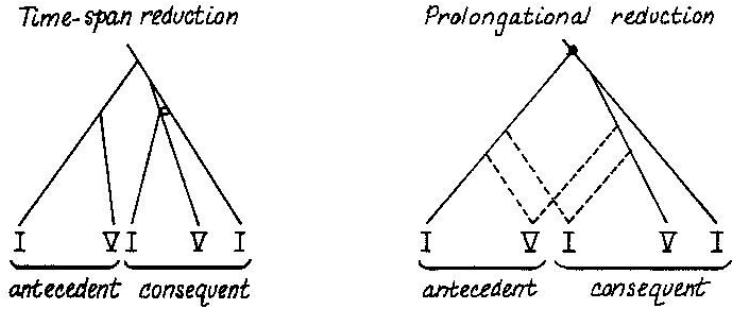
Prolongational Reduction (3)

- PRPR2 – The prolongationally most important event should be an elaboration within the time-span.

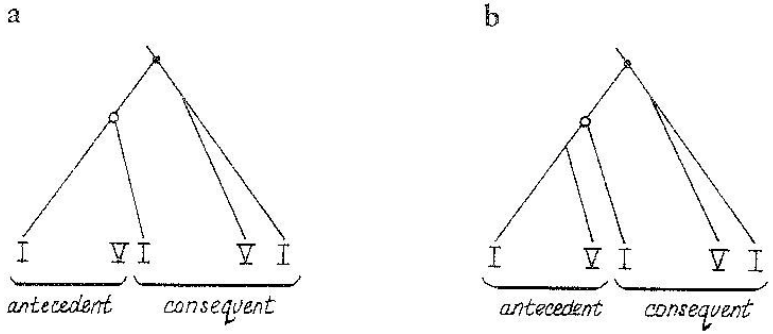


Prolongational Reduction (4)

- PRPR3 – In choosing the most important event in a region, prefer one to form a maximally stable connection with one of the endpoints (beginning/ending).



Among four choices of V or I and toward right or left, Figure (a) is most stable. This may go beyond group boundary.



Then V in the antecedent is attached to the initial I as in Figure (b).

Stability Conditions

1. (Branching condition)
   a. Right strong prolongation > right weak prolongation > right prolongation
   b. Left prolongation > left weak prolongation > left strong prolongation
2. (Pitch-collection condition)

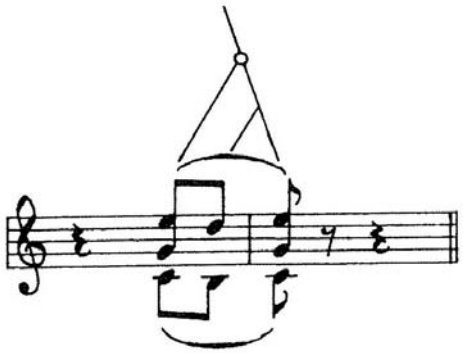
   A connection between two events is more stable if they involve or imply a common diatonic collection.

Stability Conditions – cont'd

3. (Melodic condition)
   a. (Distance) Melodically more stable if the distance is smaller.
   b. (Direction) Ascending is most stable in right-branching; descending is most stable in left-branching.
4. (Harmonic condition)
   a. (Distance) Harmonically more stable if their roots are closer on the circle of fifths.
   b. (Direction) Ascending along the circle of fifths is most stable in right-branching; descending along the circle (subdominant to dominant) is most stable in left-branching.

Prolongational reduction (5)

- PRPR4 –The most important event is an elaboration of the prolongationally more important of the endpoints.



Prolongational reduction (6)

- PRPR5 – parallelism
- PRPR6 – Normative prolongational structure
  a. a prolongational beginning
  b. a prolongational ending
  c. (a right-branching prolongation as elaboration of beginning)
  d. a right-branching progression as elaboration of ending.
  e. a left-branching 'subdominant` progression

Implementing GTTM

Keiji Hirata
NTT

Implementing Music Theory

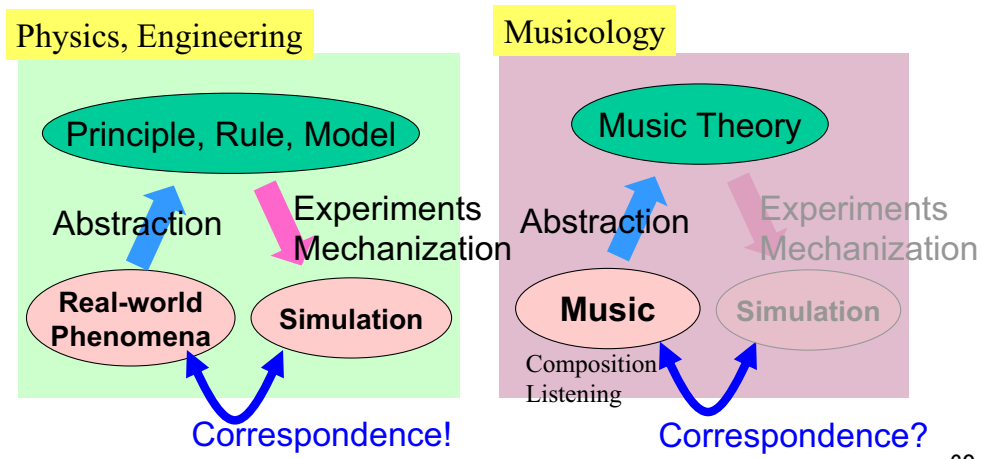
Formalizing Music Analysis
Implementing Music Analyzer

Musical concepts will be formalized:

Melody, Rhythm, Harmony…
Score, Performance, Insight…

- Describing them mathematically
- Constructing a computer program

Motivation and Benefits



Testable, Comparable, Refutable ➡ "Correct" music theory
(GTTM, p.112)

Benefits of Implementing Music Theory

Lerdahl and Jackendoff (1983), p. 55:
Our theory cannot provide a computable procedure for determining musical analyses. However, achieving computability in any meaningful way requires a much better understanding of many difficult musical and psychological issues than exists at present.

Temperley (2001), p. 14:
If the parameters of the rules can be specified, the output of the rule system for a given input can be determined in an objective way, making the theory truly testable.

Why GTTM?

- Rules modularly describe musical insight and concepts
  - ➡ Rule format easily accumulates knowledge
- GTTM is well organized in defining concepts
- Rule descriptions in GTTM are relatively rigid

➡ Relatively easy to translate rule description into computer programs

- Melody, rhythm, and harmony all in a single framework
- Developed based on *reduction* concept
  - ➡ Reduction relation corresponds to "is-a" relation

➡ Wide applicability of theory to working systems

Ambiguities

Ambiguity in Music Analysis

- Music Understanding Itself
- Rule Definitions
  - Concepts are just implicitly presented
  - Concepts needed for mechanization are not presented (nor suggested)
  - Concepts are presented but incomplete

Ambiguity in Music Understanding

Lerdahl and Jackendoff (1983), GTTM, p. 9:
In music, … grammaticality per se plays a far less important role, since almost any passage of music is potentially vastly ambiguous—it is much easier to construe music in a multiplicity of ways. The reason for this is that music is not tied down to specific meanings and functions, as language is. In a sense, music is pure structure, to be "played with" within certain bounds. The interesting musical issues usually concern what is the most coherent or "preferred" way to hear a passage.

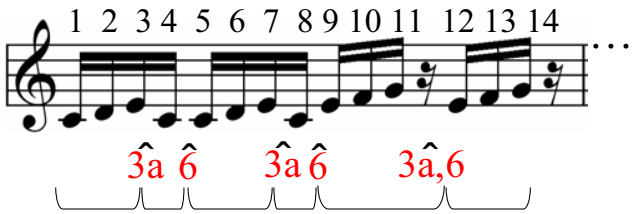
We assume a piece may have more than one correct analysis result

Ambiguity in Rule Definitions (1/3)

- Concepts/Rules are just implicitly presented

Rule Conflict (GTTM, p. 54)
GPR1: Strongly avoid groups containing a single event

GPR3a, GPR6



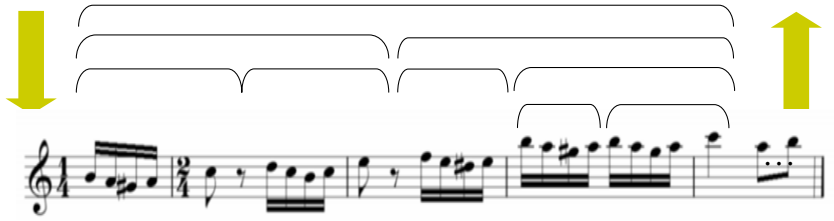
Ambiguity in Rule Definitions (2/3)

- Concepts/Rules needed for mechanization are not presented (nor suggested)

Working Algorithm for Acquiring Hierarchy:

Non-Local

GPR5: ***Top-down*** building of grouping structure

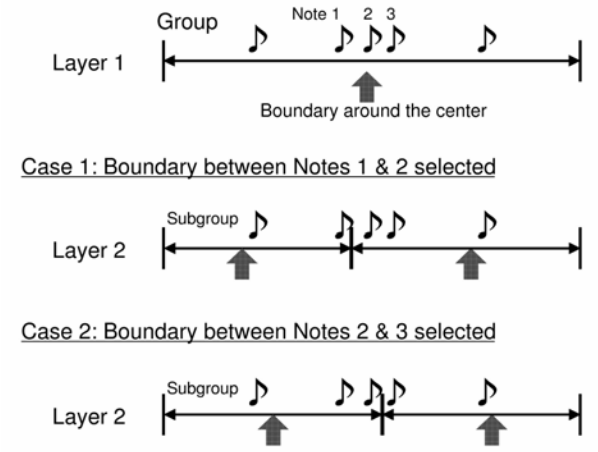


GPR2, GPR3: ***Bottom-up*** building of grouping structure

Local

Non-local Effect in Top-down Segmentation

GPR5:
Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length



Ambiguity in Rule Definitions (3/3)

•Concepts/Rules are presented but incomplete

Vague/obscure definition

GPR4: (Intensification) When effects selected by GPR2/3 are *relatively more pronounced*, a larger-level group boundary *may* be placed

GPR6: When two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups (p.52)

Circular definition

GPR7: Time-Span and Prolongational Stability ⇄ Time-Span Reduction ⇄ MPR9: Time-Span Interaction

Equivocal definition

Stability: Metrical stability (p. 68)
Pitch stability (p. 117)
Stable time-span reduction (GPR7) ⇒ Lerdahl, F. (2001) Tonal Pitch Space

Solutions

- Mainly for Grouping Structure Analysis

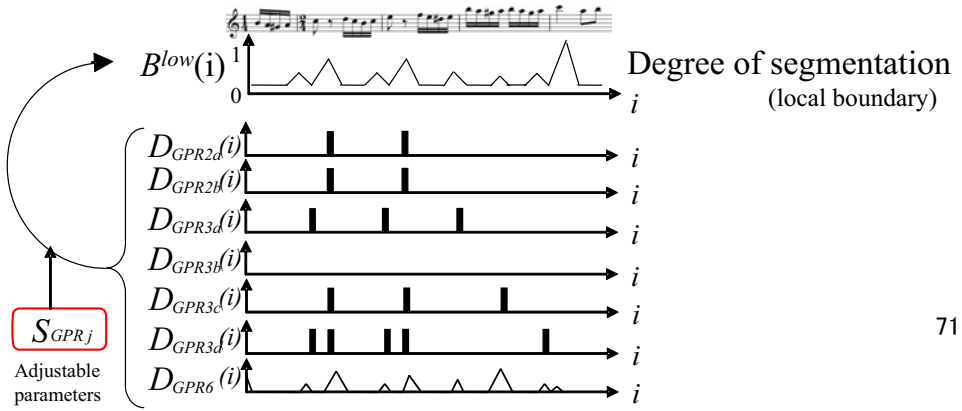
Example: Resolving Rule Conflict

Weight for controlling strength of each rule

- Introducing parameters $D_{\text{rule}}(i)$ and S_{rule}

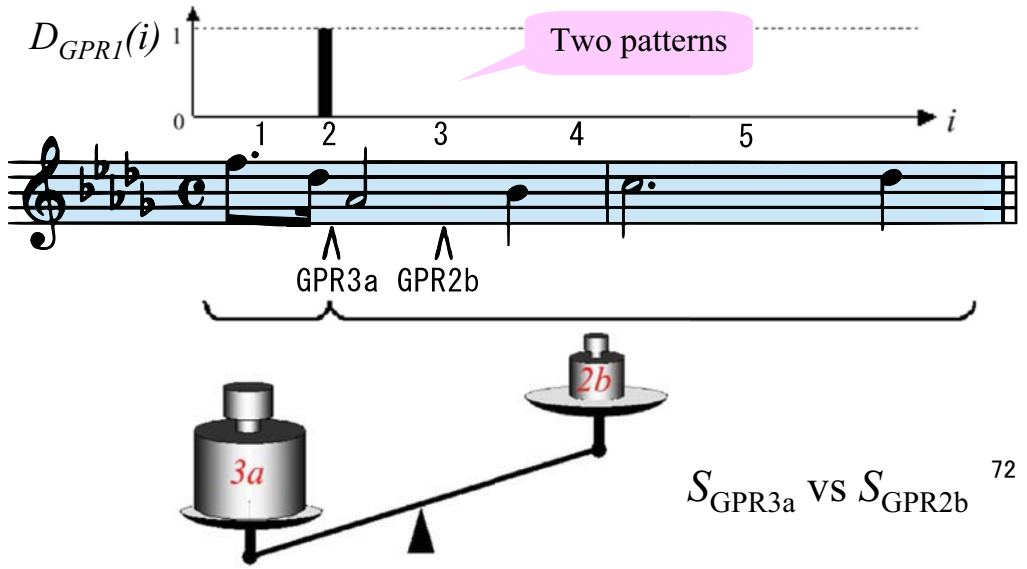
Degree of each rule being held at note i

$$B^{low}(\mathrm{i}) = \sum_{j \in \{2a,\, 2b,\, 3a,\, 3b,\, 3c,\, 3d,\, 6\}} S_{GPRj} \times D_{GPRj}(i)$$



Resolving Rule Conflict: In case of GPR1

- GPR1: Strongly avoid groups containing a single event



S_{GPR3a} vs S_{GPR2b}

Example: Working Algorithm for Acquiring Hierarchical Grouping Structure

Local

GPR1
GPR2: proximity
GPR3: change
GPR4: intensification of GPR2/3

} Bottom-up

Non-local

GPR5: Symmetry
GPR6: Parallelism

} Top-down

→ Can be preprocessed

Circular

GPR7: Time-Span and Prolongational Stability

Integration of Top-down and Bottom-up

Given interval [***s***,***e***] at a level

GPR 2, 3, and 6 → Bottom-up

1. $D^{\text{low}}(i) = \begin{cases} 1 & \text{if } B^{\text{low}}(i) > \mathrm{T}^{\text{low}} \wedge D_{\text{GPR1}}(i) \\ 0 & \text{otherwise} \end{cases}$ (*Local*)

Top-down ($D_{\text{GPR5}}(i)$ varies at each iteration)

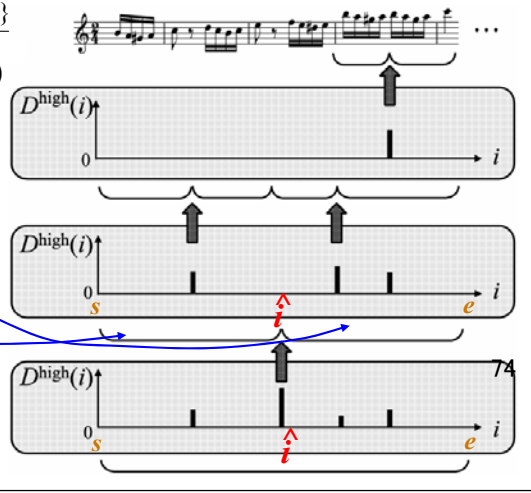
2. $B^{\text{high}}(i) = \dfrac{\sum_{j \in \{2a,\, 2b,\, 3a,\, 3b,3c,\, 3d,\, 4,5,6\}} S_{\text{GPRj}} \times D_{\text{GPRj}}(i)}{\max_{k}(\sum S_{\text{GPRj}} \times D_{\text{GPRj}}(k))}$ (*Non-local*)

3. $\hat{i} = \underset{i}{argmax}\, D^{\text{low}}(i) \times B^{\text{high}}(i)$

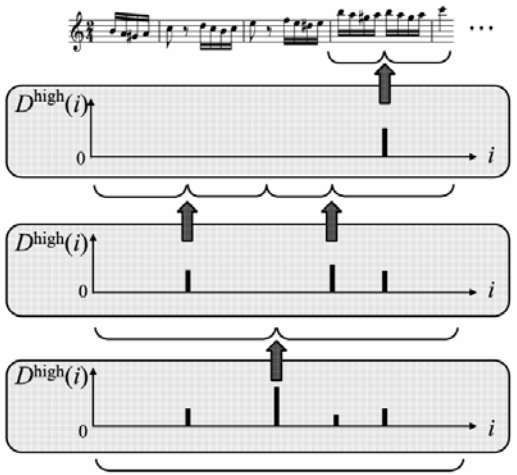
4. Go to next levels [***s***,$\hat{i}$] and [$\hat{i}$,***e***]

Iterate until group cannot be segmented

$D_{\text{GPR5}}(i)$ is recalculated at each iteration
($D_{\text{GPR6}}(i)$ is preprocessed)



Well-formedness Rule

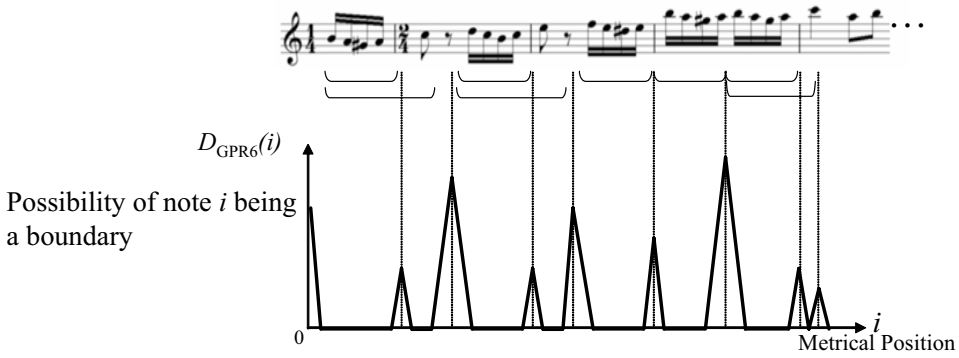


This way of generating a hierarchical grouping structure satisfies GWFR 1~5

Also in Metrical Analysis and Time-Span Reduction, algorithms are designed so that the way they generate the resulting hierarchical structure satisfies the well-formedness rules

Example: Implementing GPR6

GPR6:
When two or more segments of music can be construed as parallel, they preferably form parallel parts of groups

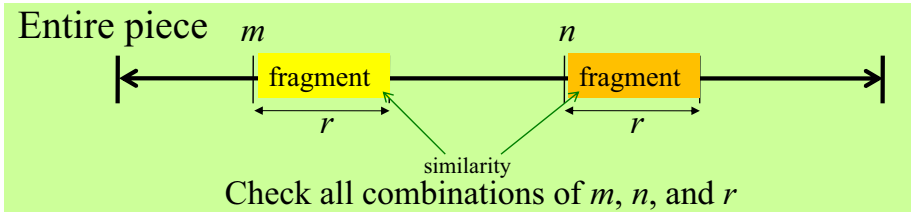


How to Calculate $D_{GPR6}(i)$, whose value rises at the beginning and end of parallel segments?

Parameters Required for Calculating $D_{GPR6}(i)$

Design Outline (How to preprocess non-local structure)

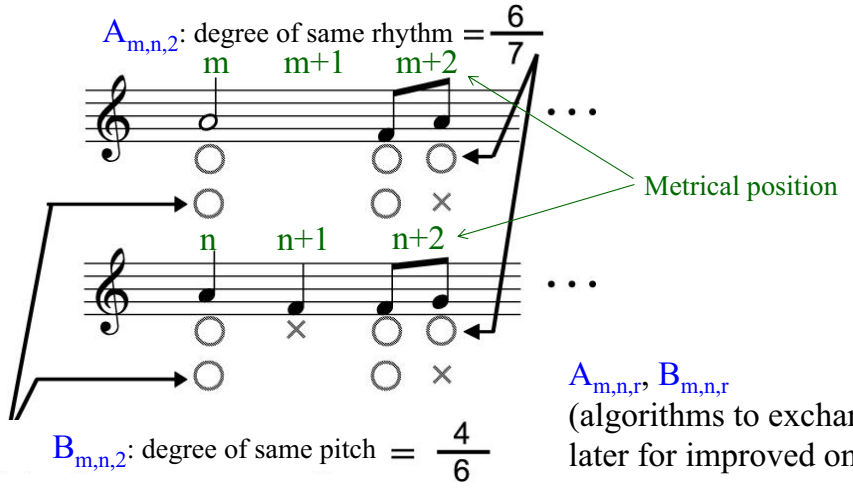
- Develop an algorithm to calculate similarity between two fragments at positions *m* and *n* of length *r*
- For every two fragments of any length within a piece, exhaustively calculate similarities



Fortunately, GPR6 is independent of other GPRs

Calculating Degrees of Same Rhythm and Pitch

Given two fragments at positions *m* and *n* of same length *r*,



$A_{m,n,r}$, $B_{m,n,r}$
(algorithms to exchange later for improved ones)

Giving Priority to Pitch Contour or Note Timing

Given two fragments at positions *m* and *n* of same length *r*,

Note timing

Value $A_{m,n,r}$ is degree of same rhythm
Value $B_{m,n,r}$ is degree of same pitch

Pitch Contour

Total similarity between fragments [m,m+r) and [n,n+r):

$$G(m,n,r) = \{ A_{m,n,r} \times (1 - W_m) + B_{m,n,r} \times W_m \} \times r^{W_l}$$

Implicit and/or lacking

Introducing two adjustable parameters:
- W_m: weighting more to pitch contour or note timing
- W_l: weighting to larger fragment

Additional Predicates and Functions for Intermediate Decisions

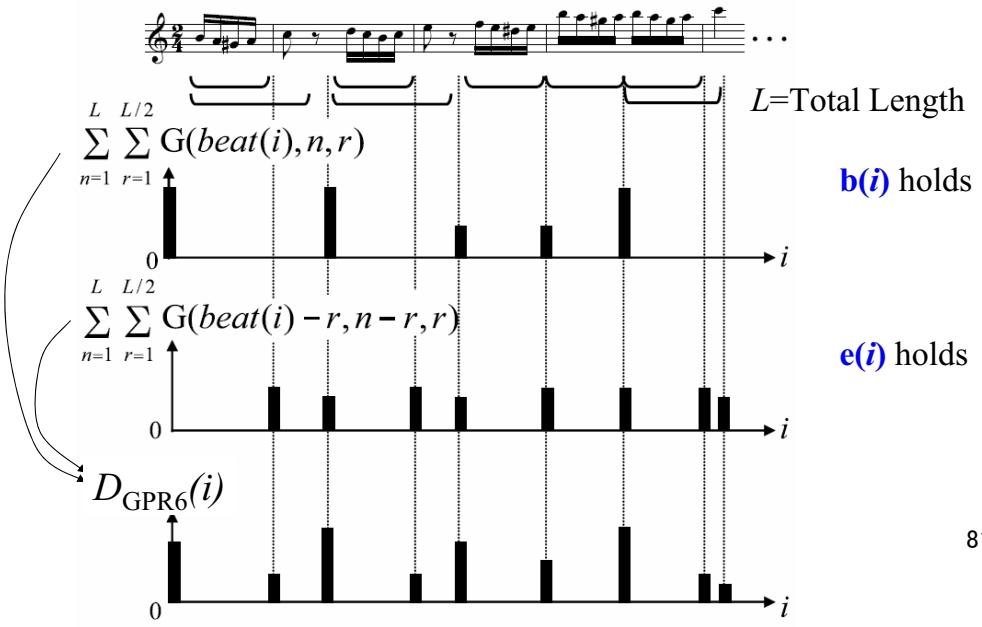
Functions

head(m)**:** returns first note *i* (*i*th note) at interval [m,m+1)
tail(m)**:** returns last note *i* at interval [m,m+1)
beat(*i*)**:** returns m if note *i* occurs in [m,m+1)

Predicates

Note *i* can occur only at beginning of parallel segment:
b(*i*) ≡ (*i*=**head(beat(*i*))** ∧ *i* ≠ **tail(beat(*i*))**)
Note *i* can occur only at end of parallel segment:
e(*i*) ≡ (*i* ≠ **head(beat(*i*))** ∧ *i* = **tail(beat(*i*))**)
Note *i* can occur at either beginning or end of parallel segment:
t(*i*) ≡ (*i* = **head(beat(*i*))** ∧ *i* = **tail(beat(*i*))**)

We Almost Come to $D_{GPR6}(i)$



Final Stage of Calculating $D_{GPR6}(i)$

More precisely,

$$A(i) = \sum_{n=1}^{L}\sum_{r=1}^{L/2} \begin{cases} G(\text{beat(i)},n,r) \times (1\text{-}W_s) & \text{if } \mathbf{b(i)} \text{ holds} \\ G(\text{beat(i)-r},n\text{-}r,r) \times W_s & \text{if } \mathbf{e(i)} \text{ holds} \\ G(\text{beat(i)},n,r) \times (1\text{-}W_s) + G(\text{beat(i)-r},n\text{-}r,r) \times W_s & \text{if } \mathbf{t(i)} \text{ holds} \\ 0 & \text{otherwise} \end{cases}$$

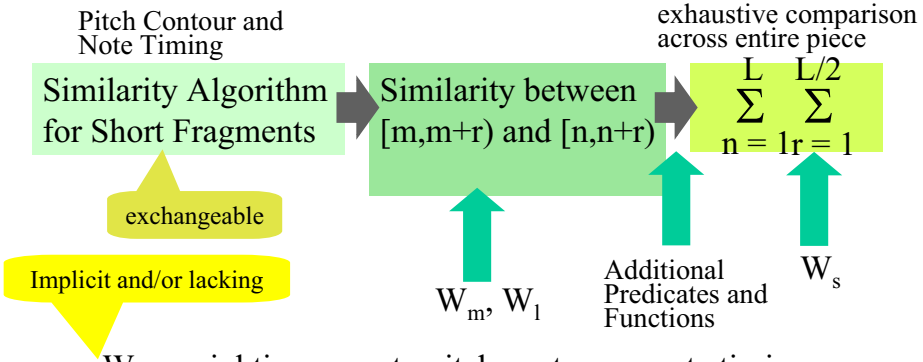
$$D_{GPR6}(i) = A(i) / \max_k(A(k))$$

Implicit and/or lacking

Introducing one more adjustable parameter:
W_s: weighting more to either beginning or end of fragment

Overview: Calculation of $D_{GPR6}(i)$

GPR6:
When two or more segments of music can be construed as parallel, they preferably form parallel parts of groups



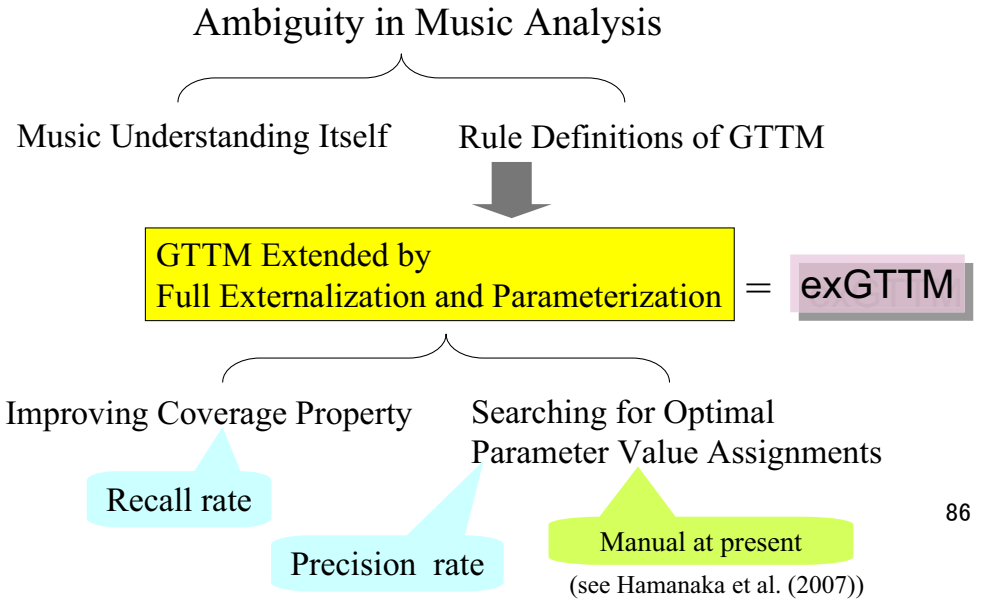
- W_m: weighting more to pitch contour or note timing
- W_l: weighting to larger fragment
- W_s: weighting more to either beginning or end of fragment

Design Approach

Approach to Handling Ambiguity in GTTM

- Supplement as many implicit or lacking parameters as possible
  ➡ Full externalization and parameterization
- Entire problem of music analysis split into
  - identifying domain of all possible results
  - searching for the most preferred result
- Focus on implementing GTTM, but without considering the human process of perceiving and recognizing music

Proposal of exGTTM



Features of exGTTM

- Introducing new parameters for
  - Resolving rule conflict
  - Supplementing implicit/lacking concepts
  - Developing a working algorithm
    (especially an algorithm for acquiring hierarchy)
- Full Externalization and Parameterization
- Restrictions in implementing GTTM
- Generating as many correct results for humans as possible

exGTTM: Restrictions in Implementing GTTM

- Only a monophony is handled
- Harmony is not taken into account
- Only ordinary heads occur in a time-span tree
- No feedback from time-span reduction to grouping and metrical analyses
- Prolongational reduction is not mechanized
- Input representation is a list of notes
  (a monophonic piano roll)

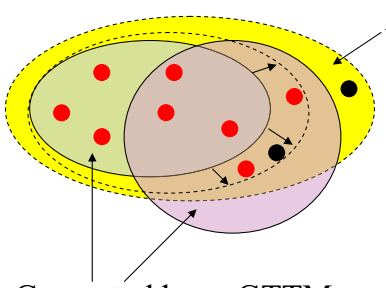
(see Hamanaka et al. (2007))

Generation of Correct Results

Goals

- Coverage of All Human Results
- Precise Controllability

Coverage Property
(recall rate)



● : Correct Time-Span Reduction for Humans

● : Wrong Time-Span Reduction

Related Work

Practical Implementations

- Temperley: Preference Rule Systems and Melisma (2001)
- Stammem & Pennycook (1994)
- Nord (1992)
- Hamanaka & Hirata: Voronoi Diagram (2002)

Temperley's Preference Rule Systems (PRS)

Temperley (2001)

Analyzing Six Aspects:

- *Meter*
- *Melodic phrase*
- Counterpoint
- Pitch spelling
- Harmony
- Key

× Well-formedness Rules and Preference Rules

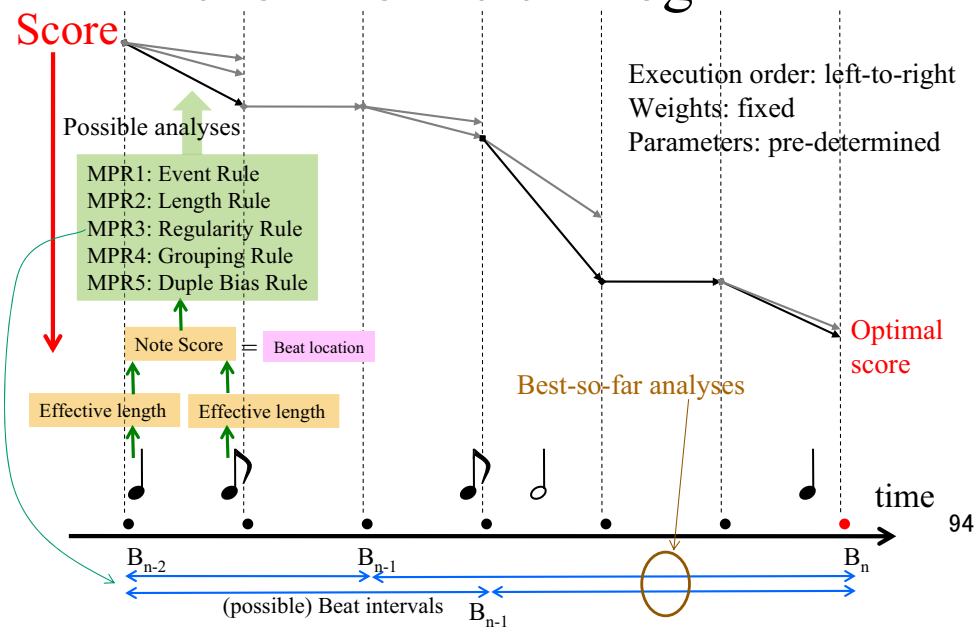
The Melisma System of Temperley & Sleator

Available at: http://www.link.cs.cmu.edu/music-analysis/

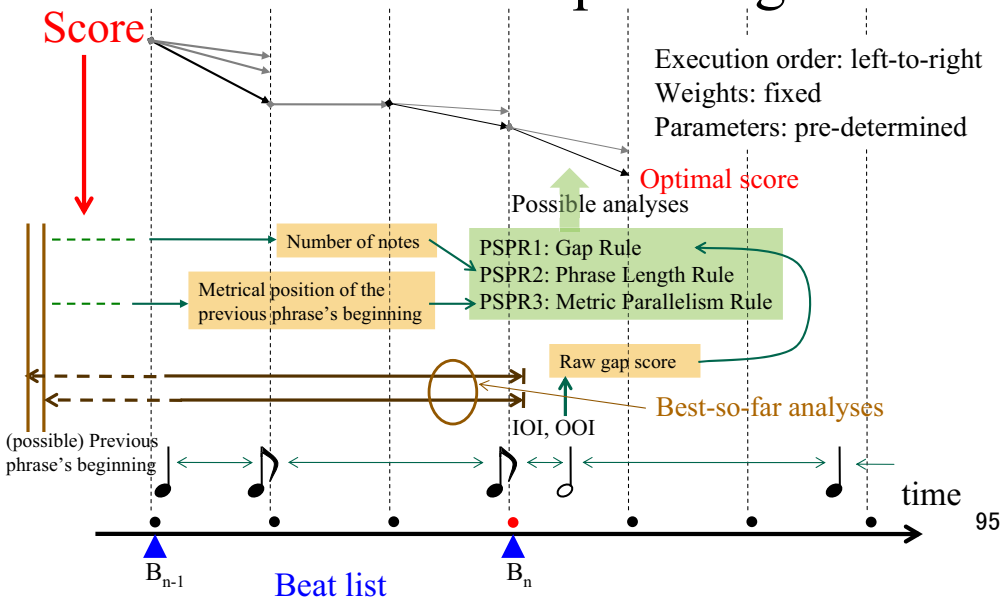
Implementation Problems in Melisma:

- How the system generates and evaluates individual possible analyses
- How Melisma evaluates all possible well-formed analyses of a piece and finds optimal analysis

Dynamic Programming in Melisma's Meter Program



Dynamic Programming in Melisma's Grouper Program



PRS vs. exGTTM

- Well-formedness Rules + Preference Rules
  → Optimization Problem

Constructing Hierarchical Structures

PRS (Melisma)	exGTTM (ATTA)
Preference Rule Systems	GTTM + Full Ext. & Param. + Algorithm for Hierarchy
Left-to-right by Dynamic Programming	Top-down & Bottom-up by Constraint Satisfaction
Ambiguity, Revision, Expectation, and Style in Musical Perception ♣	Ambiguity and Style in Musical Perception
Numerical score of how well the PRS applies to a piece ♥	(Degrees $D_{rule}(i)$ and Weights S_{rule})

♣Temperley 2001, p. 205 ♥Temperley 2001, p. 293

Nord's Implementation of GTTM

- This Prolog implementation covers all sub-theories of GTTM
- An early forerunner for later practical implementations

However:

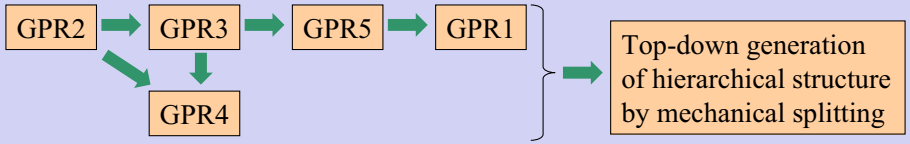
- Straightforward and oversimplified interpretation of rules
- Some rules not implemented

⇒ Very limited applicability

E.g.,
"prefer" into "always choose"
"may be" into "should be"

Implementation can successfully analyze only a very few pieces

Example: Implementation of GPR

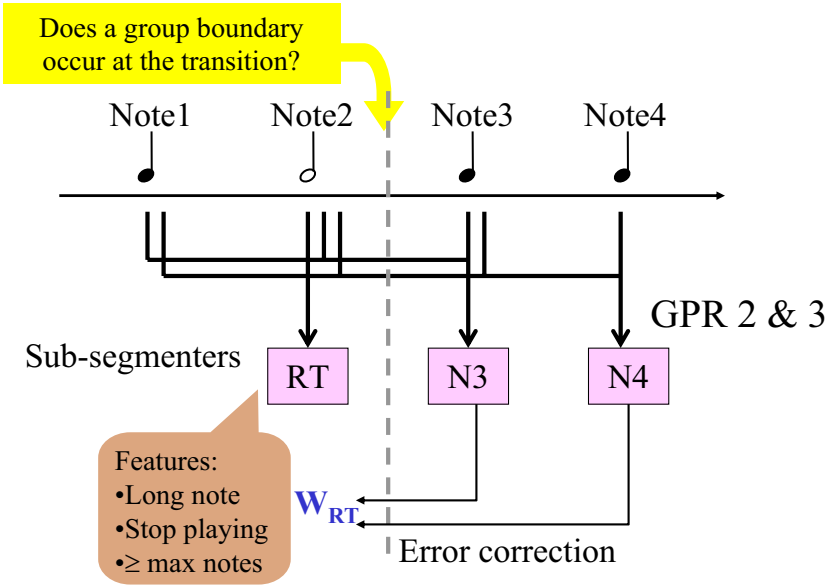


- Execution order is fixed
- There are few parameters
- Parameter values are given *a priori* without musicological reasons

Nord's Interpretations:

GPR1: Where a single-event occurs, a boundary containing "GPR5" chosen
GPR2: } Literally translated into a program
GPR3: }
GPR4: "relatively more pronounced" if time interval or pitch difference is larger than *a priori* threshold value
GPR5: periodically held
GPR6 } Not implemented
GPR7 }

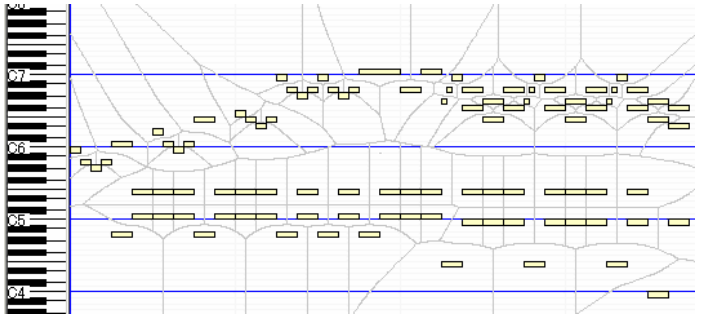
Stammem & Pennycook's Real-time Segmentation



Grouping Polyphony by Voronoi Diagram

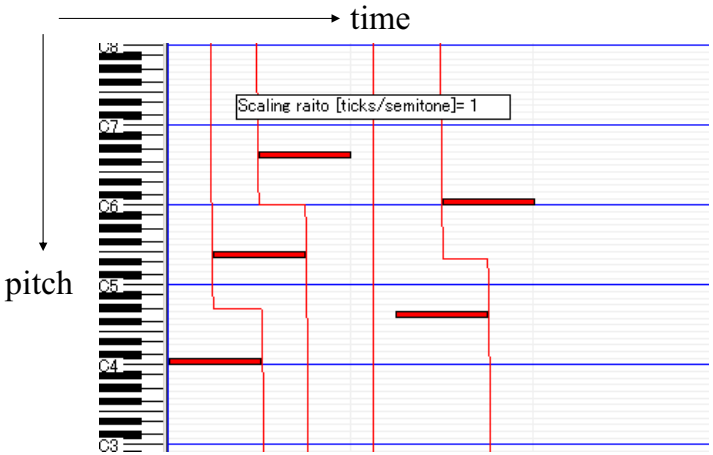
Hamanaka and Hirata (2002)

- Grouping Polyphony
- Conflict between GPR2 (time interval) and GPR3a (register) in polyphony
- Setting the *scaling ratio* of time to pitch for conflict resolution
- Applying Voronoi diagram to a piano roll
- Merging adjacent cells



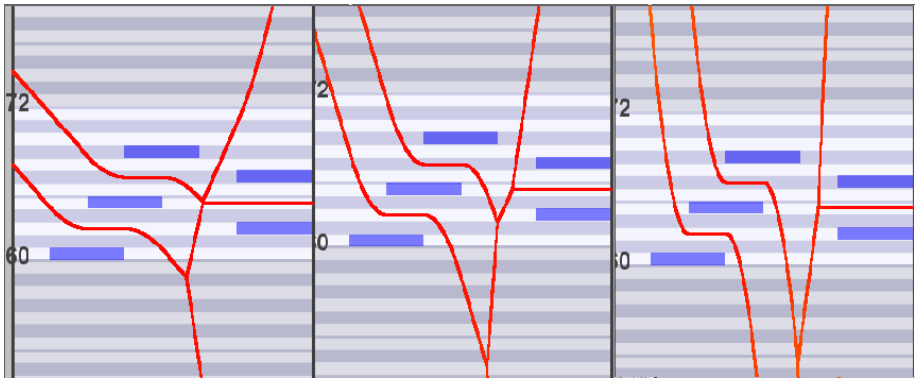
Changes in Scaling Ratio

Shape of a Voronoi diagram depends on scaling ratio of time to pitch



Snapshots of Change in Scaling Ratio

1 semitone=110 ticks 1 semitone=100 ticks 1 semitone=90 ticks



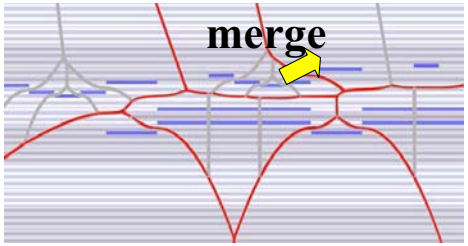
Weak ← Time Scale → Strong

Strong ← Pitch Scale → Weak

Merging Cells for Hierarchical Grouping

Introduce a simple algorithm:

the smallest cell is first merged to the nearest group



Result of our method

Correct data by an expert

Correctness rate = approx. 70%

Automatic Time-span Tree Analyzer: ATTA

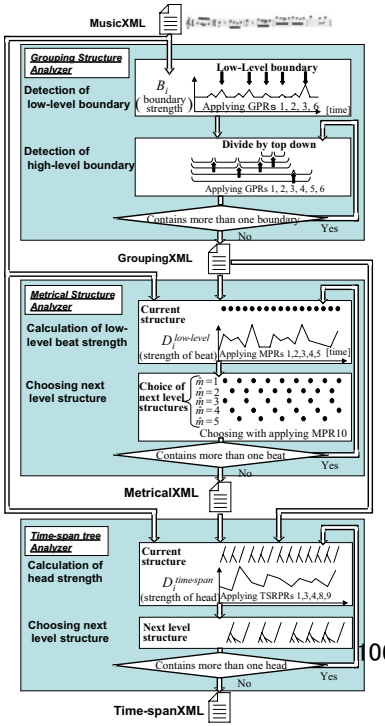
Masatoshi Hamanaka

University of Tsukuba

Grouping Structure Analyzer

Overview of ATTA

- Three analyzers
- Input: score (musicXML)
- Output: groupingXML, metricalXML, time-spanXML



Demo

Grouping Preference Rules: GPR

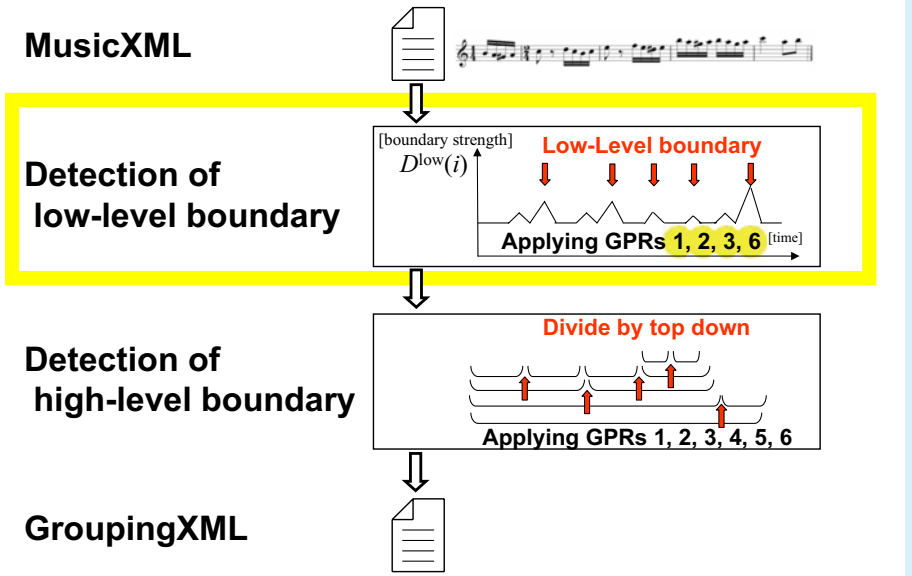
- Applied rules

GPR1(alternative form), GPR2 (Proximity), GPR3(Change), GPR4(Intensification), GPR5(Symmetry), GPR6(Parallelism)

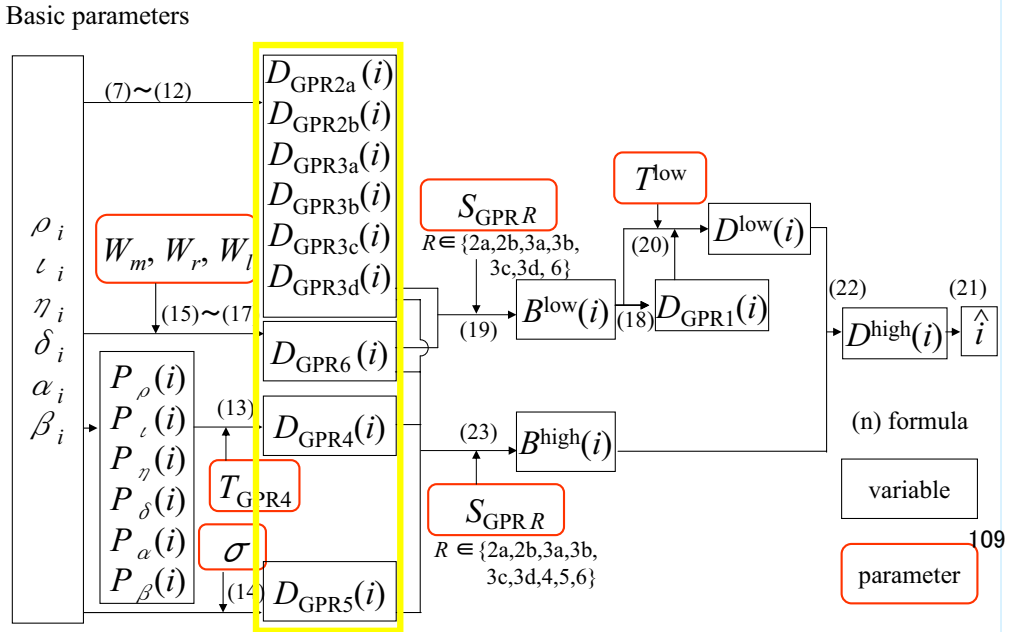
- Unapplied rules

GPR7(Time-span and Prolongational Stability)

Grouping Structure Analyzer

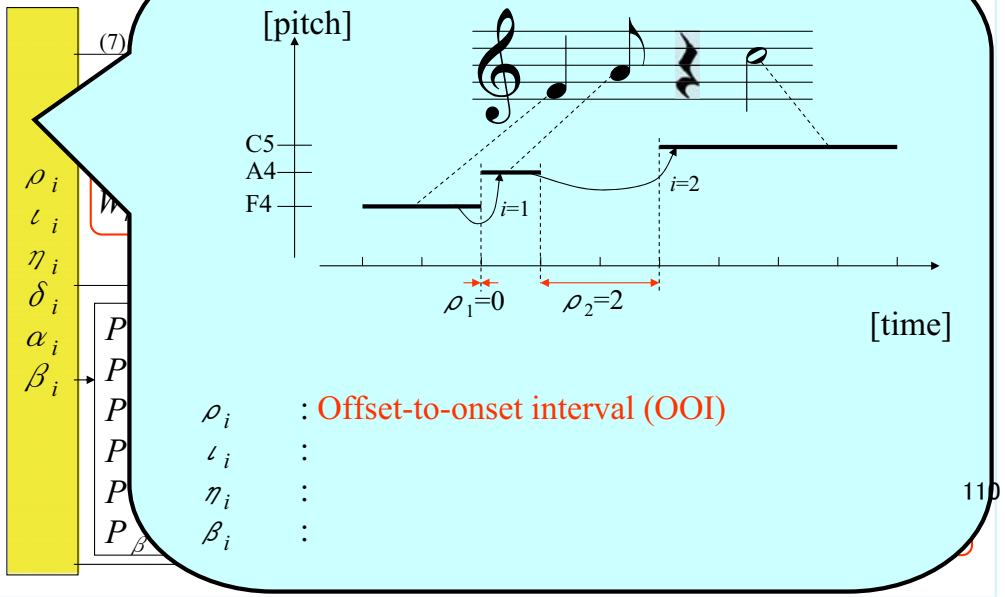


Application of GPRs



Calculation of Basic Parameters

Basic parameters



ρ_i : Offset-to-onset interval (OOI)

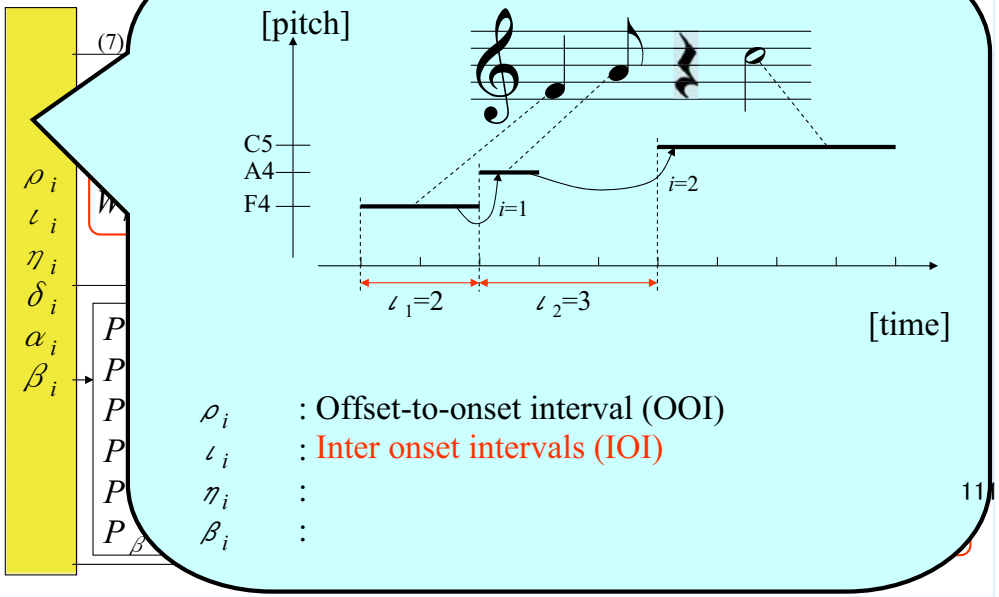
ι_i :

η_i :

β_i :

Calculation of Basic Parameters

Basic parameters



ρ_i : Offset-to-onset interval (OOI)

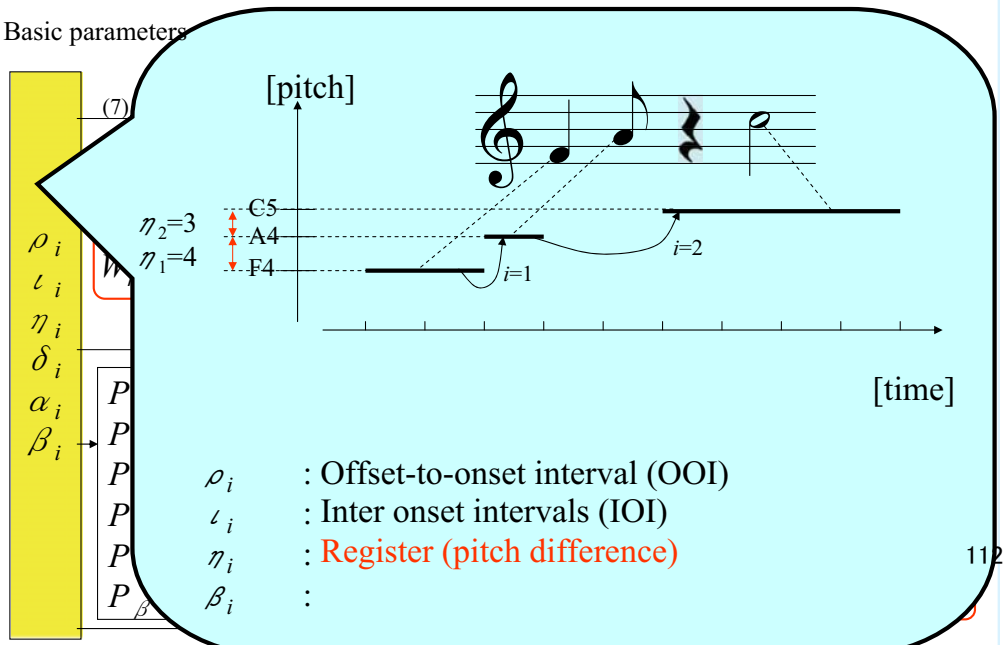
ι_i : Inter onset intervals (IOI)

η_i :

β_i :

Calculation of Basic Parameters

Basic parameters



ρ_i : Offset-to-onset interval (OOI)

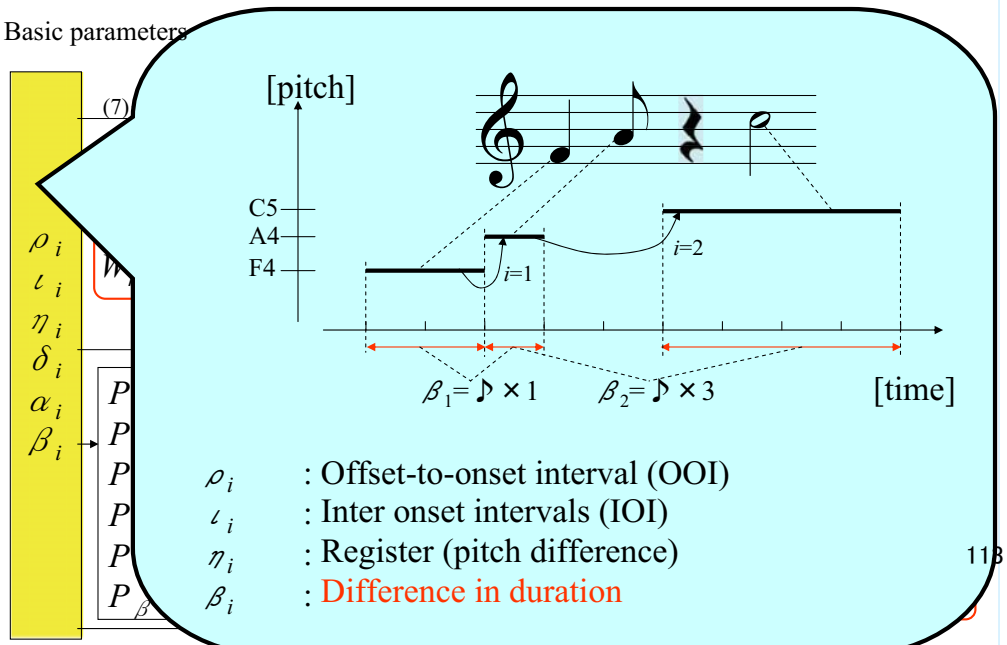
ι_i : Inter onset intervals (IOI)

η_i : Register (pitch difference)

β_i :

Calculation of Basic Parameters

Basic parameters



ρ_i : Offset-to-onset interval (OOI)

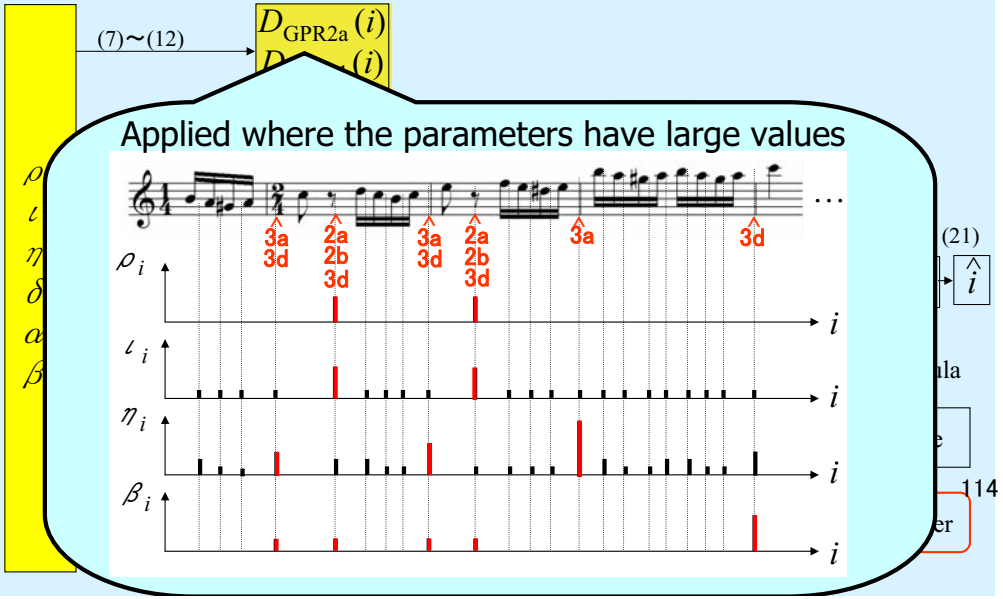
ι_i : Inter onset intervals (IOI)

η_i : Register (pitch difference)

β_i : Difference in duration

Application of GPRs 2 and 3

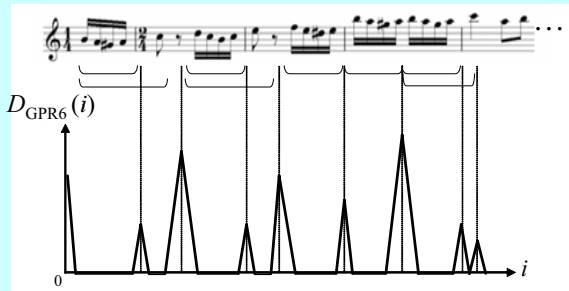
Basic parameters



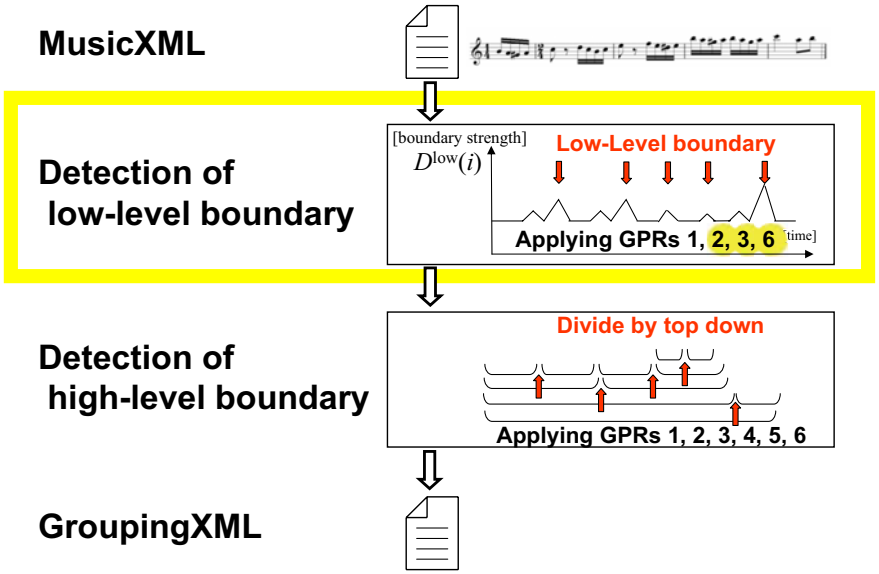
Application of GPR 2 (Parallelism)

Basic parameters

Introduce three adjustable parameters for controlling the degree of parallelism

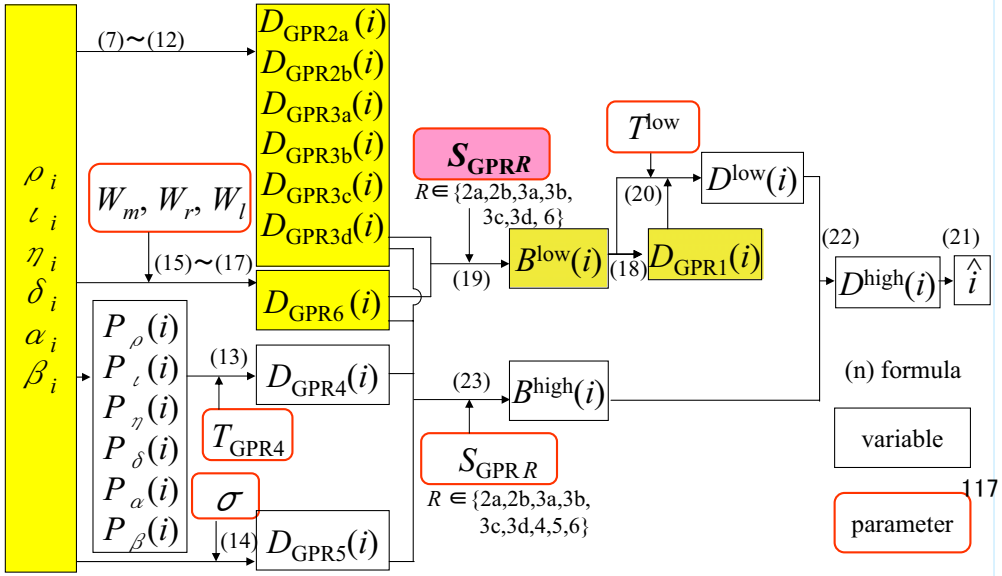


Grouping Structure Analyzer



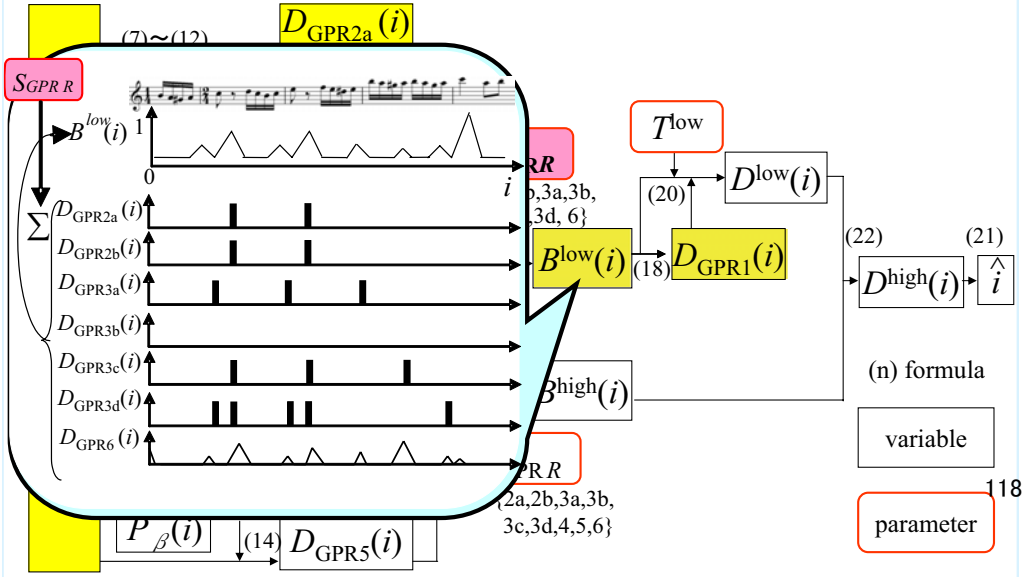
Application of GPR 1

Basic parameters



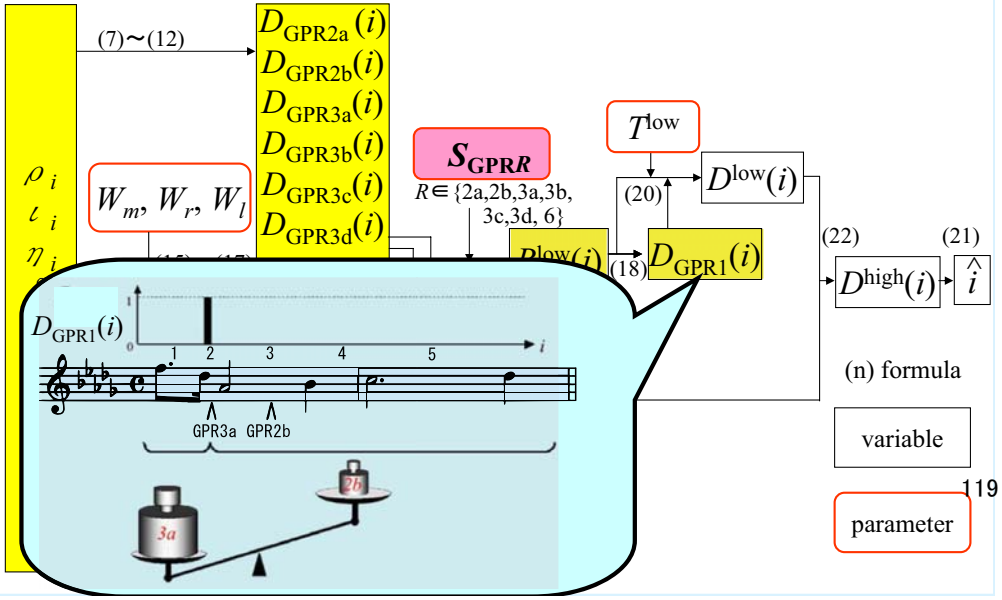
Application of GPR 1

Basic parameters



Application of GPR 1

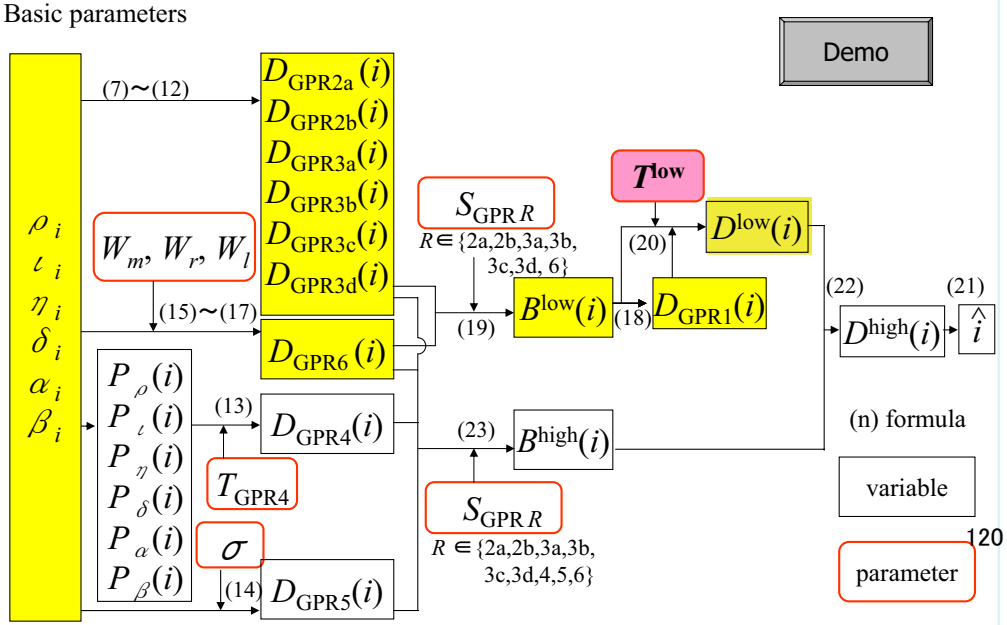
Basic parameters



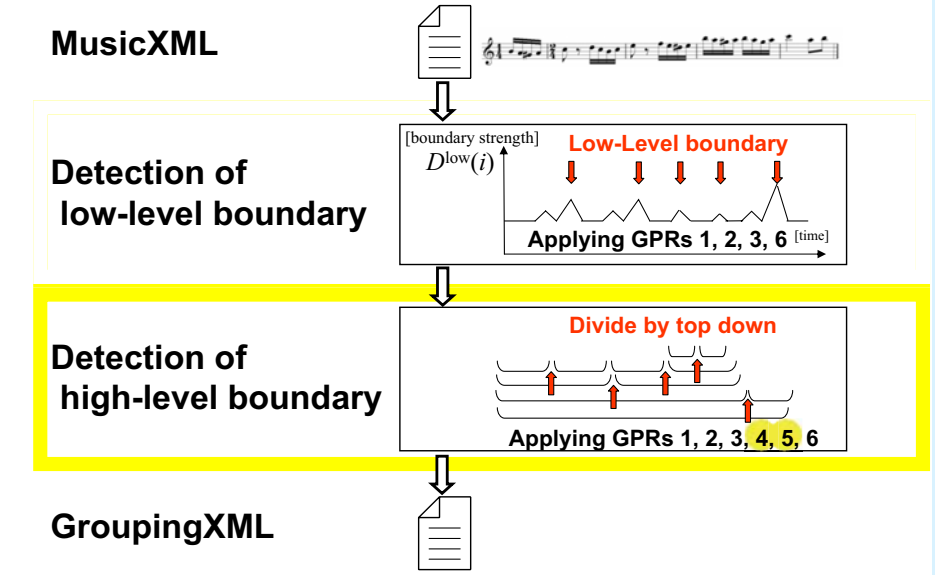
Detection of Low-level Boundary

Basic parameters

Demo



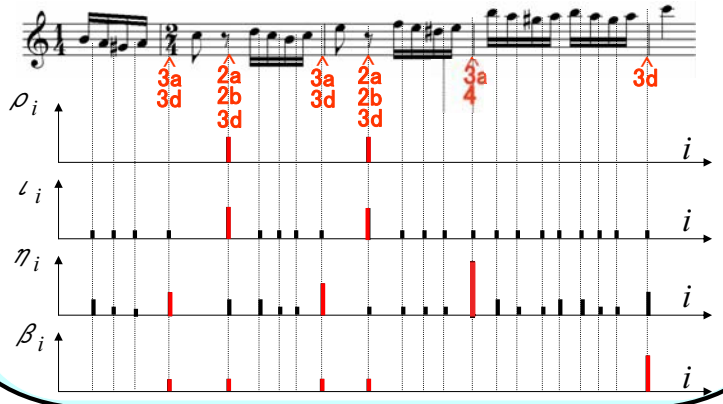
Detection of High-level Boundaries



Applicati… of GPR 4

Basic parameter

Applied where GPRs 2 and 3 are relatively pronounced

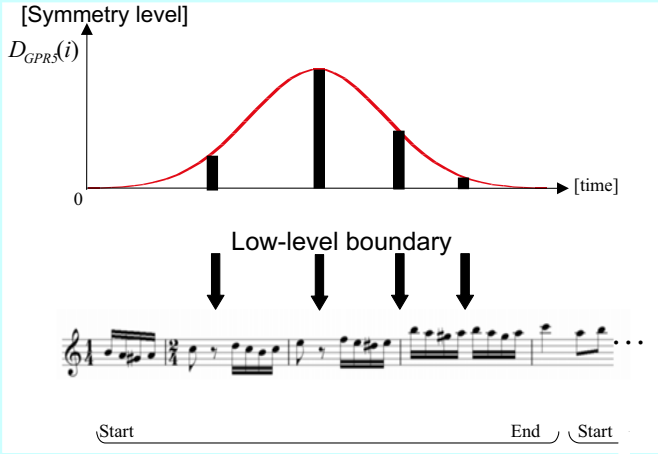


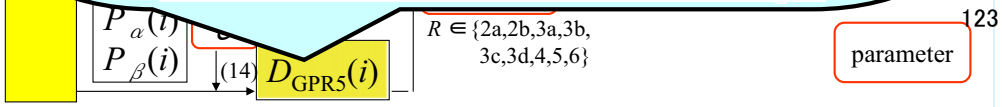
122

Application of GPR 5 (Symmetry)

•Use a normal distribution as a symmetry level

→Preference to subdivide groups into 2 parts of equal length



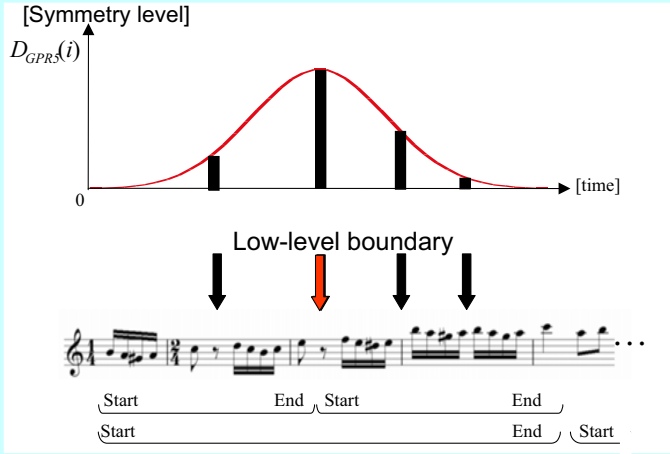


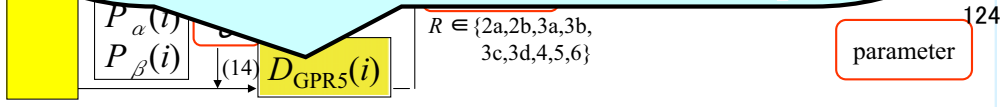
123

Application of GPR 5 (Symmetry)

•Use a normal distribution as a symmetry level

→Preference to subdivide groups into 2 parts of equal length



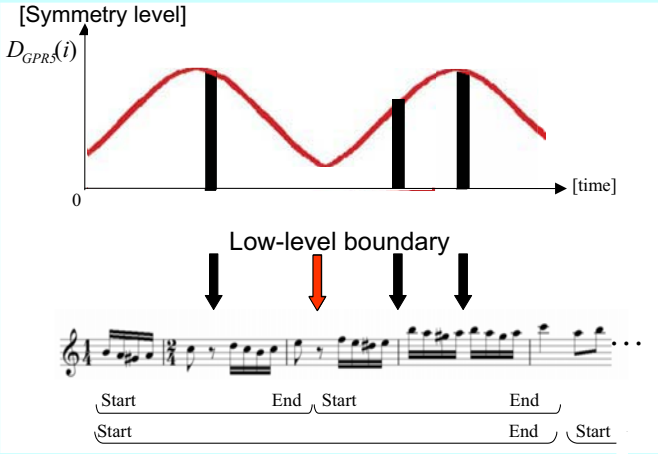


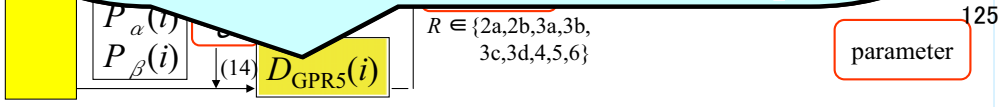
124

Application of GPR 5 (Symmetry)

•Use a normal distribution as a symmetry level

→Preference to subdivide groups into 2 parts of equal length



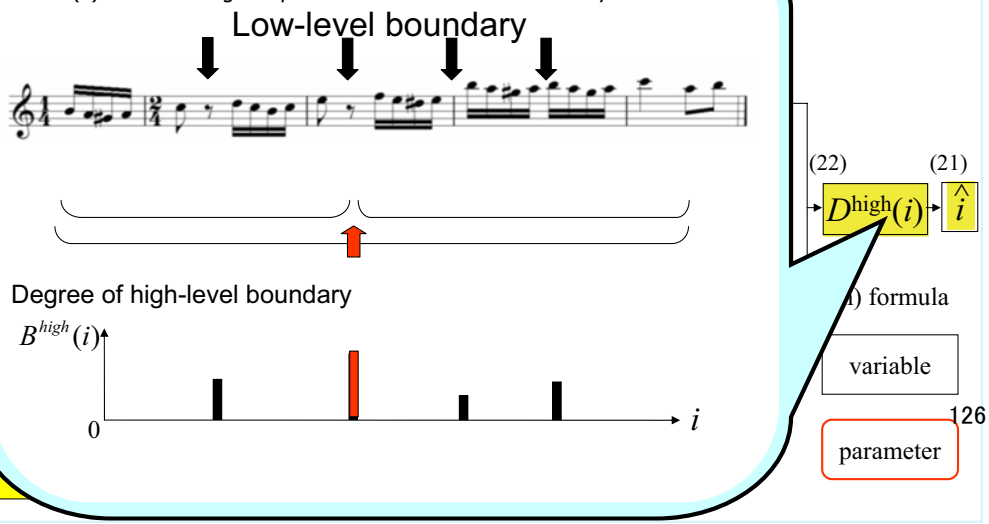


125

Detection of High-level Boundaries

Basic parameters

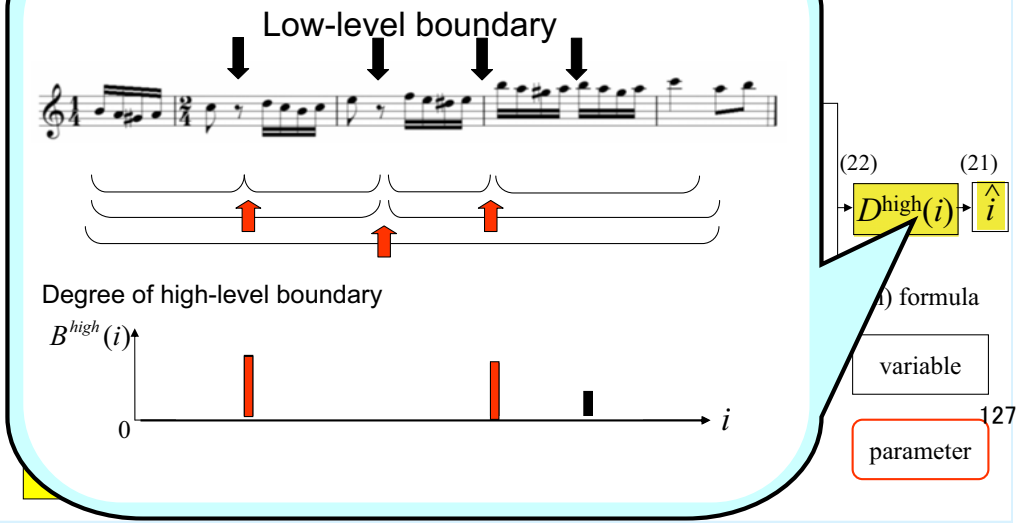
(1) Consider a whole piece as a group
(2) Calculate the degree of high-level boundary
(3) Select the highest position as the next level boundary



Detection of High-level Boundaries

Basic parameters

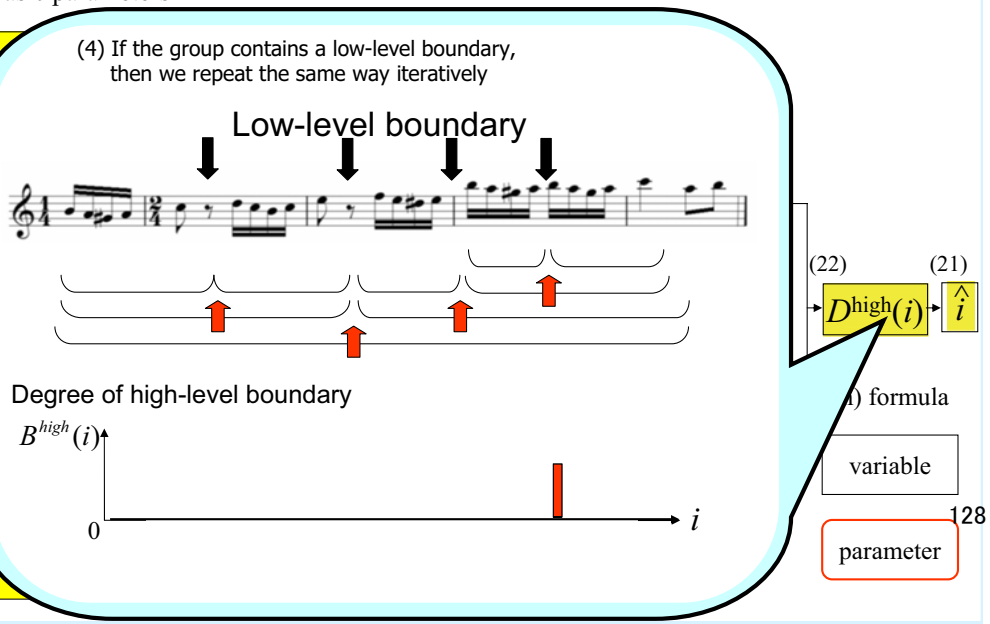
(4) If the group contains a low-level boundary,
then we repeat the same way iteratively



Detection of High-level Boundaries

Basic parameters

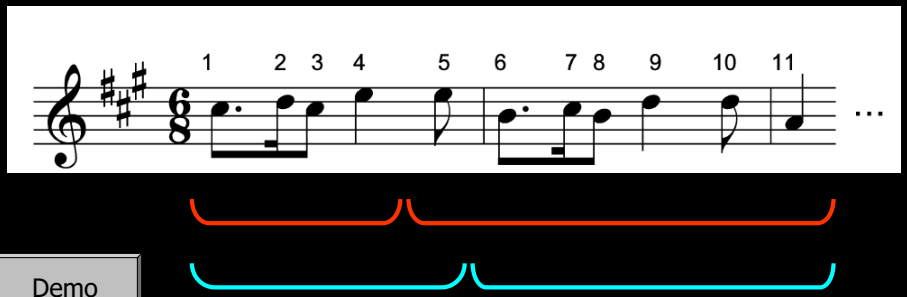
(4) If the group contains a low-level boundary,
then we repeat the same way iteratively



Example of Analysis 1 (Demo)

- Discuss whether our grouping structure analyzer can properly acquire the results
- Can interpret in two ways
- Check whether our grouping structure analyzer can output both structures.

Mozart Sonata K. 331.

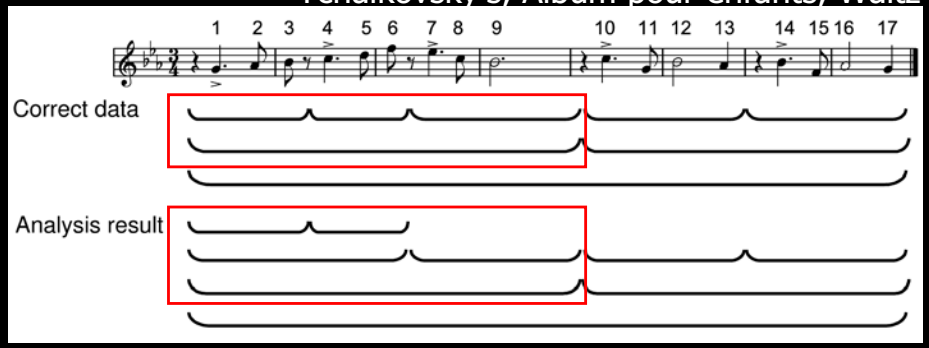


Demo

Example of Analysis 2

- Must hold GPR5, which subdivides groups into two part
- System output tends to have larger numbers of grouping hierarchies

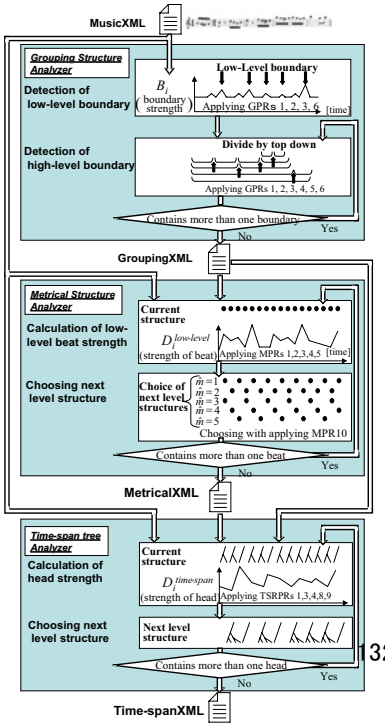
Tchaikovsky's, Album pour enfants, Waltz



Metrical Structure Analyzer

Overview of ATTA

- Three analyzers
- Input: score (musicXML)
- Output: groupingXML, metricalXML, time-spanXML



Metrical Preference Rules: MPR:

- Applied rules

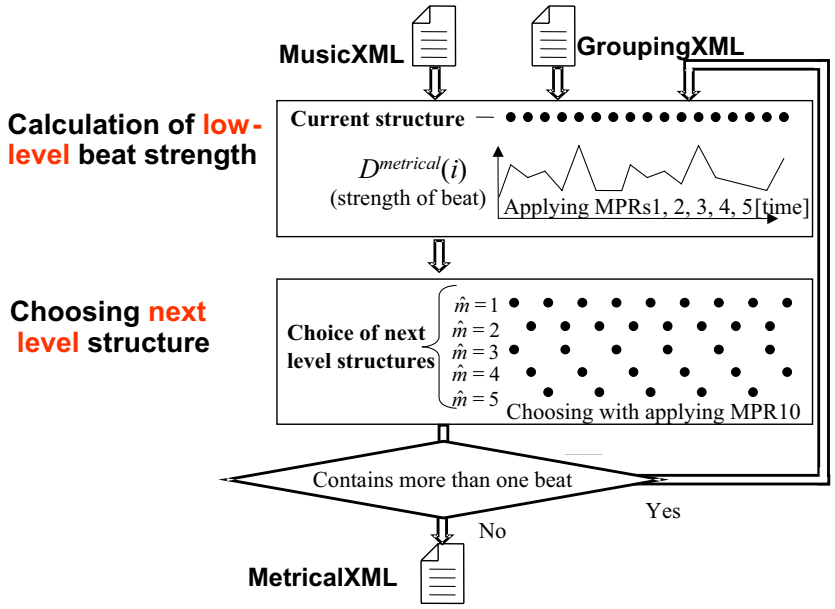
MPR1(Parallelism), MPR2 (Strong beat early), MPR3(Event), MPR4(Stress), MPR5(Length), MPR10(Binary Regularity)

- Unapplied rules

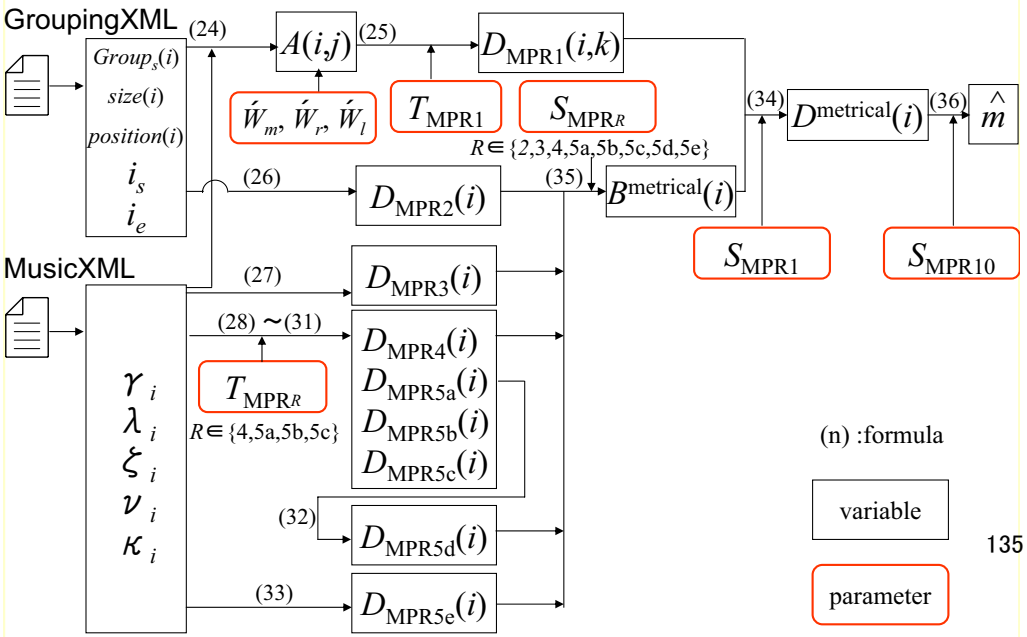
MPR6(Bass), MPR7(Cadence), MPR8(Suspension), MPR9(Time-span interaction)

Demo

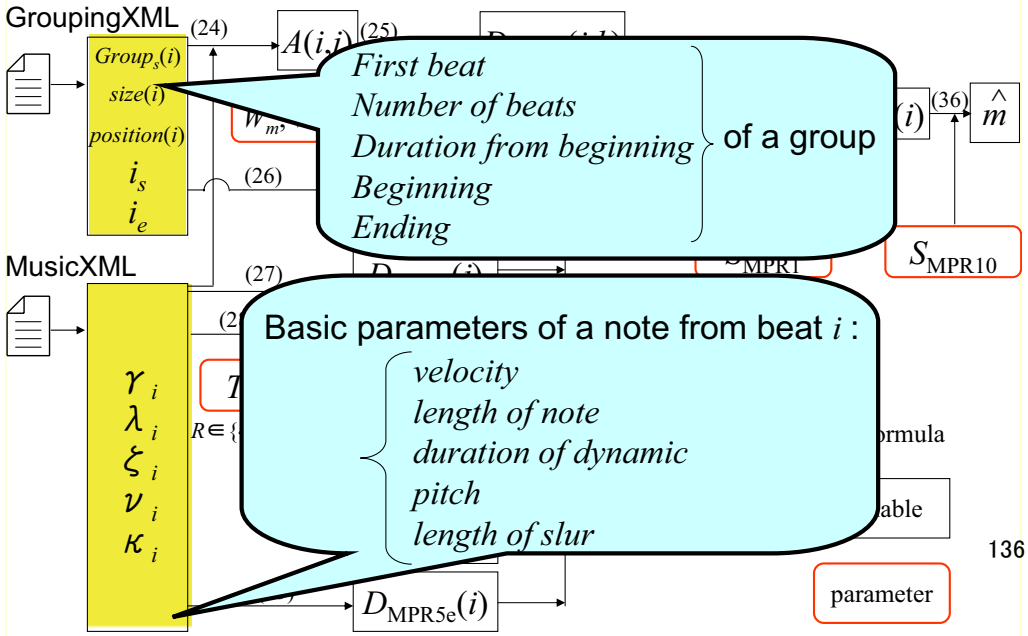
Metrical Structure Analyzer



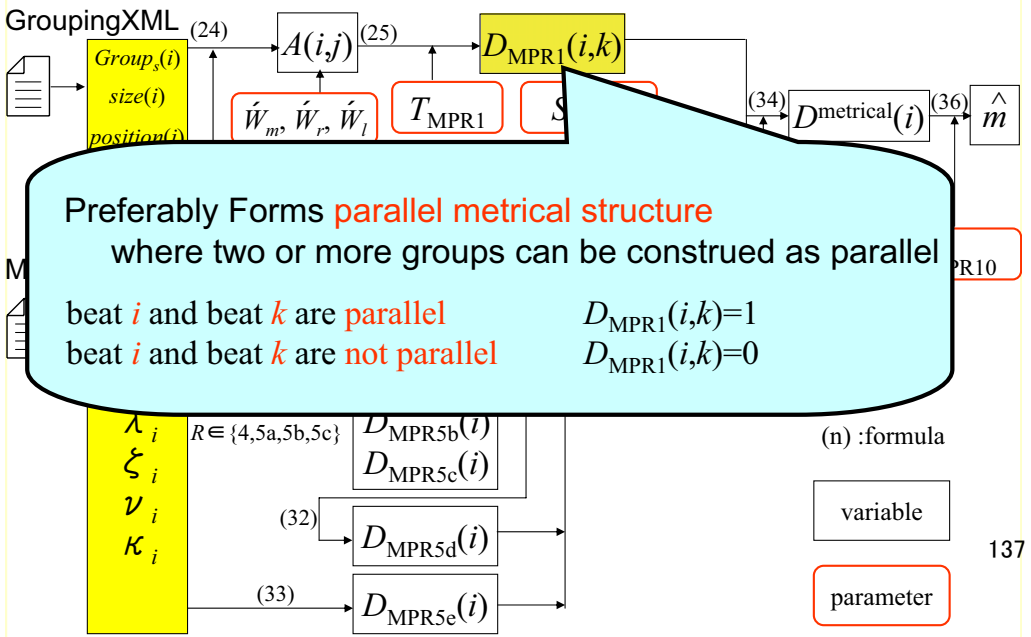
Application of MPRs



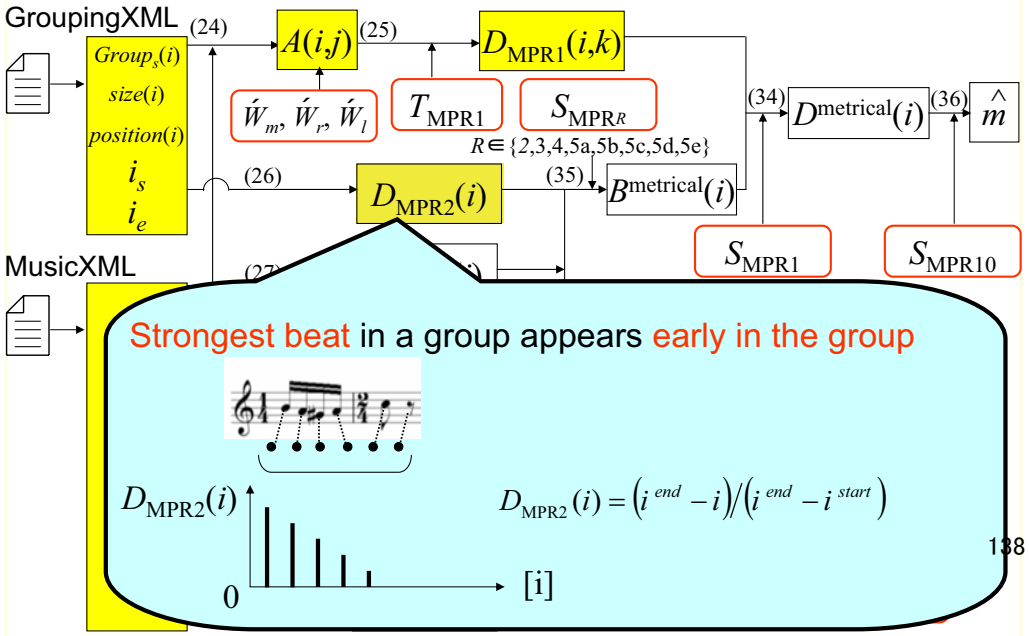
Calculation of basic parameters



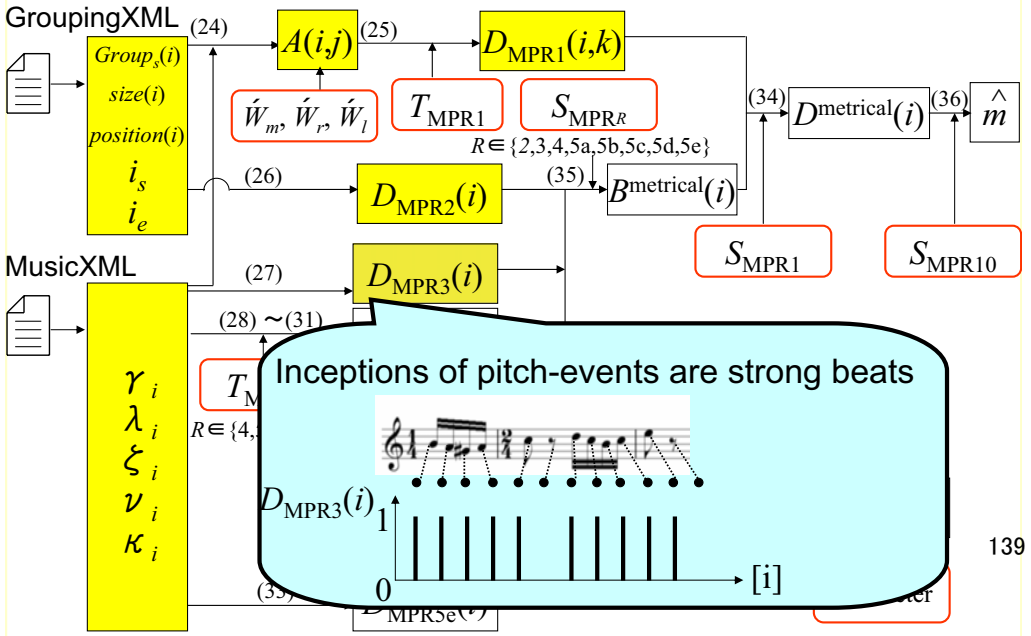
Application of MPR 1 (Parallelism)



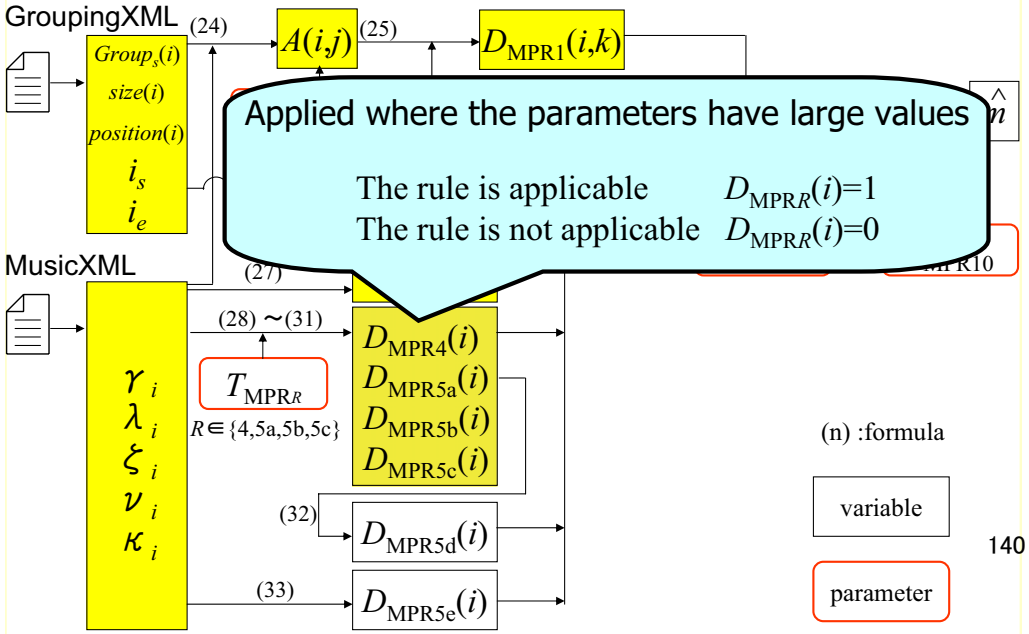
Application of MPR 2



Application of MPR 3



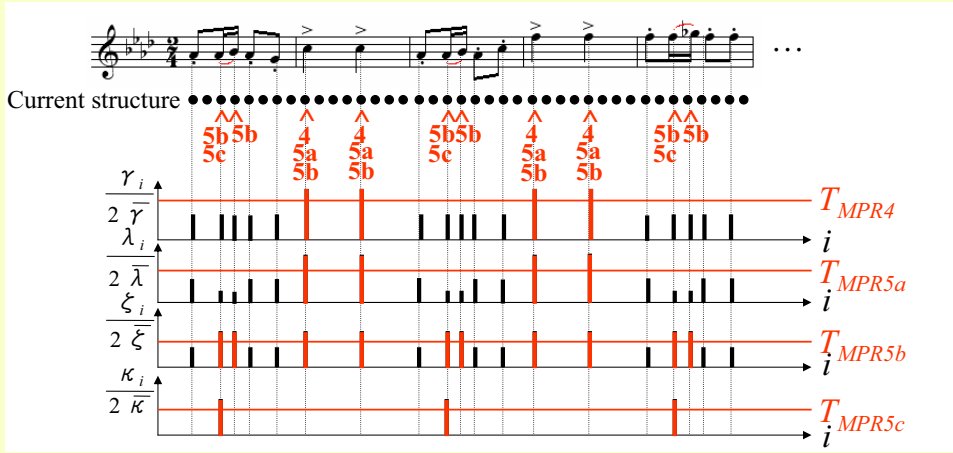
Application of MPRs 4, 5a, 5b, 5c



Application of MPRs 4, 5a, 5b, 5c

- γ_i : velocity
- λ_i : length of note
- ξ_i : duration of dynamic
- κ_i : length of slur

Each rule is applicable where the values are over the threshold T_{MPRj}



Application of MPRs 5d, 5e

MPR5d Prefers a metrical structure with a relatively long pattern of articulation

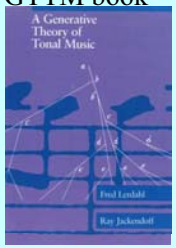
MPR5e Prefers a metrical structure with a relatively long duration of a pitch in the relevant levels of the time-span reduction

Application of MPRs 5d, 5e

MPR5d Prefers a metrical structure with a relatively long pattern of articulation

Inception of MPR5a repetition	$D_{MPR5d}(i)=1$
Otherwise	$D_{MPR5d}(i)=0$

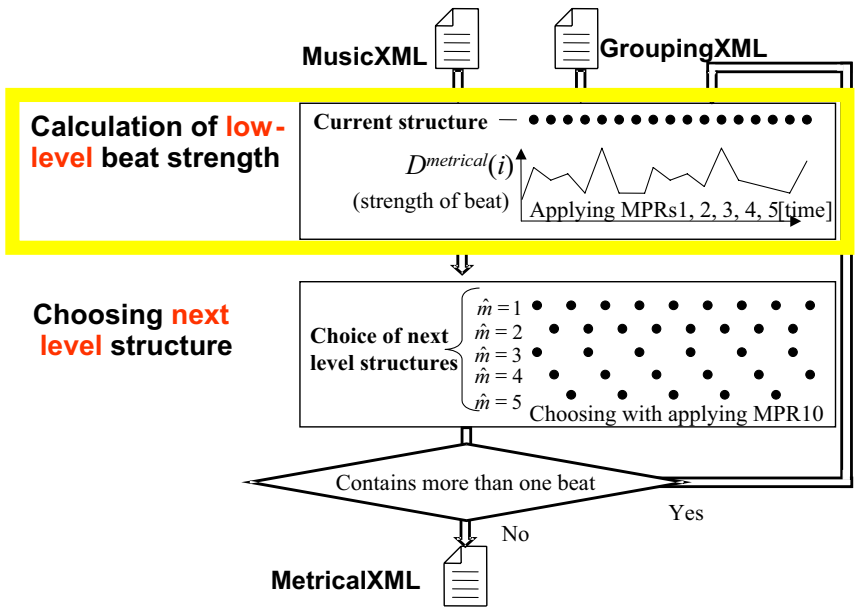
GTTM book



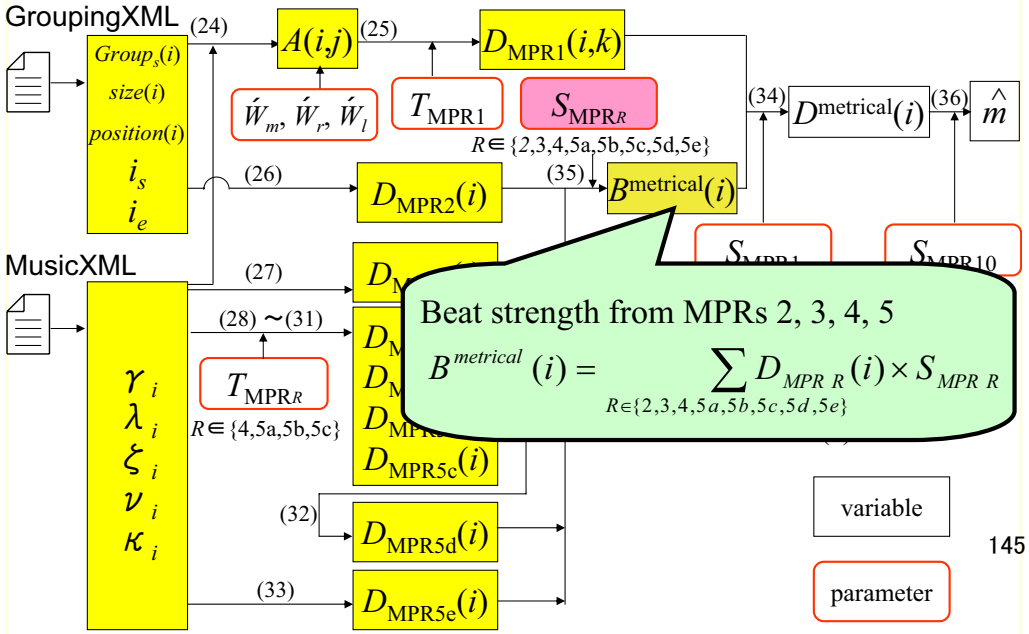
MPR5e Prefers a metrical structure with a relatively long duration of a pitch in the relevant levels of the time-span reduction

Inception of pitch repetition	$D_{MPR5e}(i)=1$
Otherwise	$D_{MPR5e}(i)=0$

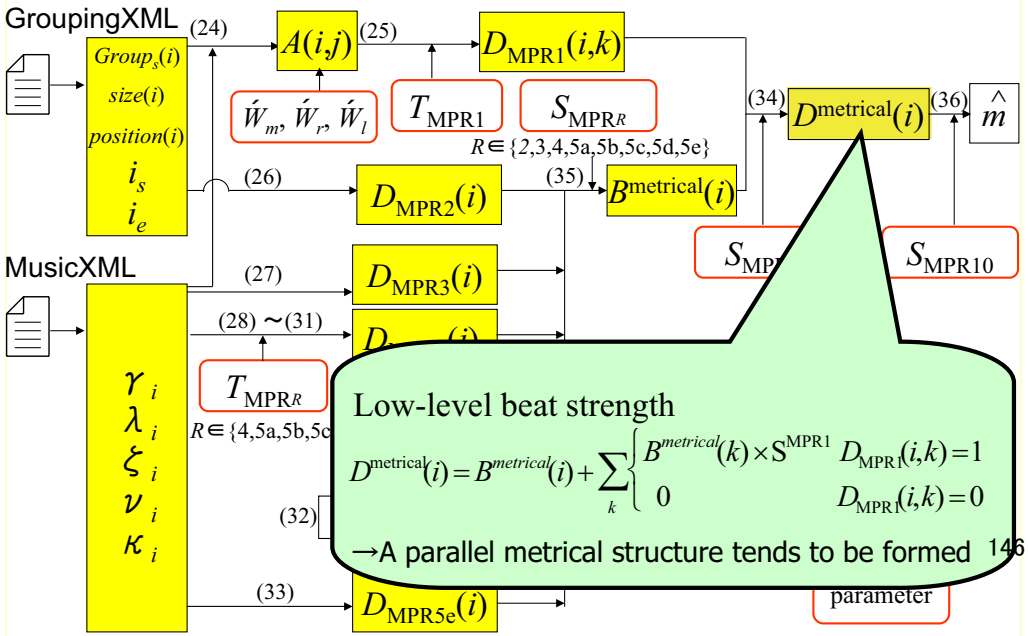
Metrical Structure Analyzer



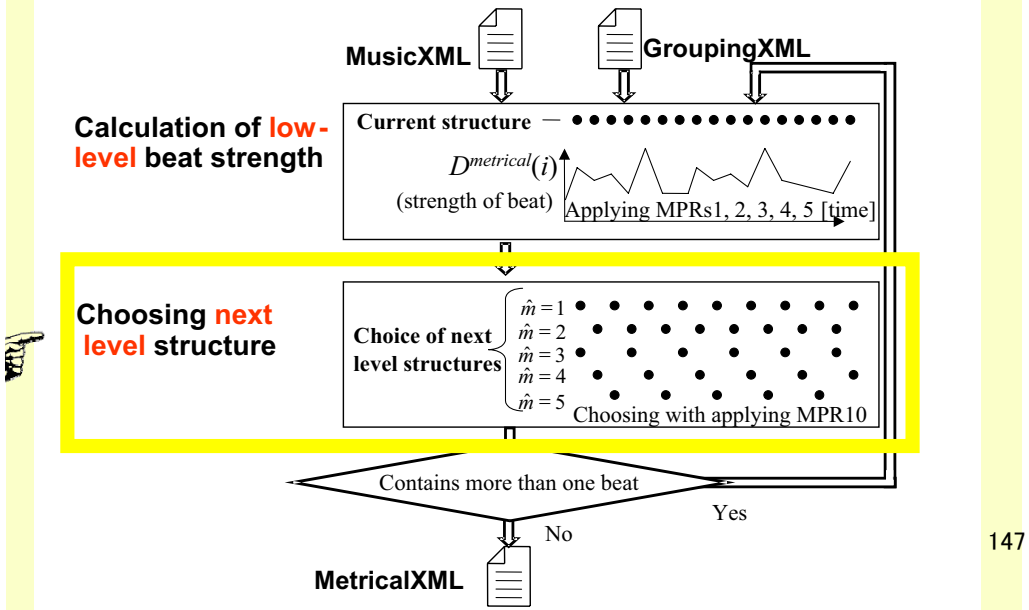
Calculation of Beat Strength



Calculation of Beat Strength

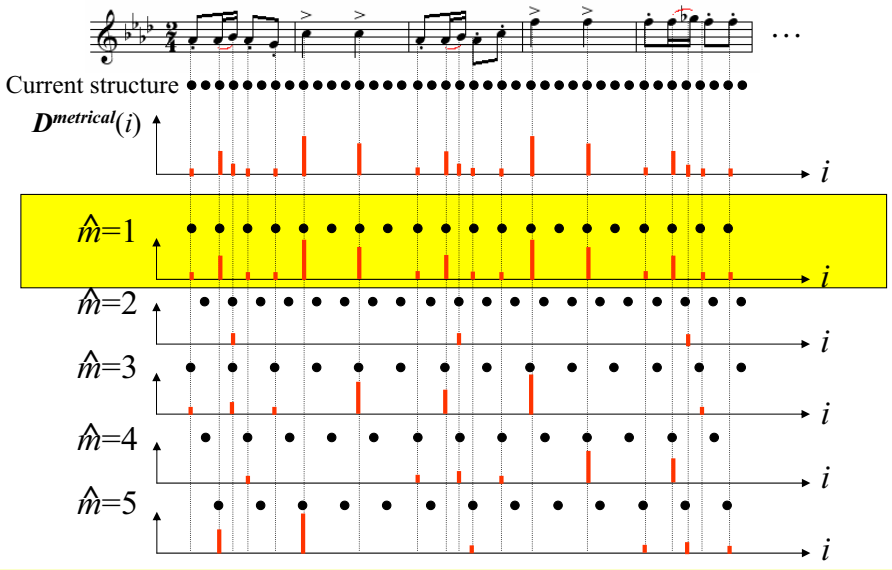


Metrical Structure Analyzer

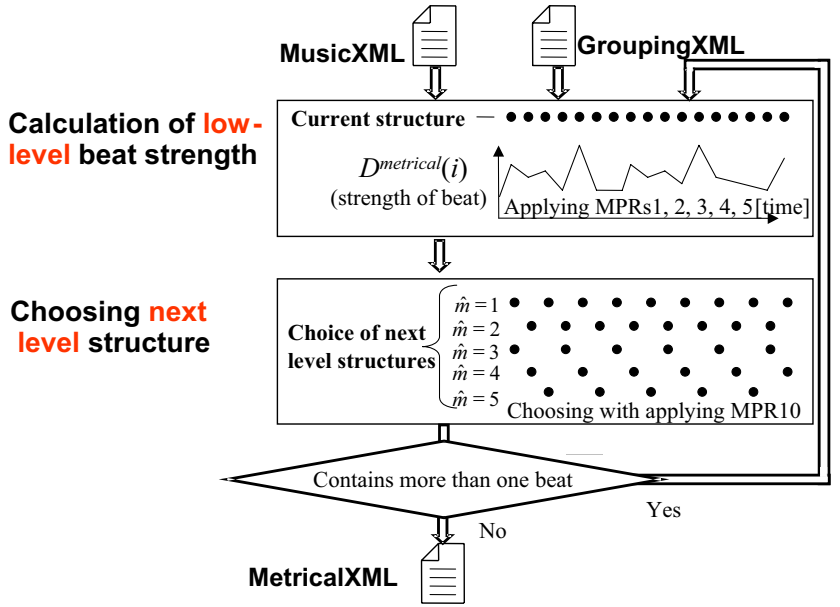


Choosing Next Level Structure

- Calculating the total remaining beat strength



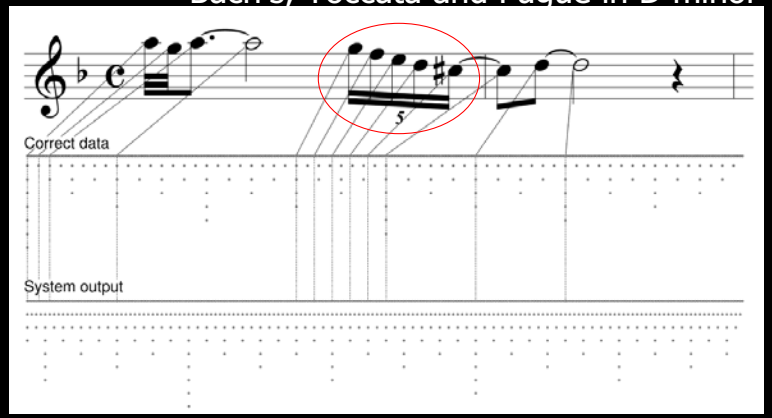
Metrical Structure Analyzer



Example of Analysis

- Most pieces were well-analyzed, but not all
- Quintuplet is not aligned in simple duple/triple time
- Cannot analyzed properly

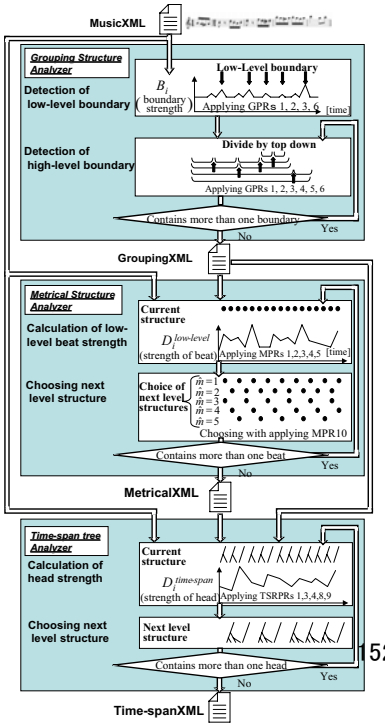
Bach's, Toccata and Fugue in D minor



Time-span Tree Analyzer

Overview of ATTA

- Three analyzers
- Input: score (music XML)
- Output: groupingXML, metricalXML, time-spanXML



Time-span Reduction Preference Rules: TSRPR

- Applied rules

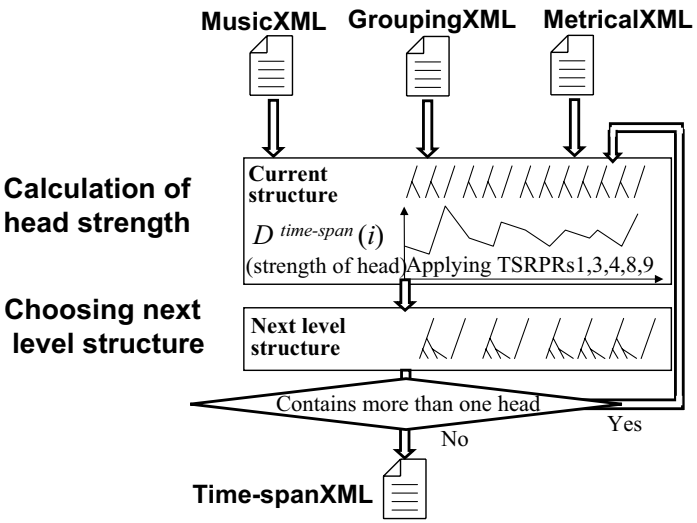
TSRPR1(Metrical position), TSRPR3 (Registral extremes), TSRPR4(Parallelism), TSRPR8(Structural beginning), TSRPR9(Structual ending)

- Unapplied rules

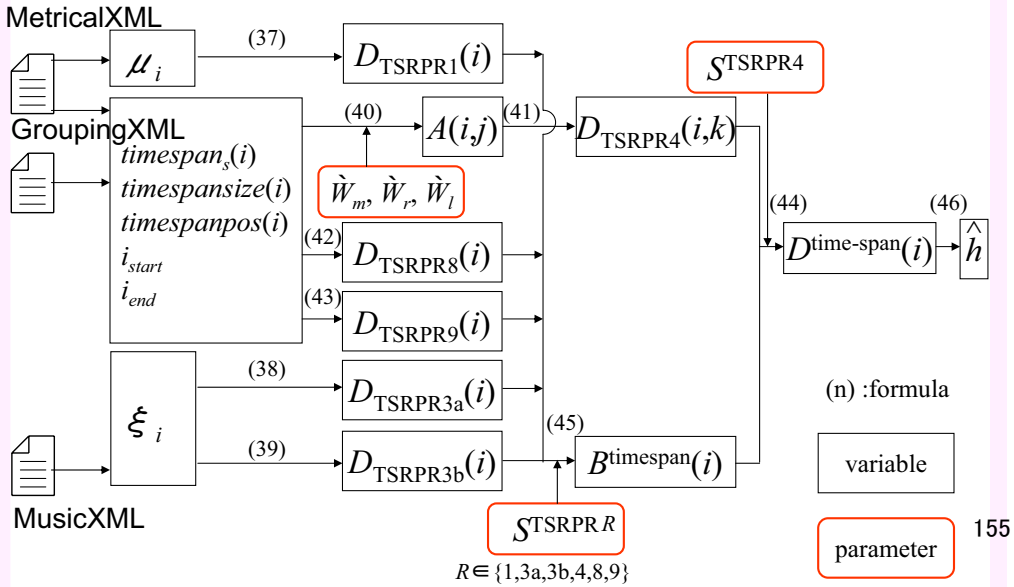
TSRPR2(Local Harmony), TSRPR5(Metrical stability), TSRPR6(Prolongational stability), TSRPR7(Cadential retention),

Demo

Time-span Tree Analyzer

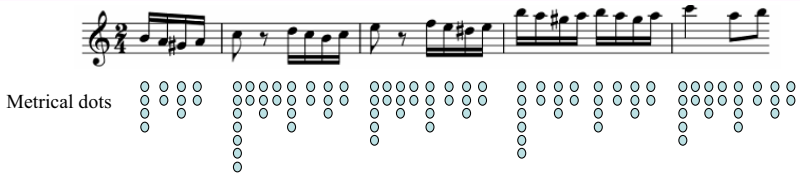


Application of TSRPRs

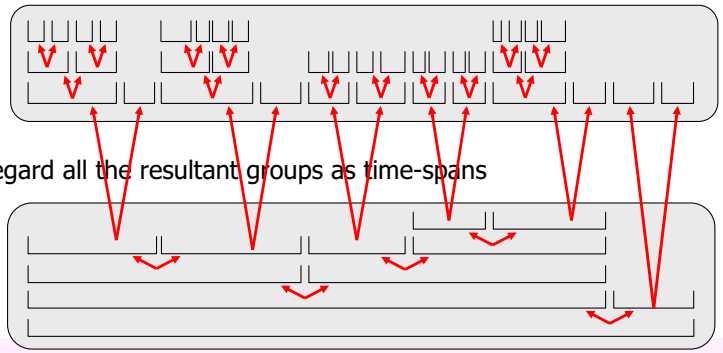


Time-span Segmentation

- Divide the entire piece into hierarchical time-spans

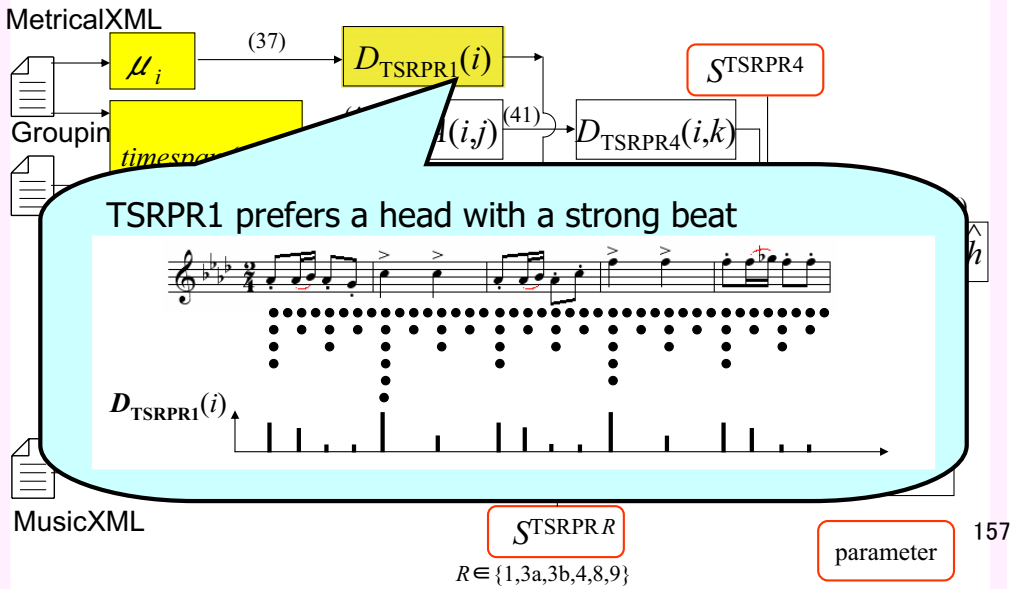


Step2: divide at strongest beat, when time-span includes more than one note

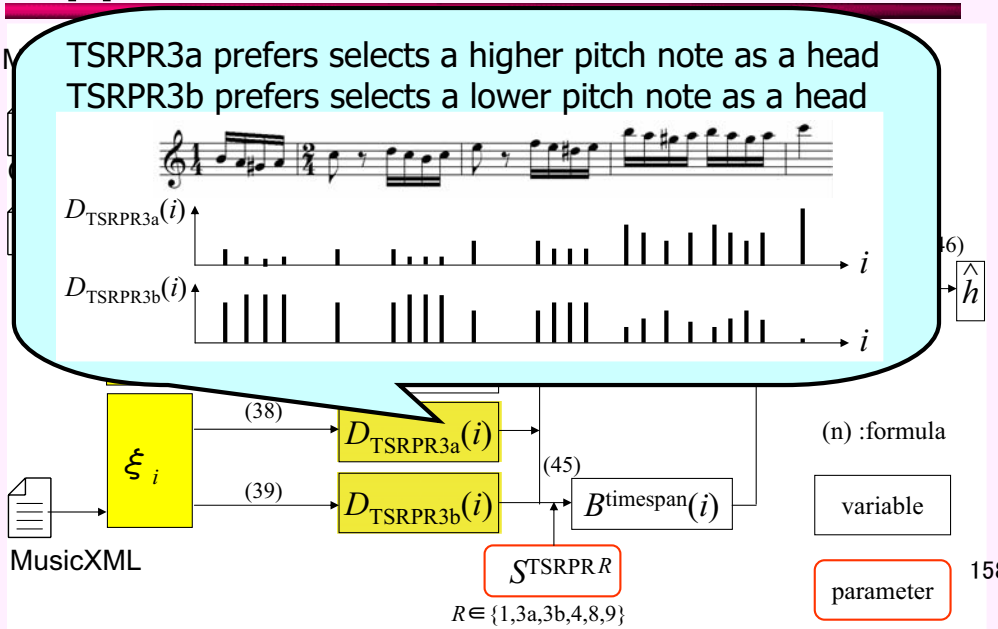


Step1: regard all the resultant groups as time-spans

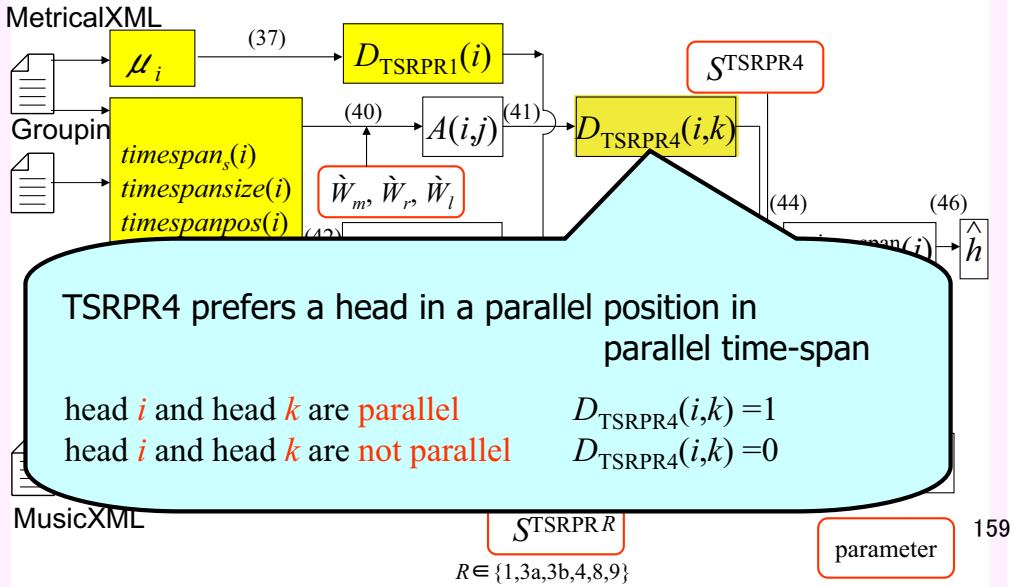
Application of TSRPR 1



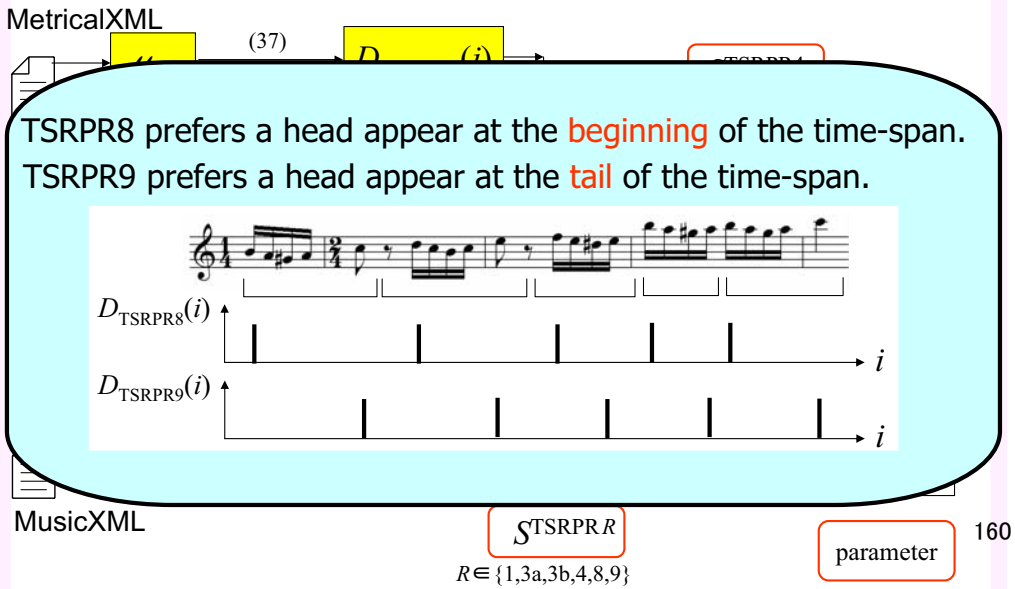
Application of TSRPRs 3a and 3b



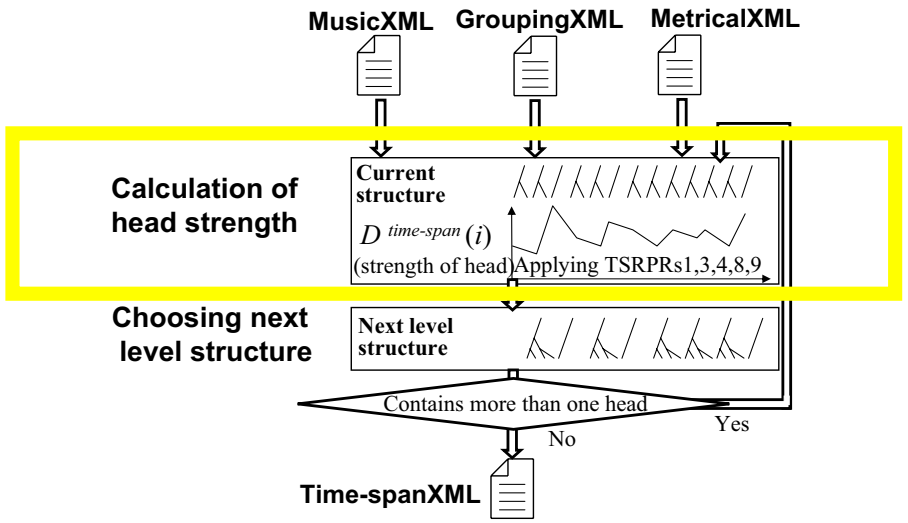
Application of TSRPR 4 (Parallelism)



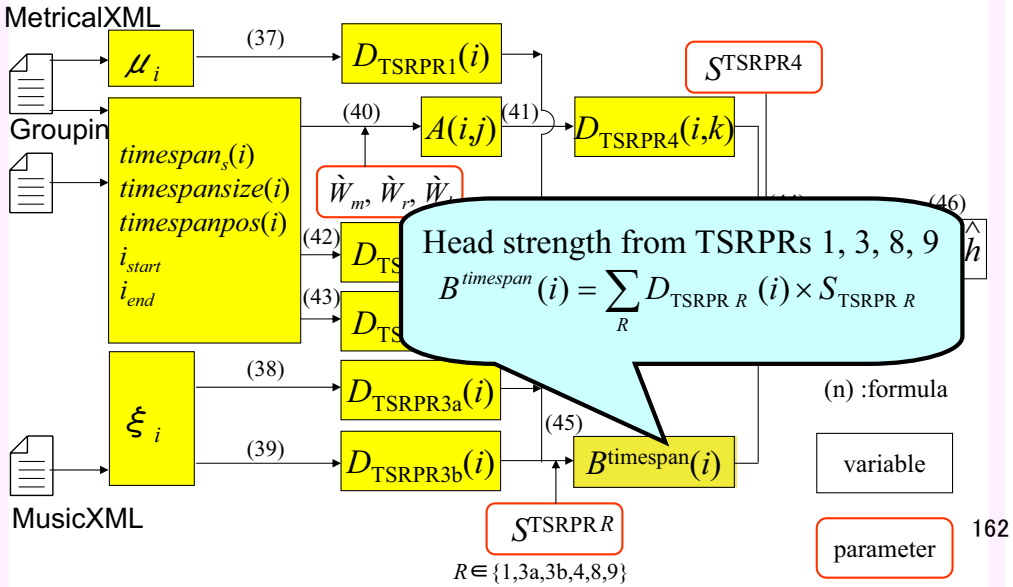
Application of TSRPRs 8 and 9



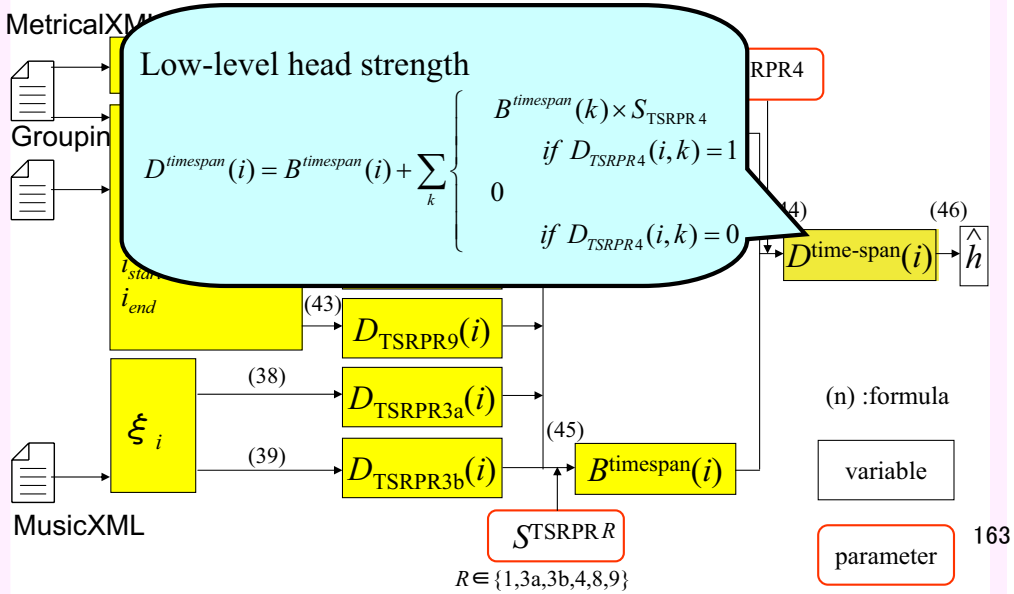
Time-span Tree Analyzer



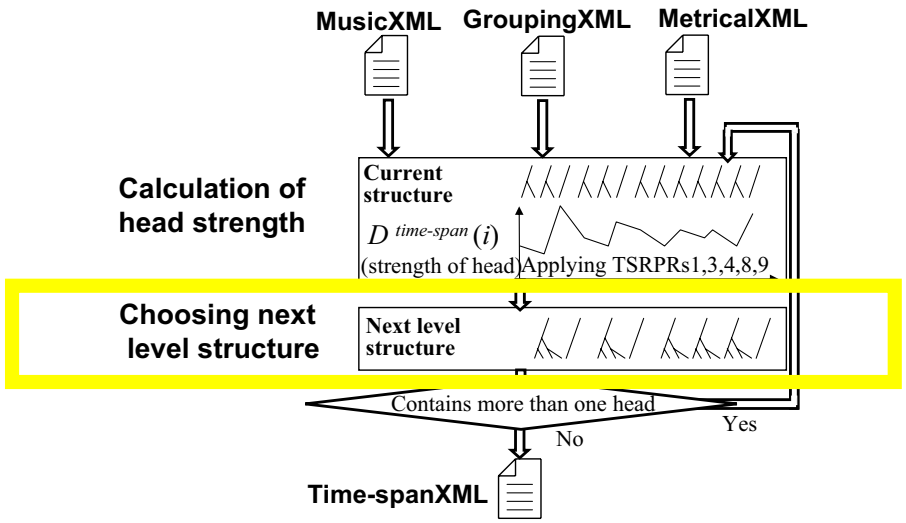
Calculation of head strength



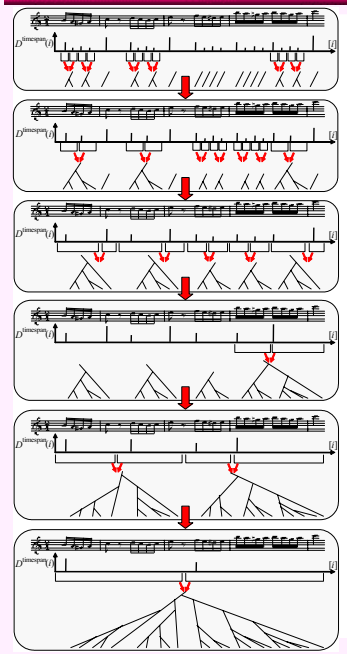
Calculation of head strength



Time-span Tree Analyzer



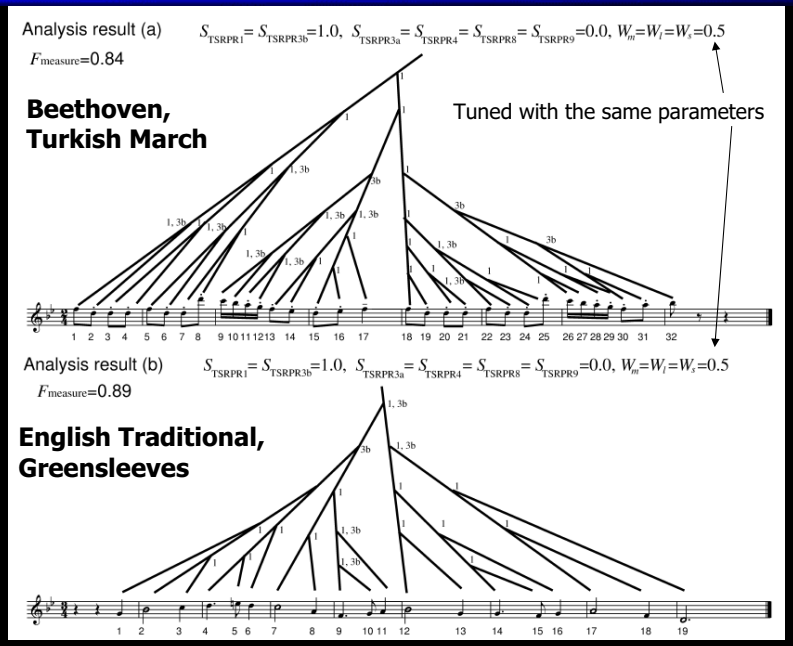
Generating a hierarchical Tree



- Iterating the calculation of the low-level head strength $D^{timespan}(i)$
- Choosing the next level heads,

$$\hat{h} = \begin{cases} i & D^{timespan}(i) \le D^{timespan}(j) \\ j & otherwise \end{cases}$$

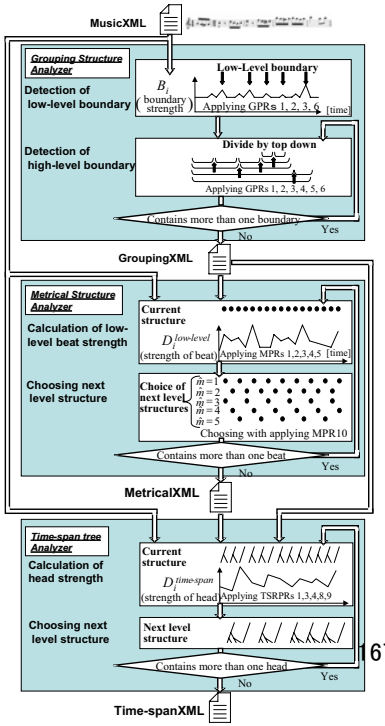
Example of Analysis



Overview of ATTA

- Three analyzers
- Input: score (music XML)
- Output: groupingXML, metricalXML, time-spanXML

Demo

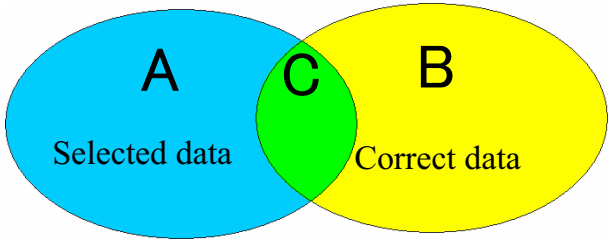


Experimental Results

Experimental results

- Evaluate performance using F-measure
- Required us to prepare the correct data

$$F_{measure} = 2 \times \frac{P \times R}{P + R} \qquad \left(\begin{array}{l} \boldsymbol{P}(Precision) = \frac{C}{A} \\ \boldsymbol{R}(Recall) = \frac{C}{B} \end{array} \right)$$



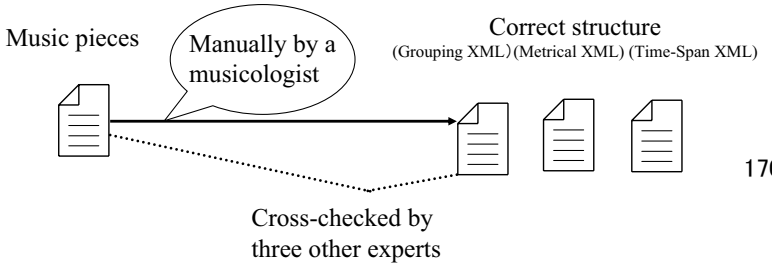
Experimental results

- Evalu
- Requ

$F_{measure}$

Correct data

- 100 pieces of 8-bar length,
  monophonic, classical music pieces
- Analyzed manually by a musicologist
- Cross-checked by three other experts



Experimental results

- **Baseline performance**

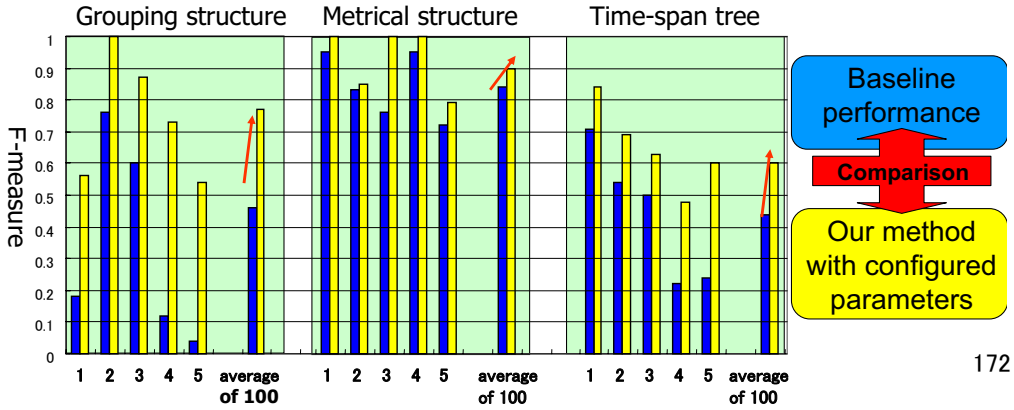
  Default parameters: S^{rule}=0.5, W_r=0.5, W_r=0.5, W_r=0.5, T^{rule}=0.5, and $\sigma = 0.05$

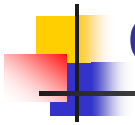
Comparison

- **Our method with configured parameters**
  - Cost about 10 minutes on average
  - Found a plausible tuning of the parameter sets

Experimental results

- All outperformed the baseline performance





Conclusion

- Implemented ATTA on a computer
  Rule reformalization and Refinement of ambiguous rules
  Resolution of preference rule conflicts by prioritizing rules
  Introduction of an algorithm for acquiring a hierarchical structure
- Experiment with correct data
  100 pieces of correct data
  Consistently out-performed the baseline F-measure

The Perl CGI version of ATTA can access
http://music.iit.tsukuba.ac.jp/hamanaka/atta/

Further Information

Further Information (1/5)

Music Theory

- Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*, The MIT Press, Cambridge.
- Lerdahl, F. (2001). *Tonal Pitch Space*. Oxford University Press.
- Schenker, H. (1979). *Free Composition (Der freie Satz)*, translated and edited by Ernst Oster, New York: Longman Inc.
- Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structure*, The University of Chicago Press, Chicago.
- Cooper, G. and Meyer, L. B. (1960). *The Rhythmic Structure of Music*, The University of Chicago Press, Chicago.
- Marsden, A. (2005). Generative Structural Representation of Tonal Music, *Journal of New Music Research*, Vol. 34, No. 4, pp. 409-428.

Music Theory and Implementation

- Temperley, D. (2001). *The Cognition of Basic Musical Structures*, The MIT Press, Cambridge.
- Cope, D. (1995). *Experiments in Musical Intelligence*, A-R Editions, Inc.

Further Information (2/5)

Implementation

- Nord, T. (1992). *Toward Theoretical Verification: Developing a Computer Model of Lerdahl and Jackendoff's Generative Theory of Tonal Music*, Ph.D. Thesis, The University of Wisconsin, Madison.
- Stammen, D. R. and Pennycook, B. (1994). Real-time Segmentation of Music using an Adaptation of Lerdahl and Jackendoff's Grouping Principles, *Proc. ICMPC1994*, pp. 269-270.

Melody Segmentation

- Cambouropoulos, E. (2001). The Local Boundary Detection Model (LBDM) and its application in the study of expressive timing, *Proc. ICMC2001*, pp. 290-293.
- Cambouropoulos, E. (2006). Musical Parallelism and Melodic Segmentation: A Computational Approach. *Music Perception*, Vol. 23, No. 3, pp. 249-267.
- Ferrand, M., Nelson, P. and Wiggins, G. (2003). Memory and Melodic Density: A Model for Melody Segmentation, *Proc. XIV CIM 2003*, pp. 95-98.
- Thom, B. , Spevak, C. and Hothker, K. : Melodic segmentation: evaluating the performance of algorithms and musical expert, *Proc. ICMC2002*, pp. 65-72 (2002).
- Hamanaka, M. and Hirata, K. (2002). Applying Voronoi Diagrams in the Automatic Grouping of Polyphony, *Information Technology Letters*, Vol. 1, No. 1, pp. 101-102.
- Selfridge-Field, E. (1998). Conceptual and Representational Issues in Melodic Comparison, *Computing in Musicology 11*, The MIT Press, pp. 3-64.
- Hewlett, W. B. ed. (1998). Melodic Similarity: Concepts, Procedures, and Application. *Computing in Musicology*, Vol. 11, The MIT press, Cambridge.

Further Information (3/5)

Application

- Hirata, K. and Hiraga, R. (2003). Ha-Hi-Hun plays Chopin's Etude, *Working Notes of IJCAI-03 Workshop on Methods for Automatic Music Performance and their Applications in a Public Rendering Contest*, pp. 72-73.
- Hirata, K. and Matsuda, S. (2003). Interactive Music Summarization based on Generative Theory of Tonal Music, *Journal of New Music Research*, Vol. 32, No. 2, pp. 165-177.

Music Representation and AI

- Hirata, K. and Aoyagi, T. (2003). Computational Music Representation on the Generative Theory of Tonal Music and the Deductive Object-Oriented Database, *Computer Music Journal*, Vol. 27, No. 3, pp. 73-89.
- Grachten, M., Arcos, J.-L., and Mantaras, R. L. (2004). Melodic Similarity: Looking for a Good Abstraction Level. *Proc. ISMIR2004*, pp.210-215.
- Balaban, M., Ebcioğlu, K., and Laske, O. (Eds) (1992). Understanding Music with AI: Perspectives on Music Cognition. The MIT Press.
- Hofstadter, D. (1979). Gödel, Escher, Bach: An Eternal Golden Braid. Basic Books.
- Hofstadter, D. (1985). Metamagical Themas. Basic Books.

Further Information (4/5)

Our Work

- Hamanaka, M., Hirata, K. and Tojo, S. (2004). Automatic Generation of Grouping Structure based on the GTTM, *Proc. ICMC2004*, pp. 141-144.
- Hamanaka, M., Hirata, K. and Tojo, S. (2005). Automatic Generation of Metrical Structure based on the GTTM, *Proc. ICMC2005*, pp. 53-56.
- Hamanaka, M., Hirata, K. and Tojo, S. (2005). ATTA: Automatic Time-span Tree Analyzer based on Extended GTTM, *Proc. ISMIR2005*, pp. 358-365.
- Hamanaka, M., Hirata, K. and Tojo, S. (2006). Implementing "A Generative Theory of Tonal Music", *Journal of New Music Research*, Vol. 35, No. 4, pp. 249-277. (To Appear)
- Hamanaka, M., Hirata, K. and Tojo, S. (2007). FATTA: Full Automatic Time-span Tree Analyzer, *Proc. ICMC2007*, Vol. I, pp. 153-156.
- Oka, Y., Hamanaka, M., Hirata, K. and Tojo, S. (2007). Optimal Parameter Set Acquisition for exGTTM. *Proc. Music-AI 2007*, pp. 95-106.

Further Information (5/5)

Language Theory

- Jackendoff, R. (1983). *Semantics and Cognition*. The MIT Press.
- Jackendoff, R. (1993). Patterns in the Mind, Harvester Wheatsheaf.
- Jackendoff, R. (2003). Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford University Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge. The MIT Press.
- Pinker, S. (1994). *The Language Instinct*. Perennial (HarperCollins).

Web Pages

- Heikki, V. (2000). Lerdahl and Jackendoff Revisited:
  http://www.cc.jyu.fi/~heivalko/articles/lehr_jack.htm
- David Meredith's Web Site:
  http://www.titanmusic.com/

The End