



**Tree
Automata
Techniques and
Applications**

HUBERT COMON MAX DAUCHET RÉMI GILLERON
FLORENT JACQUEMARD DENIS LUGIEZ CHRISTOF LÖDING
SOPHIE TISON MARC TOMMASI

For:

$$\begin{aligned}x + 0 &\rightarrow x \\x + s(y) &\rightarrow s(x + y)\end{aligned}$$

the normal forms are: $0, s(0), s(s(0)), \dots$

This set is generated by the following Regular Tree Grammar with one non-terminal q :

$$q \rightarrow 0 \quad q \rightarrow s(q)$$

In general, it is the case for sets of rewrite rules with a *linear* left-hand-side (no repetition of variable).

- Normal Forms are the terms that do not embed a left-hand-side,
- Regular Tree Grammar (or equivalently Tree Automata) can do linear pattern matching,
- these languages are closed by complementation.

This is no more true for rules with *non-linear* left-hand-side such as

$$(y \cdot x)/x = y$$

or

$$\textit{ins}(x, \textit{ins}(x, y)) = \textit{ins}(x, y).$$

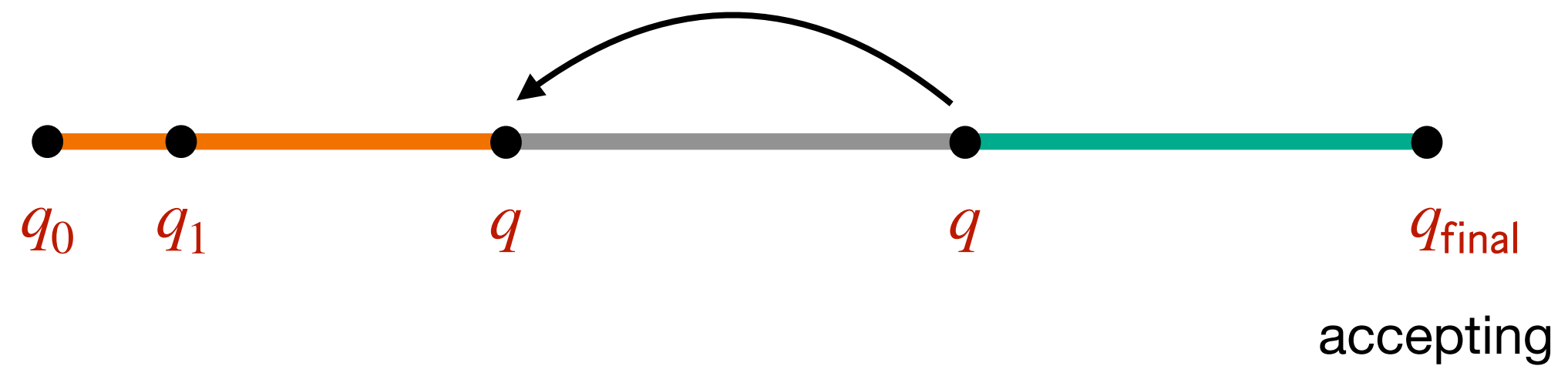
For such cases, tree automata were extended (in Lille) with some *constraints*.

With some local tests of *disequalities* between subtrees during tree automata computation, one can characterise sets of normal form of any set of rewrite rules.

For instance, for the above rule, when detecting a pattern $(y \cdot x)/x'$, the automaton must check that $x \neq x'$.

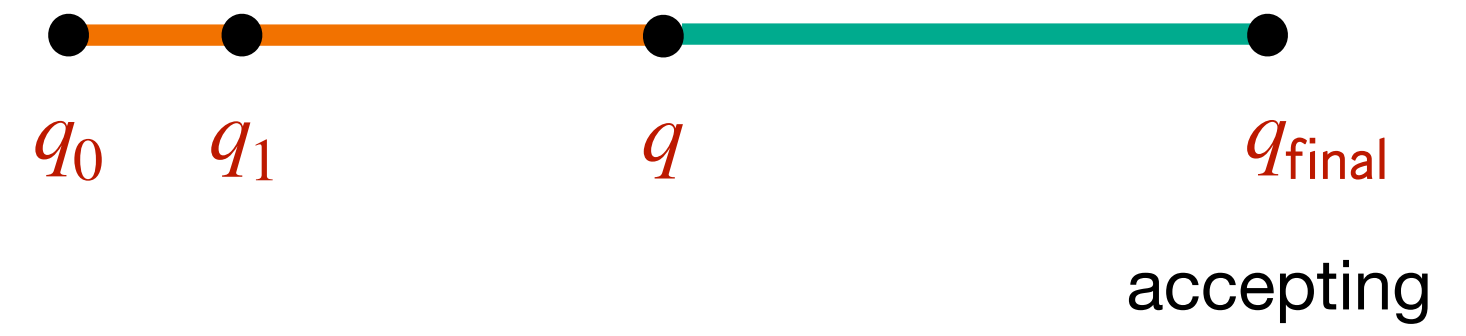
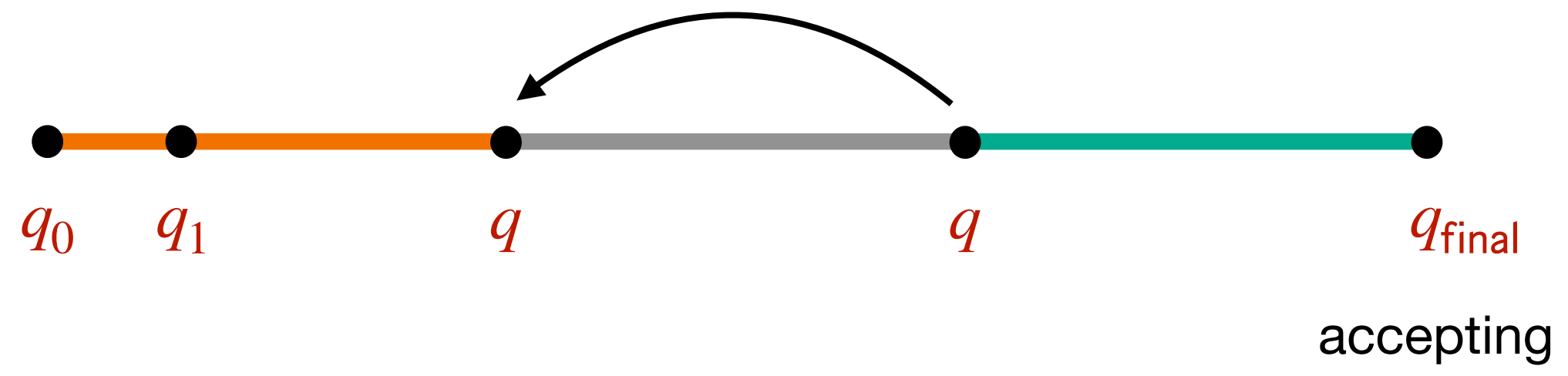
Pumping Lemmas

NFA's run



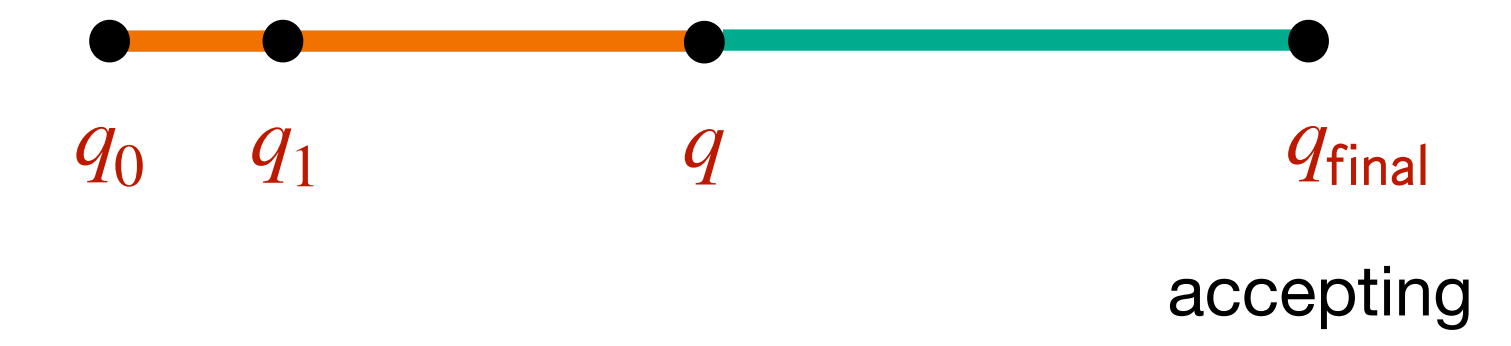
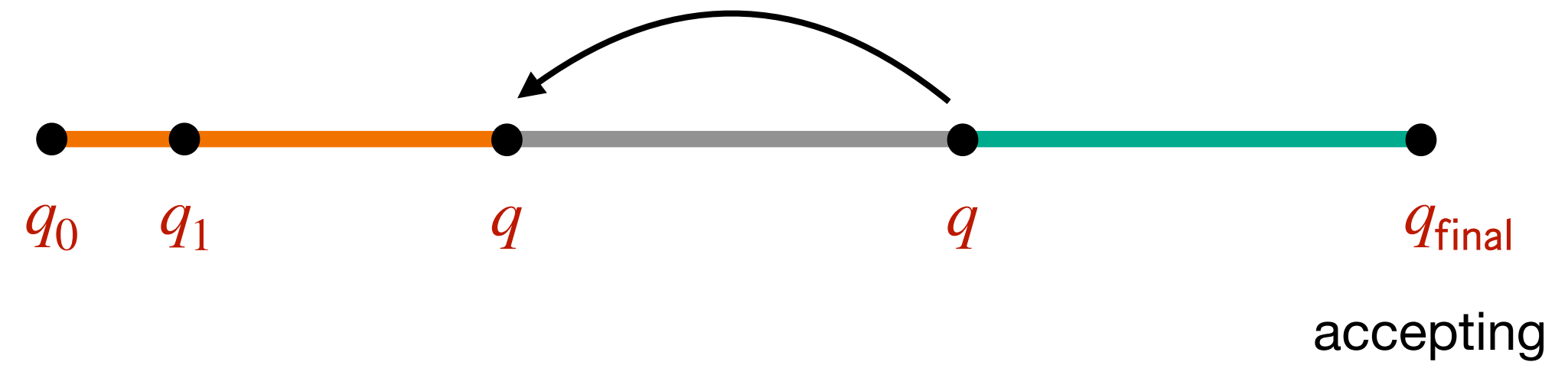
Pumping Lemmas

NFA's run

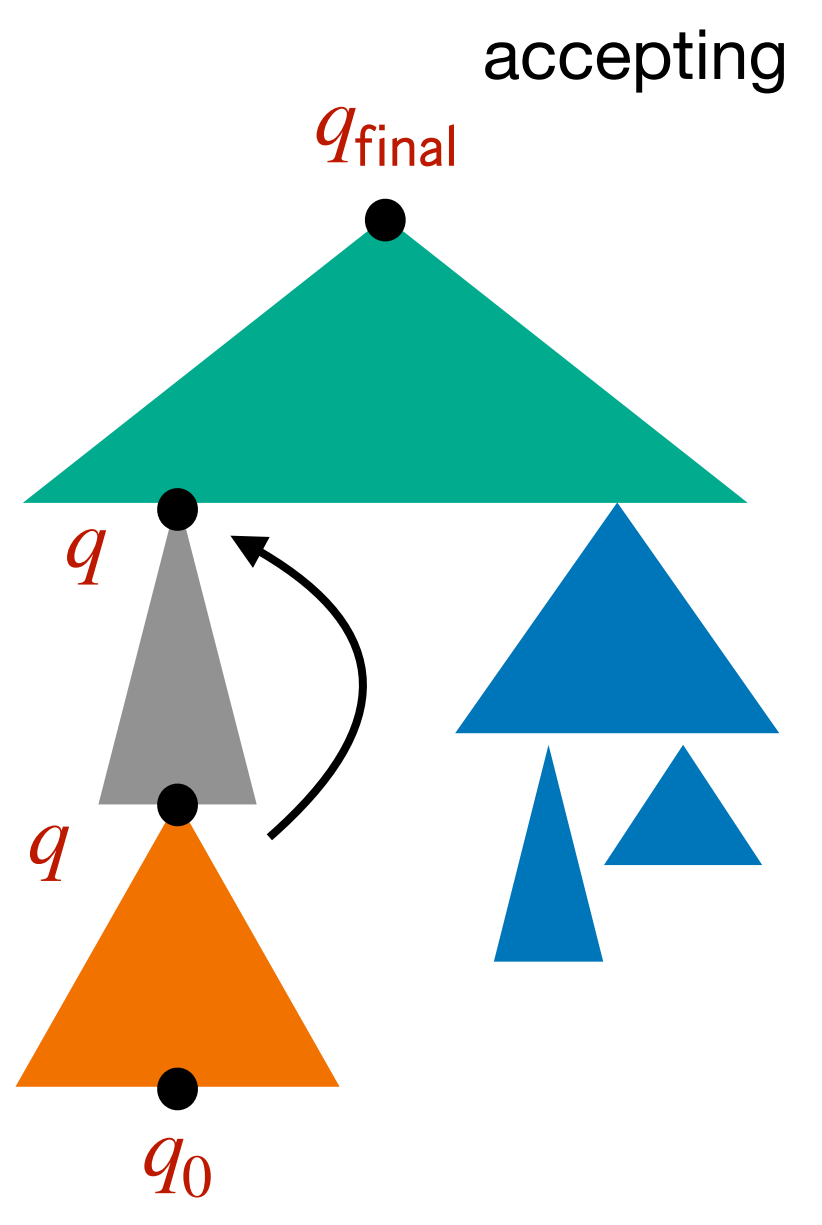


Pumping Lemmas

NFA's run

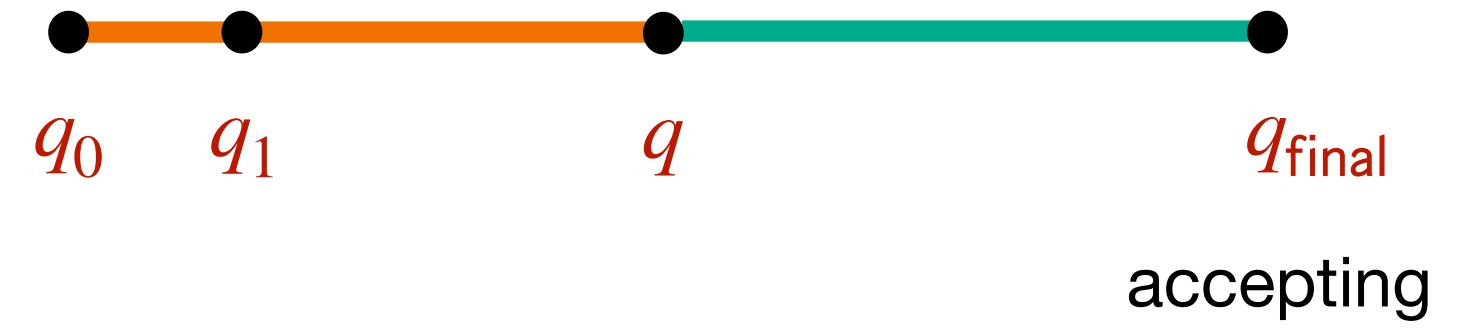
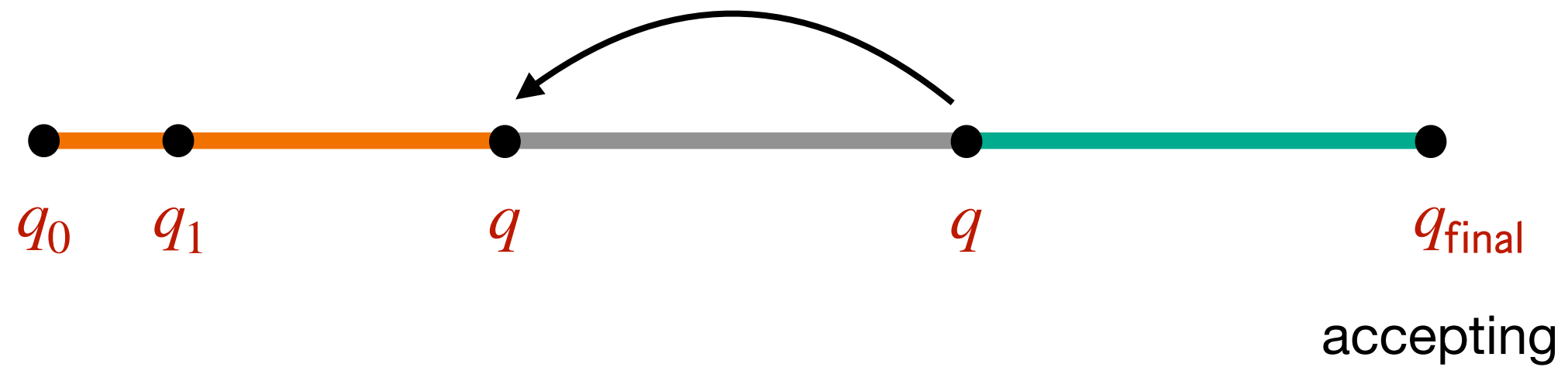


Tree Automaton's run

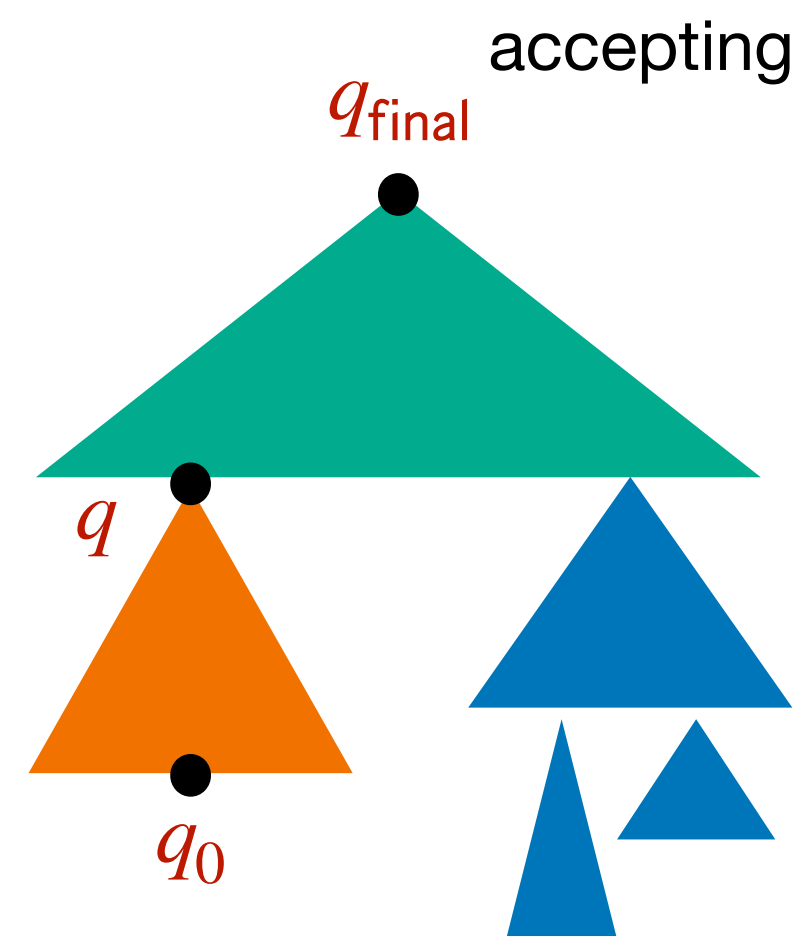
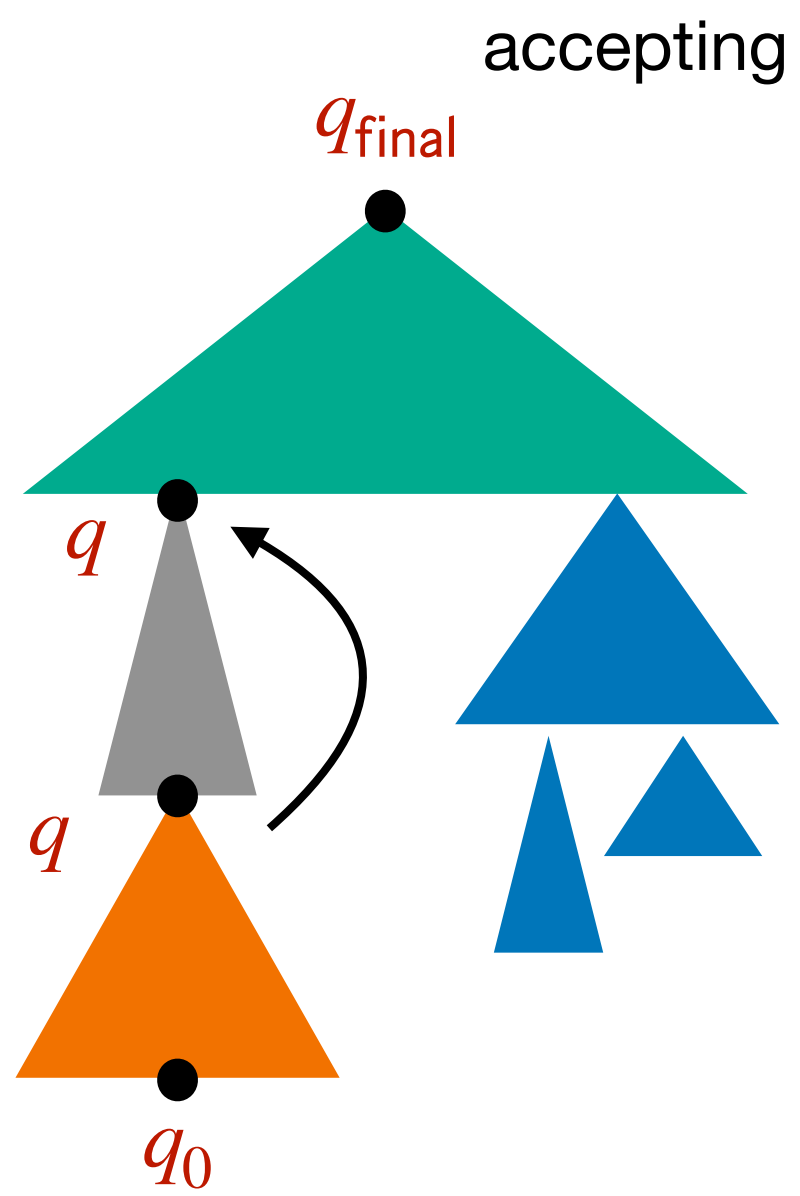


Pumping Lemmas

NFA's run

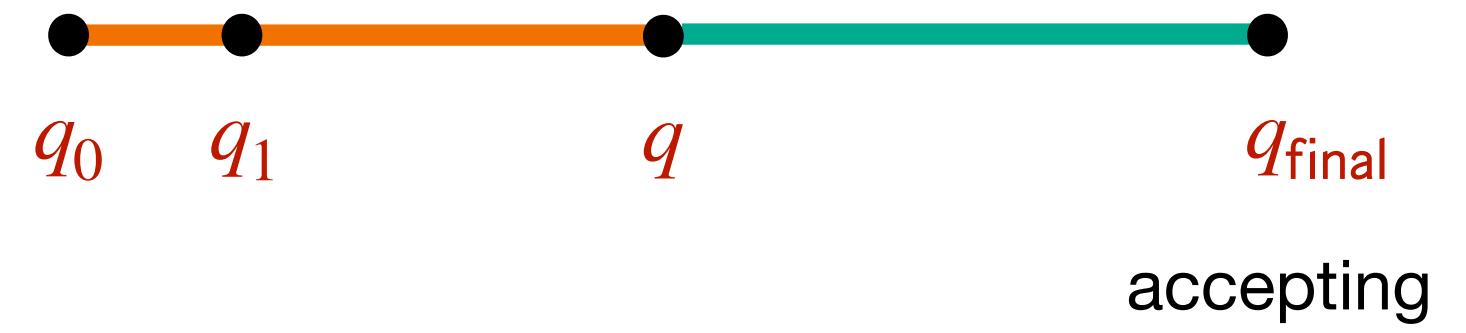
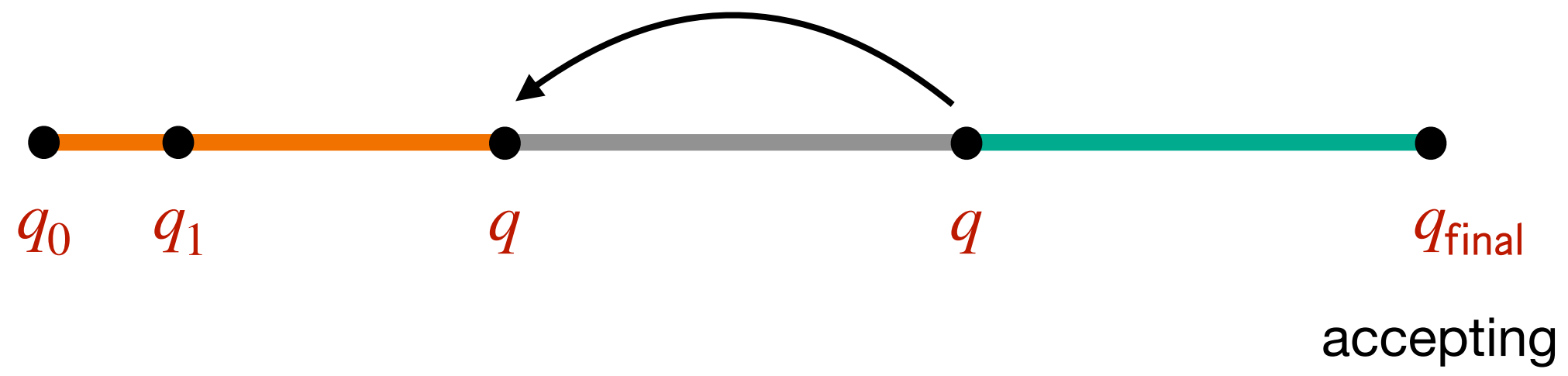


Tree Automaton's run

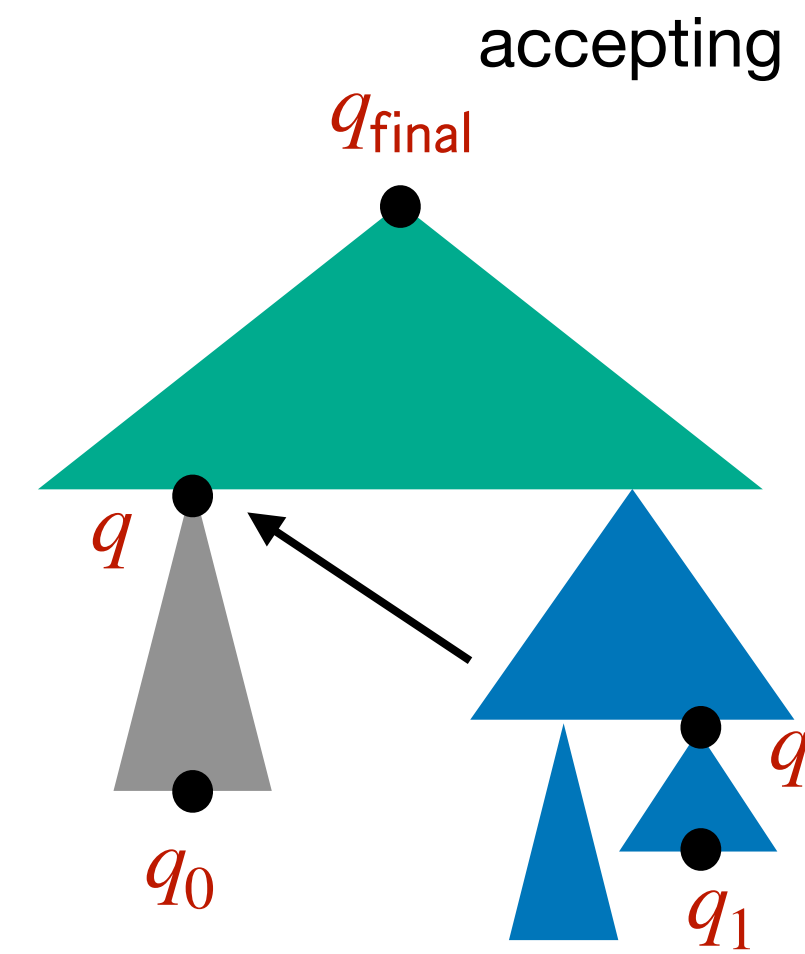
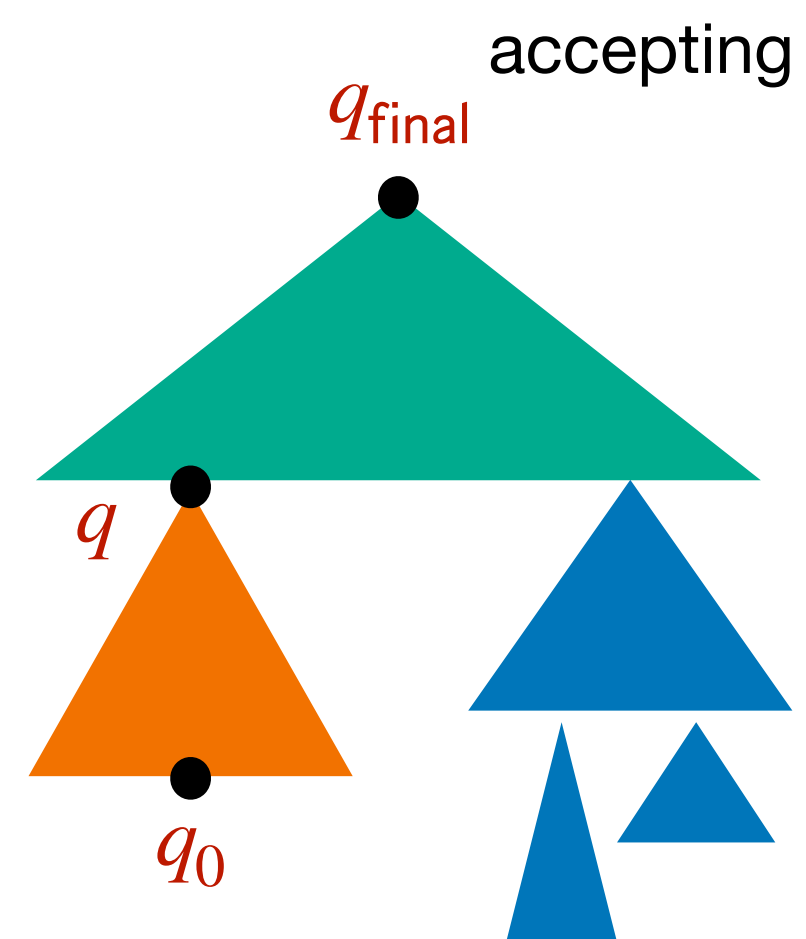
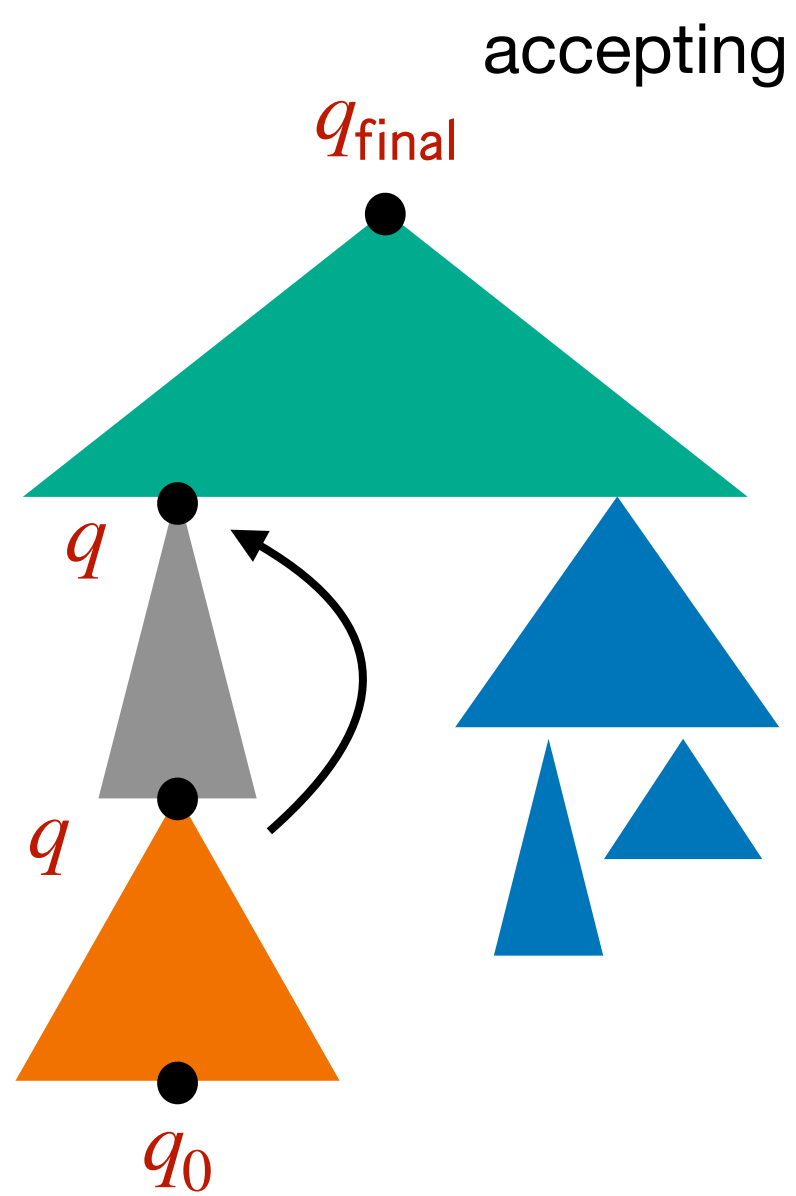


Pumping Lemmas

NFA's run

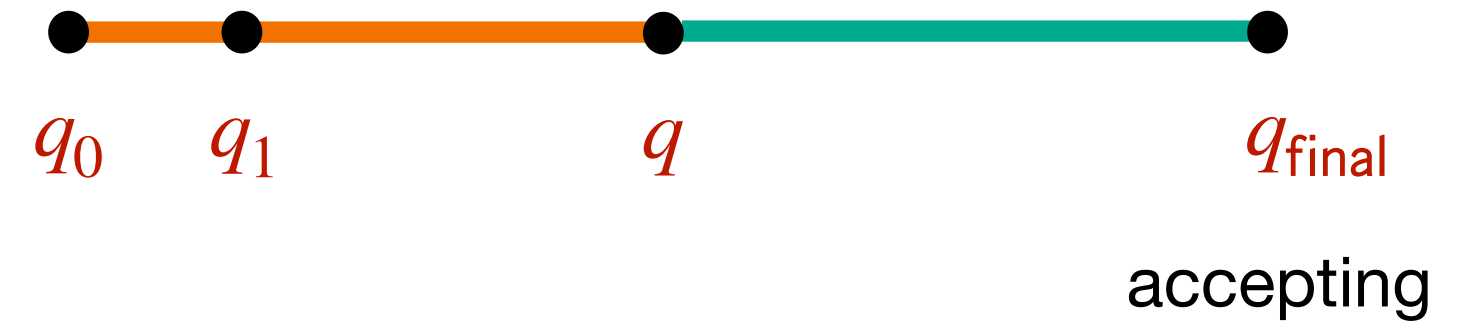
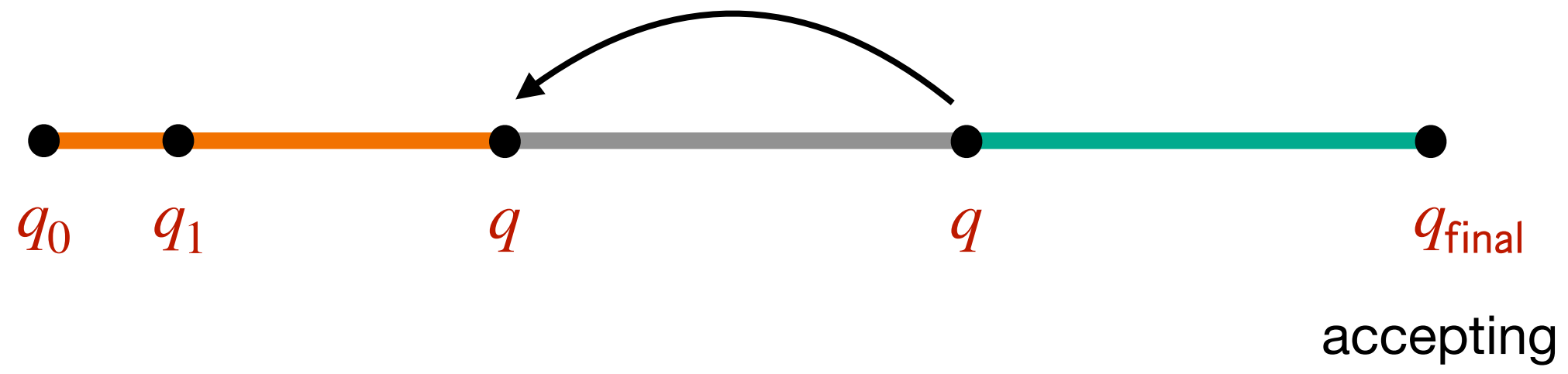


Tree Automaton's run

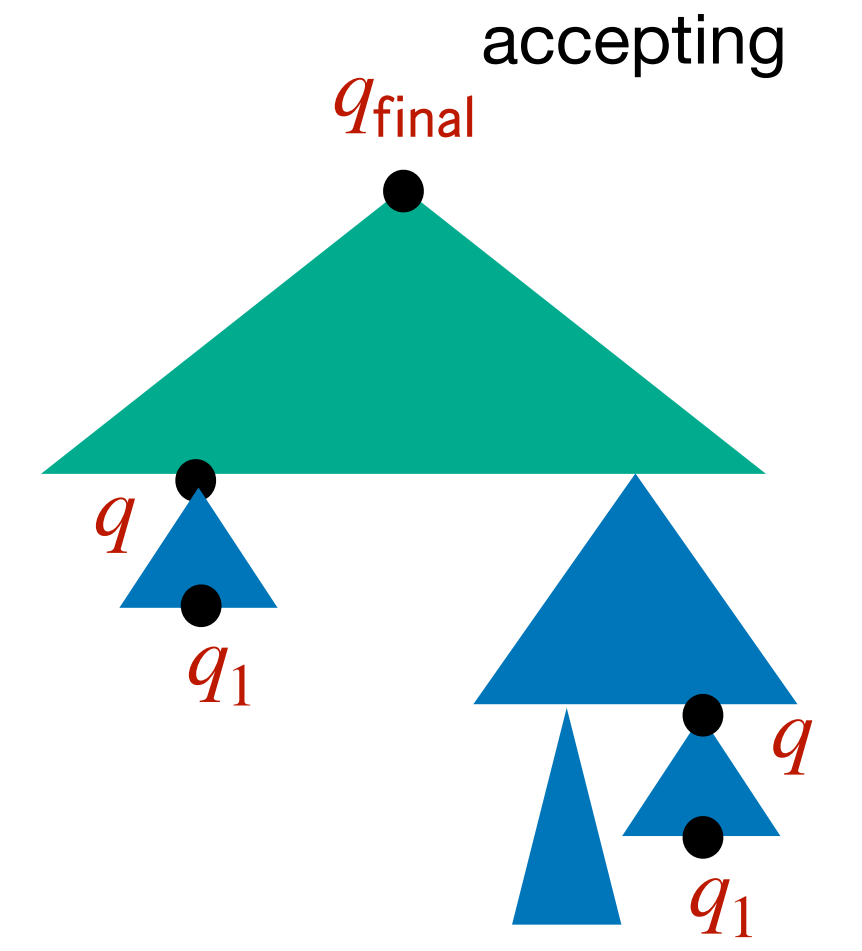
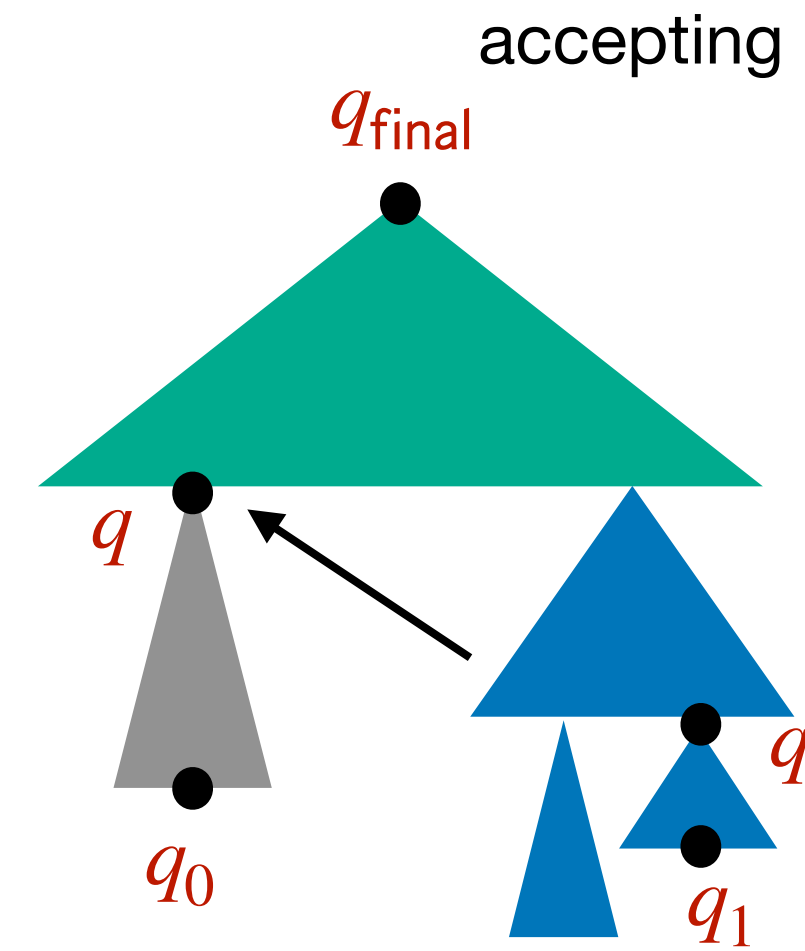
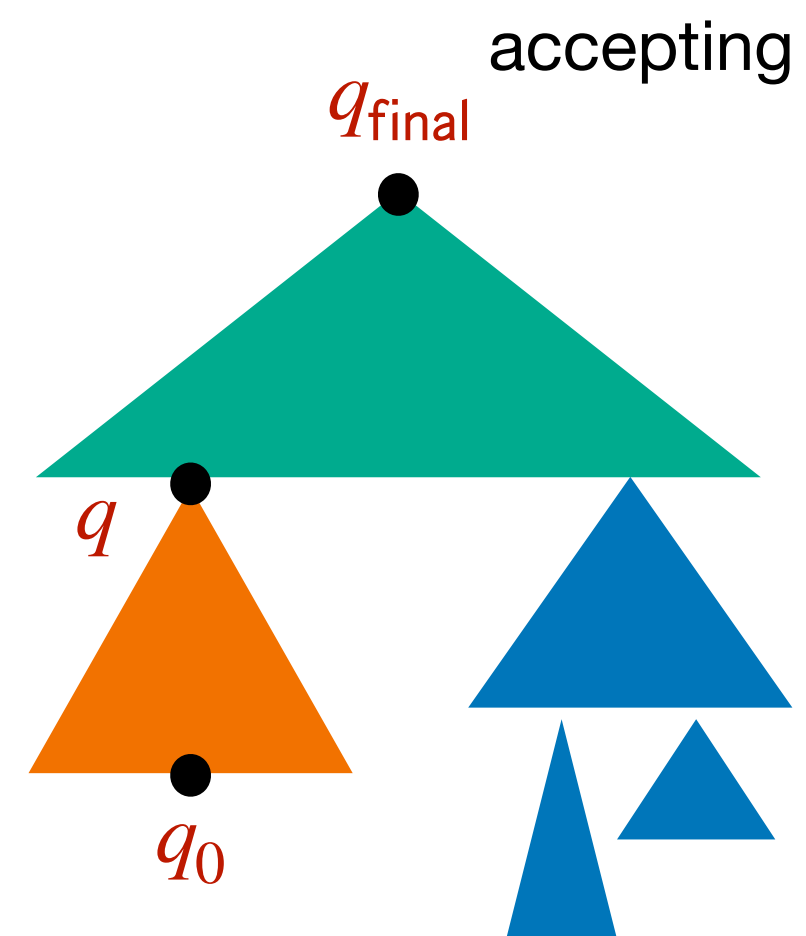
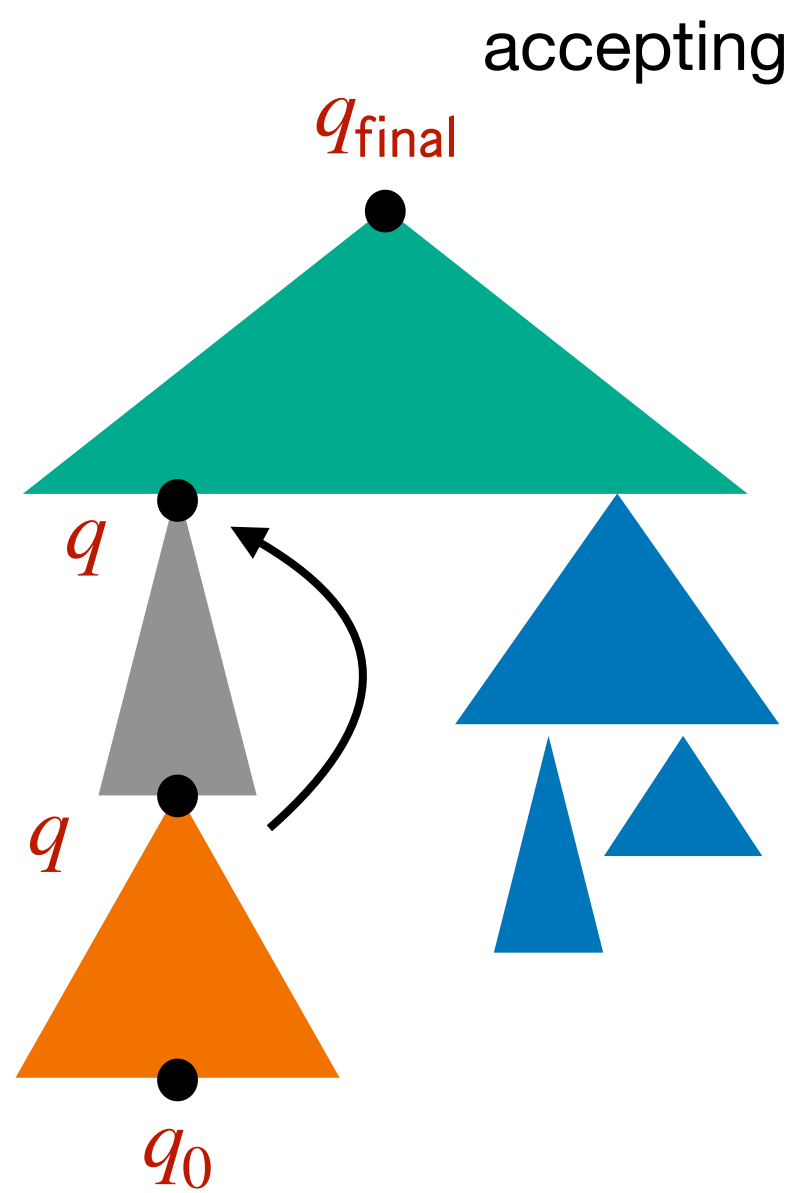


Pumping Lemmas

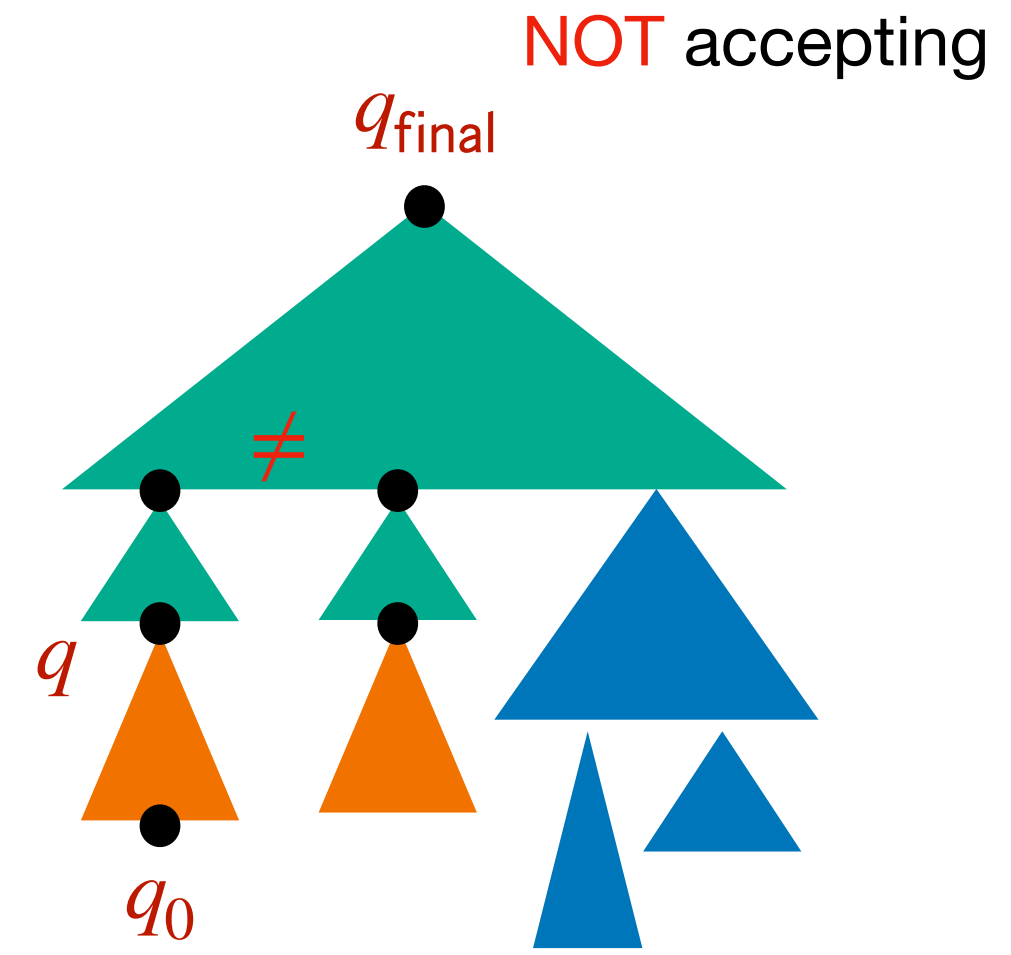
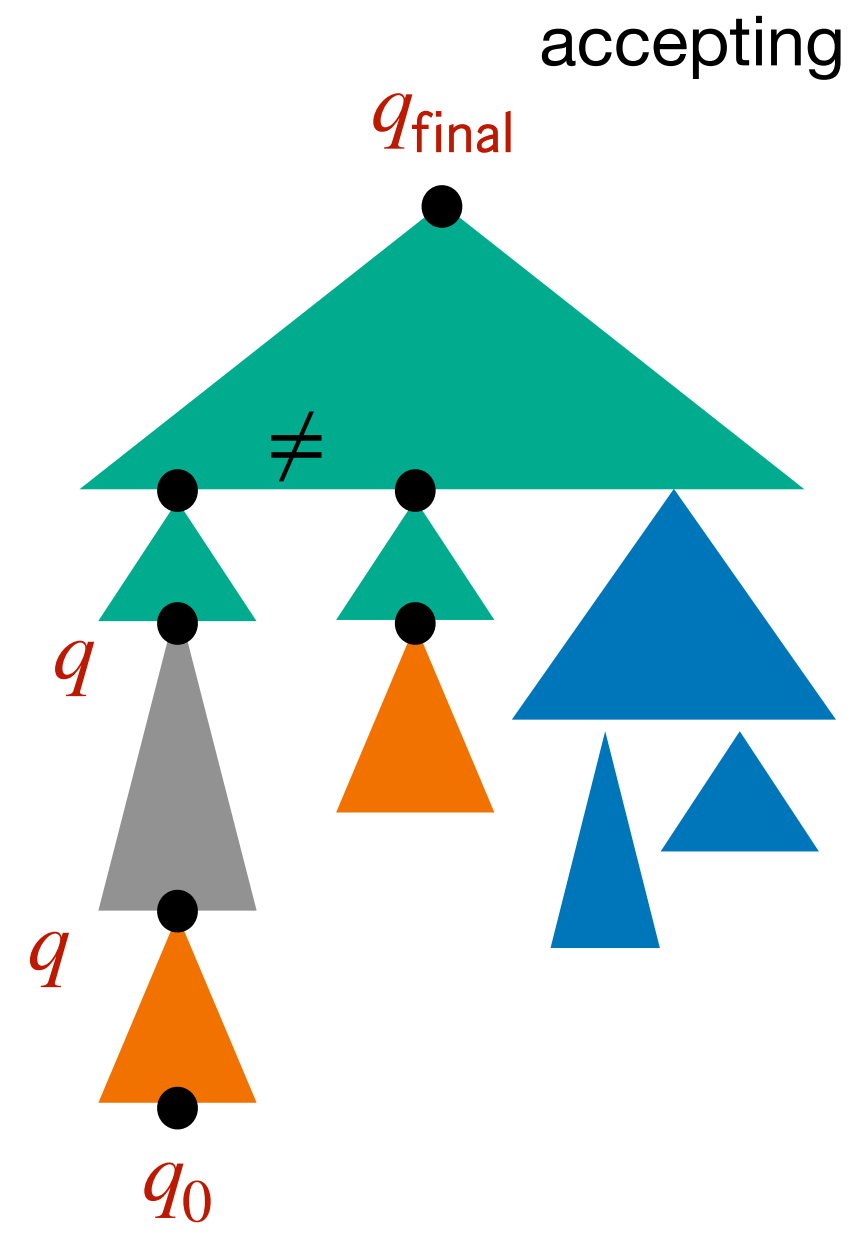
NFA's run



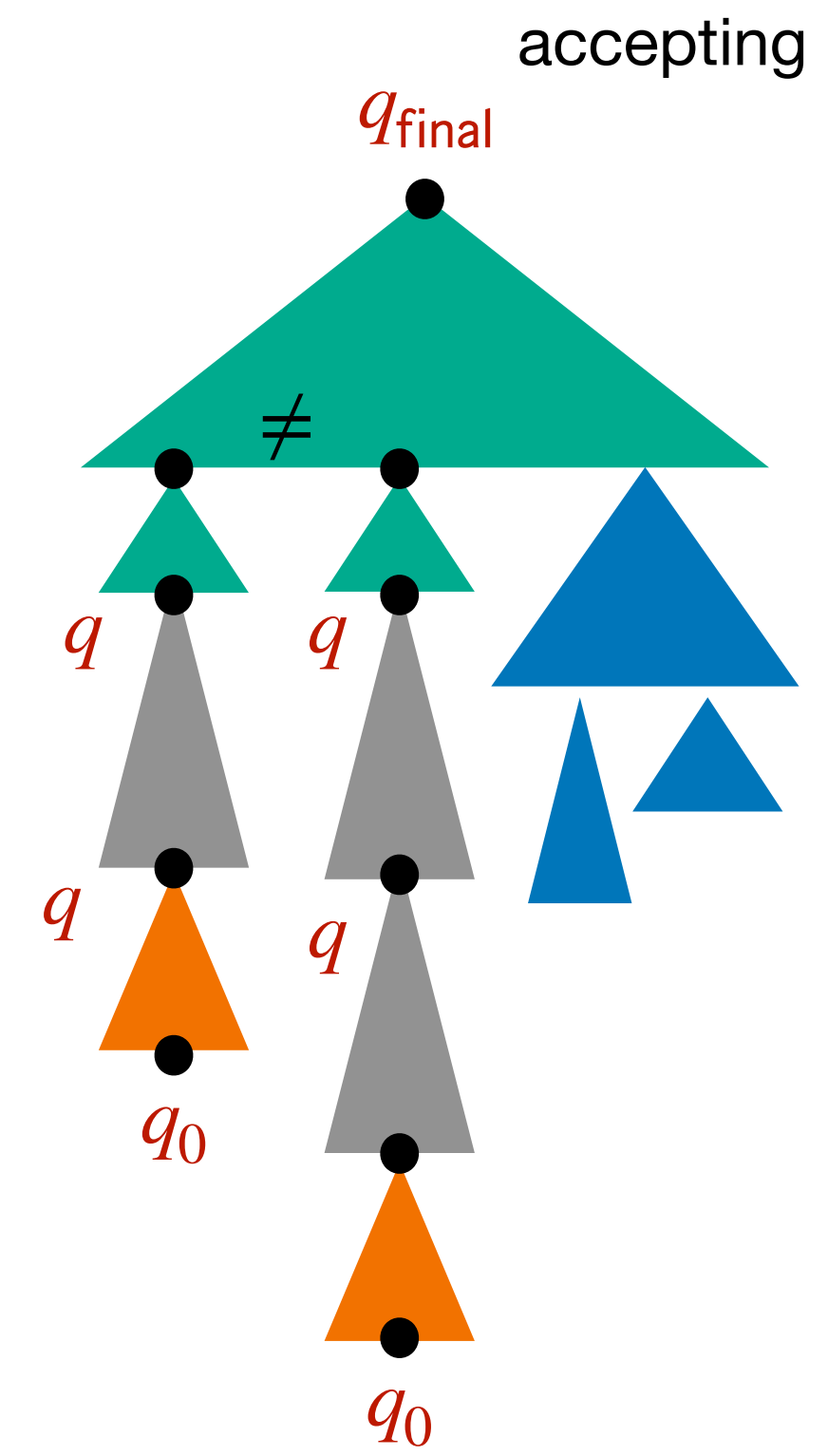
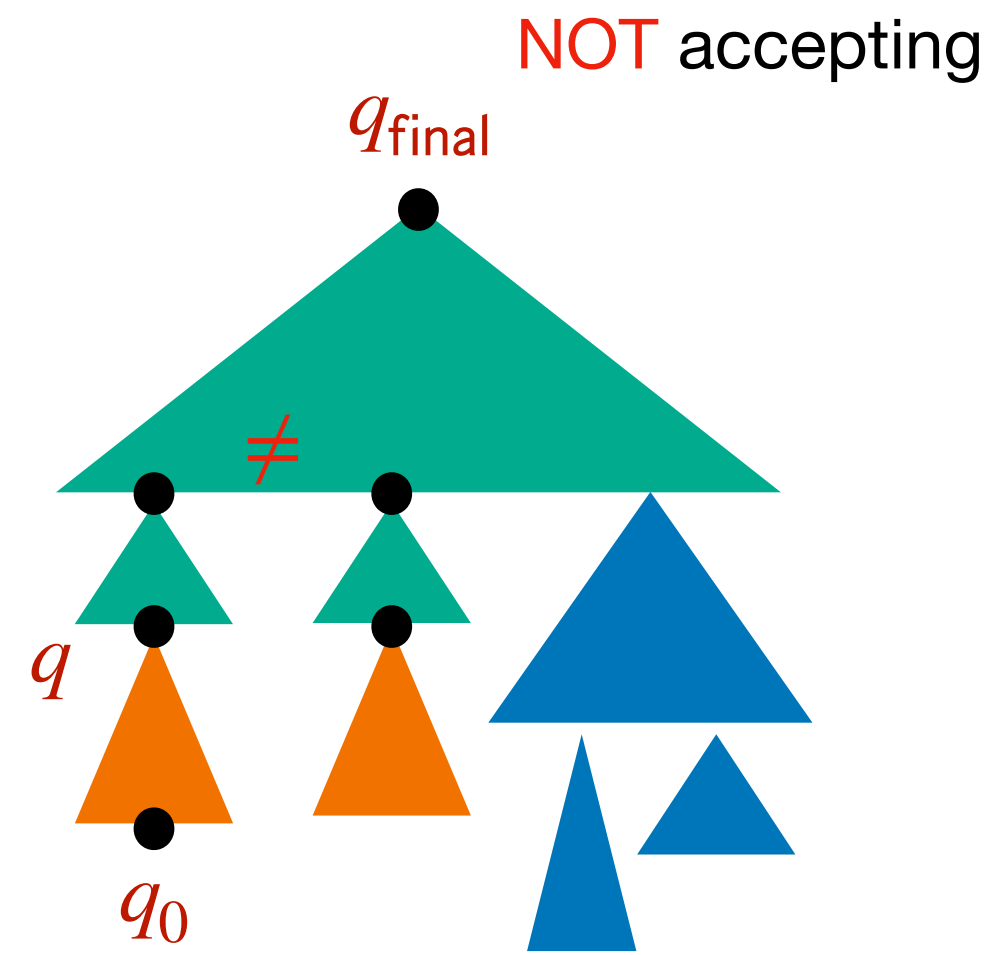
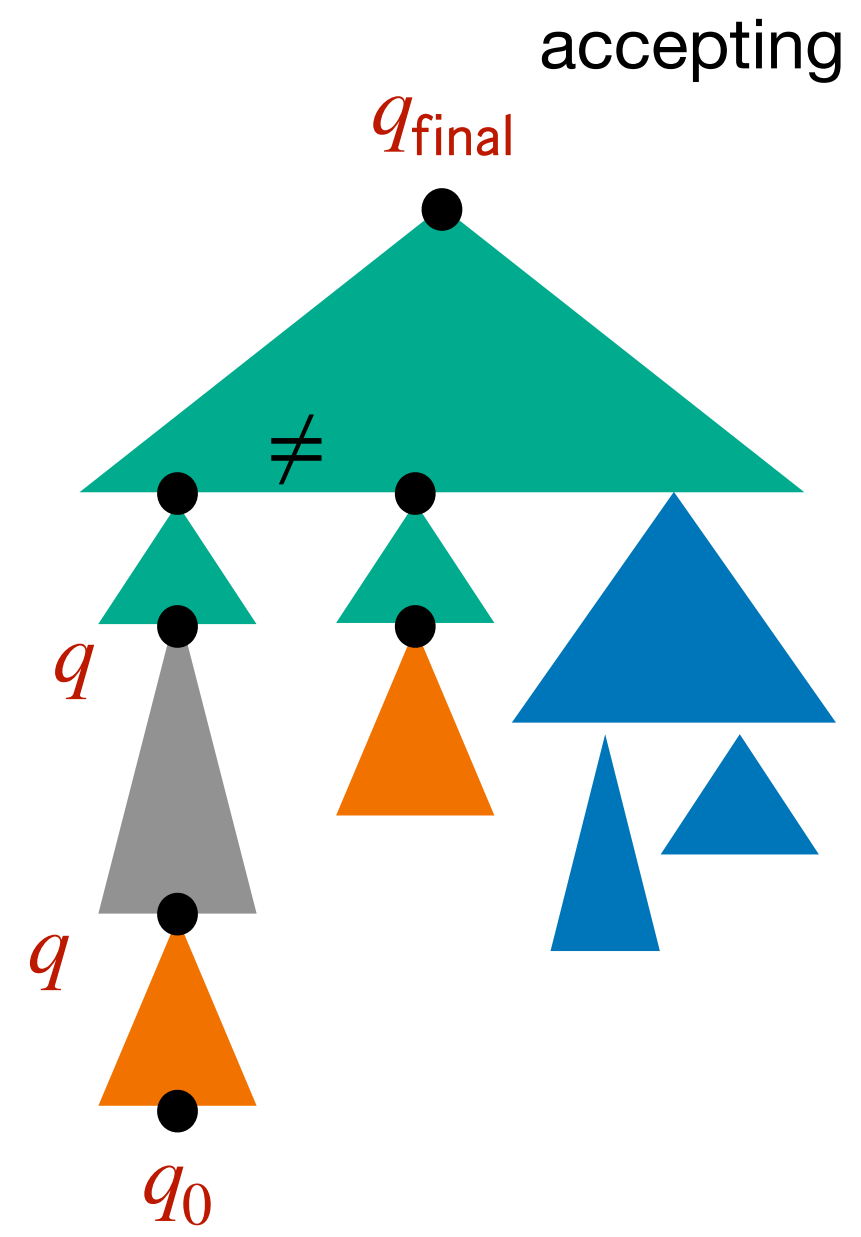
Tree Automaton's run



Constrained Tree Automaton's run



Constrained Tree Automaton's run



Ground Reducibility is EXPTIME-complete

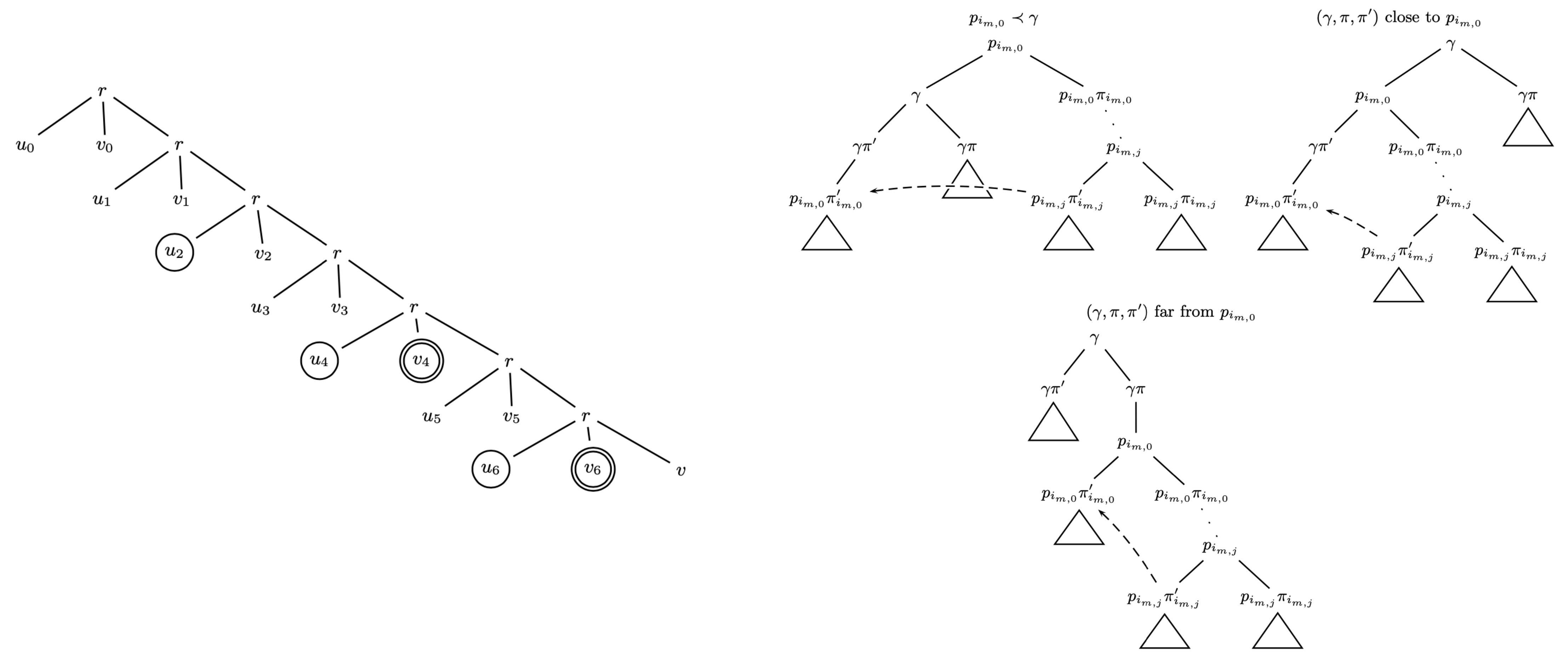


Fig. 8. (γ, π, π') far from $p_{i_m,0}\pi'_{i_m,0}$.

Ground Reducibility is EXPTIME-complete

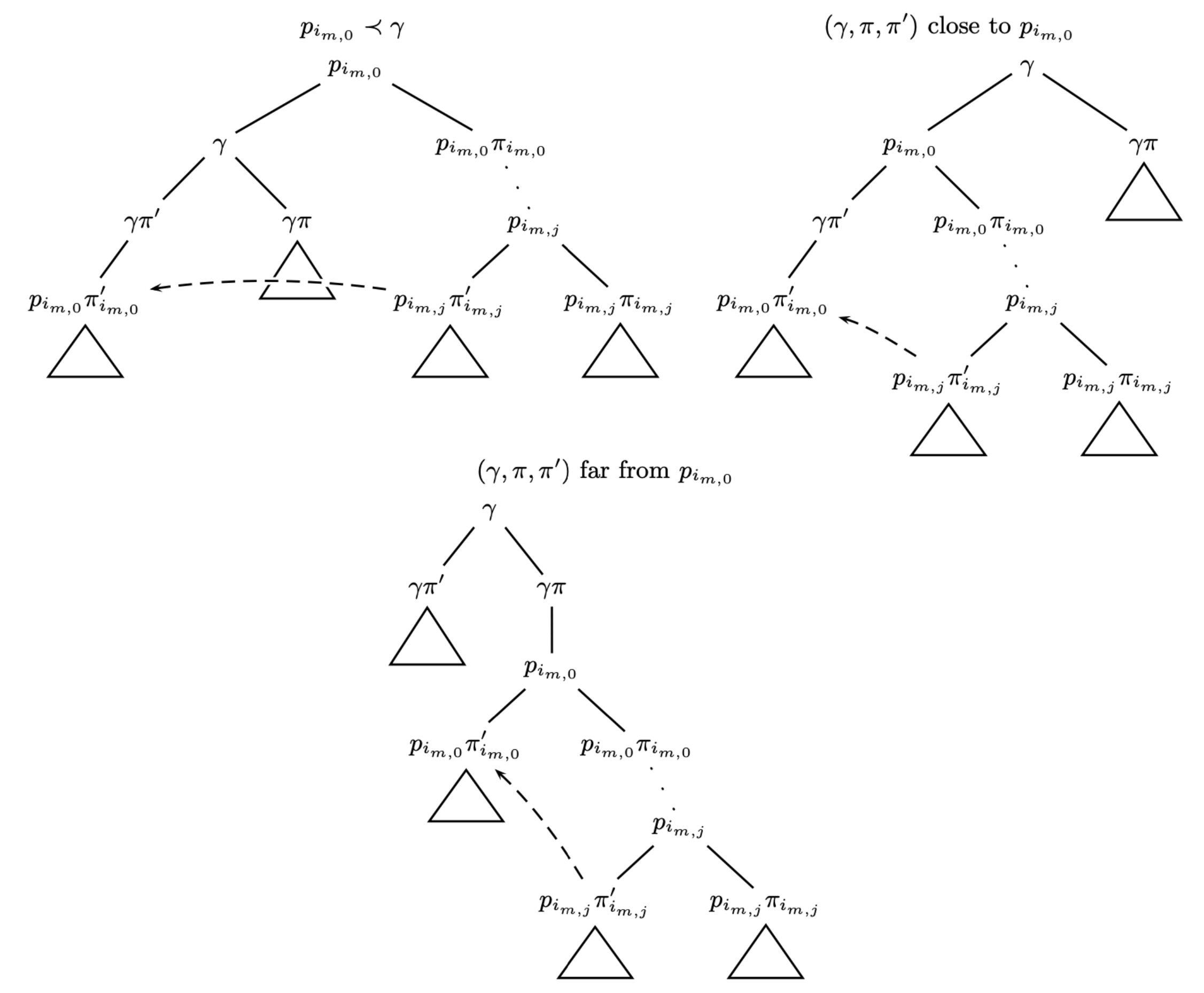
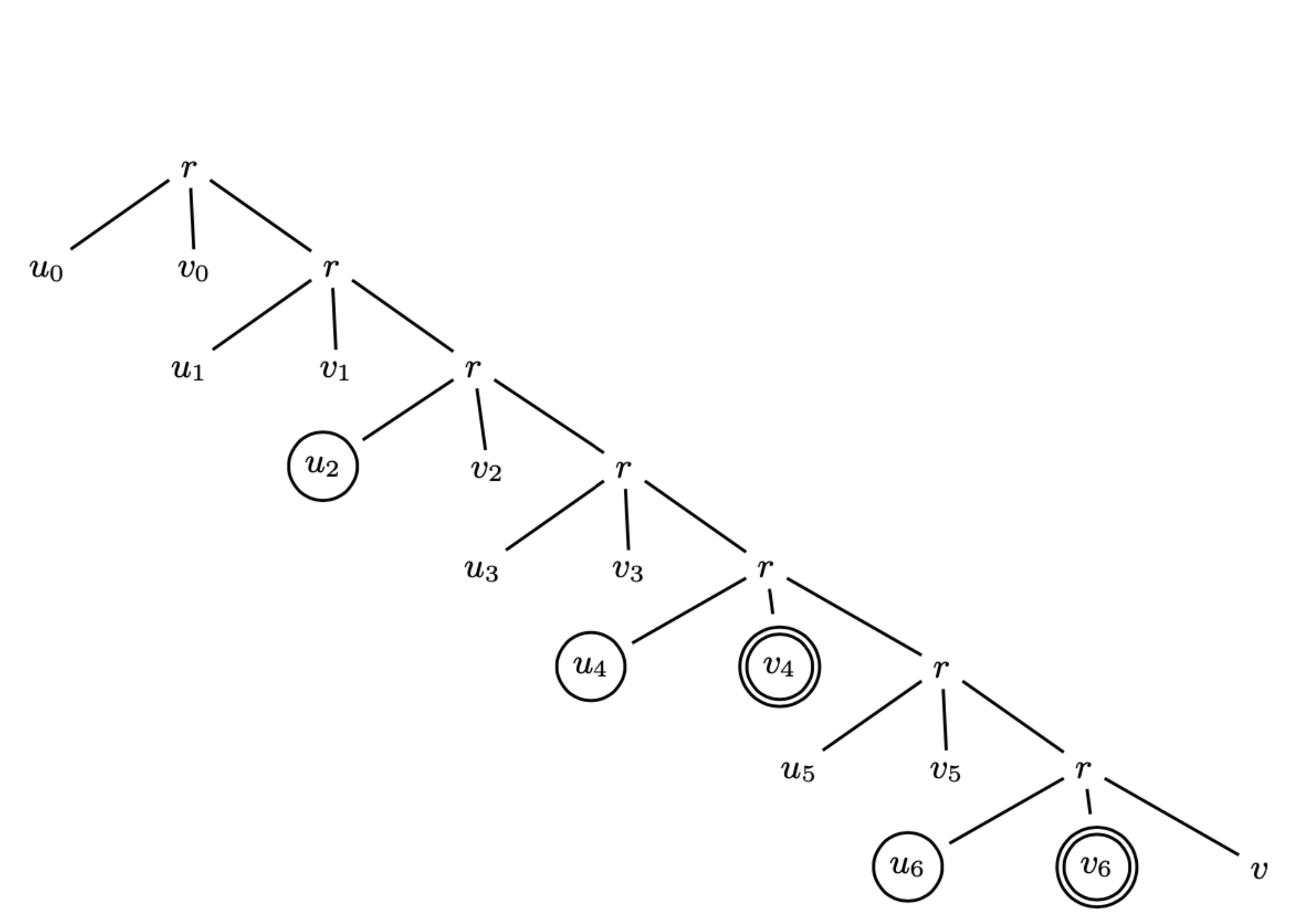
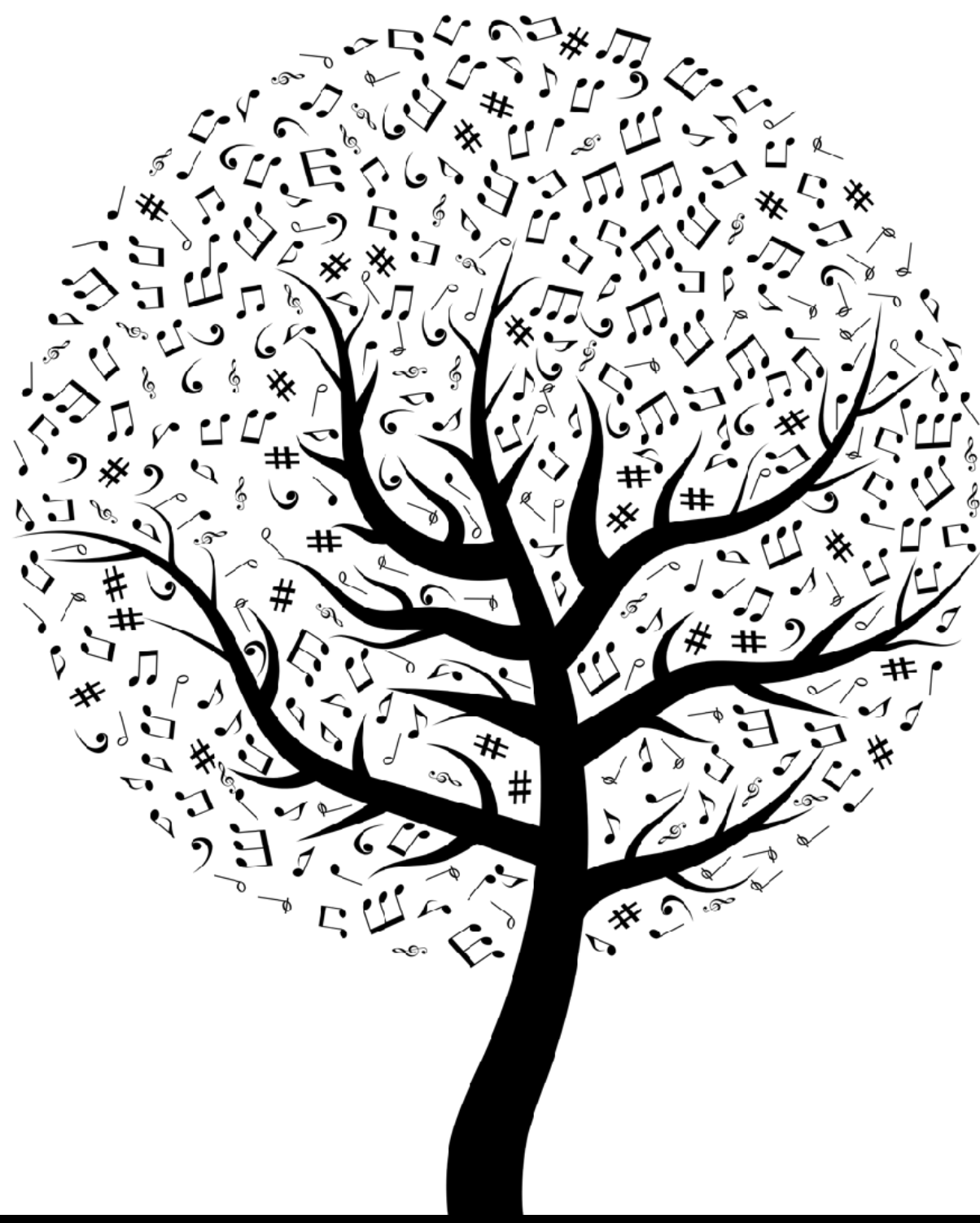


Fig. 8. (γ, π, π') far from $p_{i,m,0} \pi'_{i,m,0}$.

encore vos lemmes de pompage
à la *mord-moi-Inœud* !?





TATA
à la musique





Philippe Rigaux
CNAM

Florent Jacquemard
Inria

Raphaël Fournier-S'niehotta
CNAM

Lydia Rodriguez-de la Nava
PhD (Codex, Inria)

Tiange Zhu
PhD (Polifonia, EU)

post-doc (Collabscore, ANR)

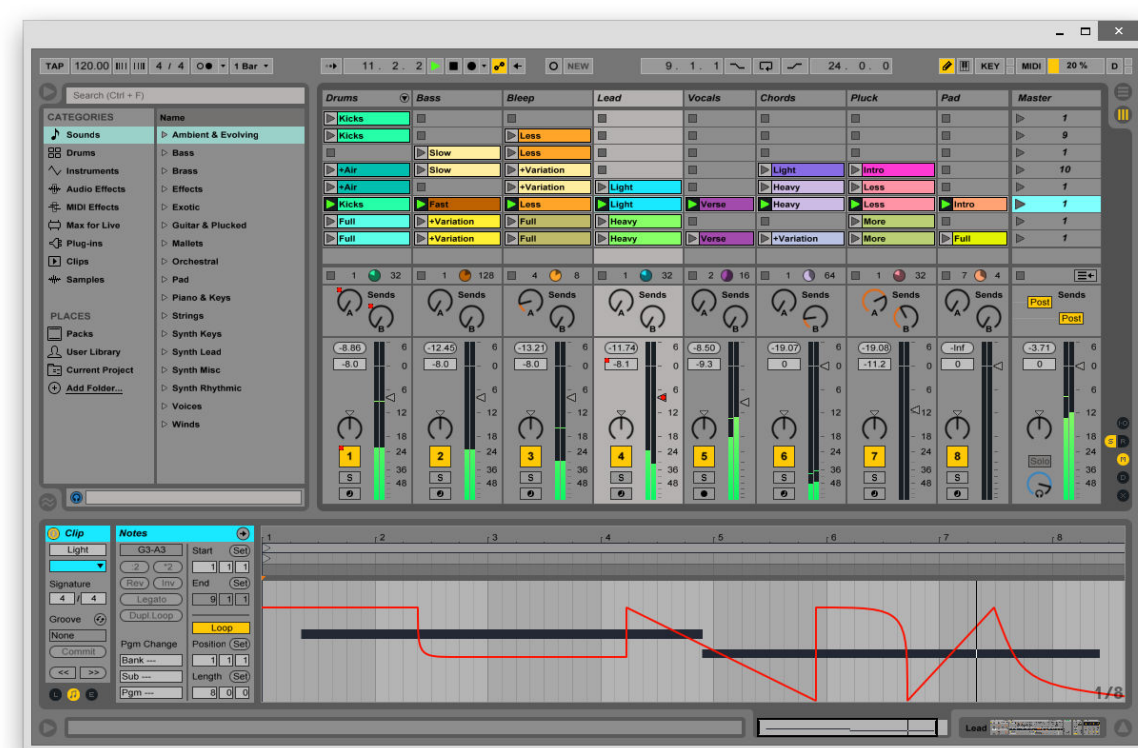
Music Notation Processing

The image displays a musical score for E. Granados' Goyescas, featuring piano and guitar parts. The score is written in a key signature of two flats (B-flat major or D-flat minor) and a 3/4 time signature. The piano part is in the upper staves, and the guitar part is in the lower staves. The score includes various musical notations such as chords, arpeggios, and ornaments. Performance instructions like "cresc. molto" and "appassionato molto" are present. The score is typeset with Lilypond.

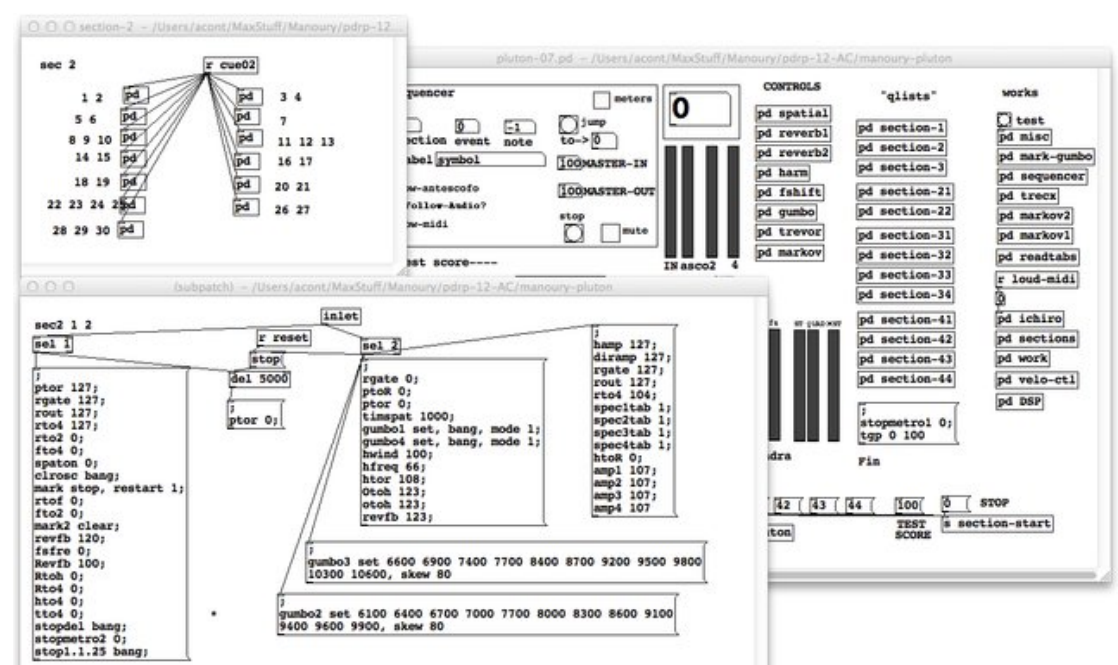
E. Granados' Goyescas
typesetted with Lilypond

Why studying Music Notation Processing?

Western Music Notation = graphical format for music practice, in use since ~1000 years (Guido d'Arezzo)



VS



(digital) music scores, a natural language for

- performers
performance : real-time reading or memoization
- composers
authoring, exchange
- teachers & students
transmission
- editors
access digital score libraries e.g. nkoda.com
- librarians
cultural heritage preservation: e.g. Gallica
- scholars (historians, musicologists...)
research, analysis

Philippe Manoury
Tensio for string quartet and electronics

Philippe Rigaux

CNAM

Florent Jacquemard

Inria

Raphaël Fournier-S'niehotta

CNAM

Lydia Rodriguez-de la Nava

PhD (Codex, Inria)

Tiange Zhu

PhD (Polifonia, EU)

post-doc (Collabscore, ANR)

Music Notation Processing

- Structured music scores models
- Music scores languages
vas-y TATA!
- Search and retrieval
- Similarity metrics
string and tree edit-distances

Applications

- Databases of digital music scores
Cultural heritage preservation [H2020 Polifonia](#)
- Computational Musicology
[neuma.huma-num.fr](#) - [IReMus UMR 8223](#)
- Optical Music Recognition, Crowdsourced correction
[ANR Collabscore](#) - [IRISA](#), [BnF](#), [Royaumont](#)
- Automated Music Transcription
[JSPS 採譜](#) [JAIST](#) & [Nagoya U.](#) [grand Yamaha Music Foundation.](#)

Conversion of a recorded music performance into a music score ~ *speech-to-text* in NLP
 a holy grail in Computer Music since 1970's

articles

Perception of melodies

H. C. Longuet-Higgins

Centre for Research on Perception and Cognition, Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK

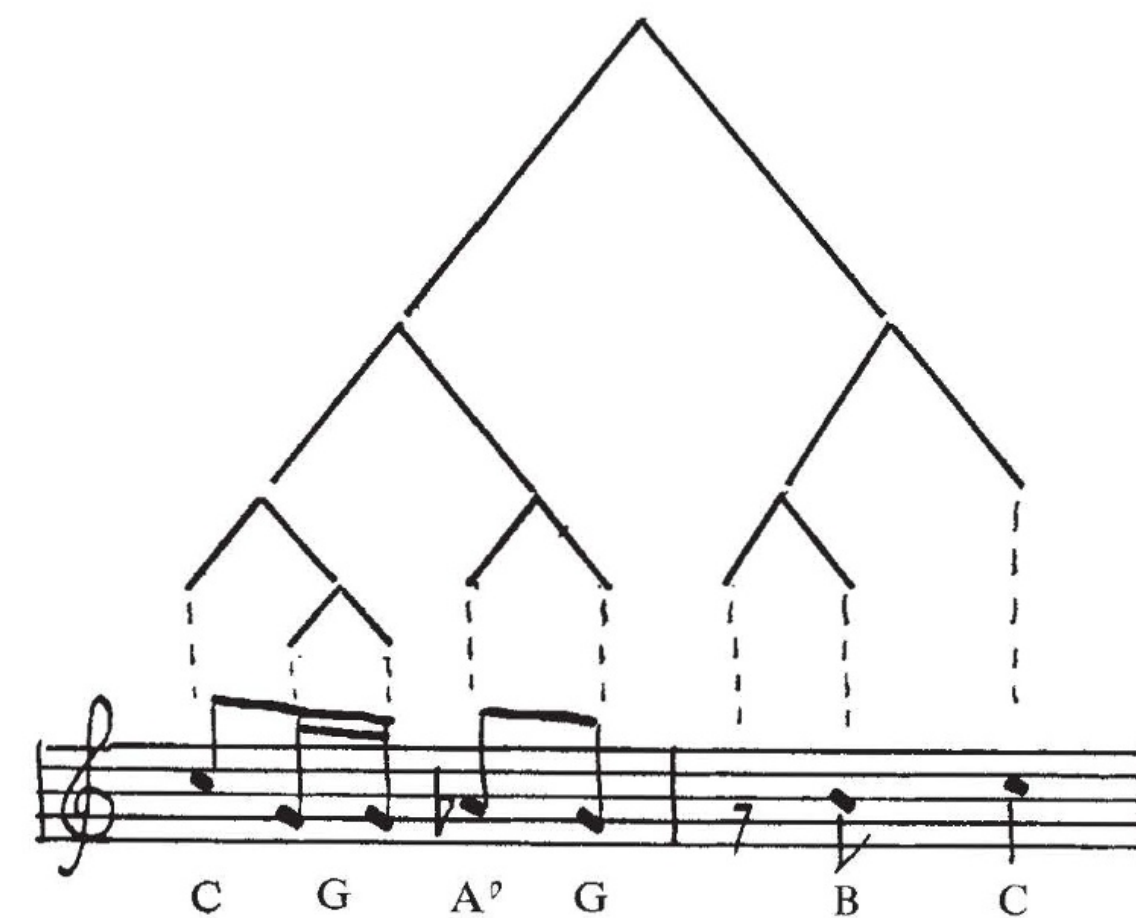
A computer program has been written which will transcribe a live performance of a classical melody into the equivalent of standard musical notation. It is intended to embody, in computational form, a psychological theory of how Western musicians perceive the rhythmic and tonal relationships between the notes of such melodies.

A SEARCHING test of practical musicianship is the 'aural test' in which the subject is required to write down, in standard, musical notation, a melody which he has never heard before. His transcription is not to be construed as a detailed record of the actual performance, which will inevitably be more or less out of time and out of tune, but as an indication of the rhythmic and tonal relations between the individual notes. How the musical listener perceives these relationships is a matter of some interest to the cognitive psychologist. In this paper I outline a theory of the perception of classical Western melodies, and describe a computer program, based on the theory, which displays, as best it can, the rhythmic and tonal relationships between the notes of a melody as played by a human performer on an organ console.

The basic premise of the theory is that in perceiving a melody the listener builds a conceptual structure representing the rhythmic groupings of the notes and the musical intervals between them. It is this structure which he commits to memory, and which subsequently enables him to recognise the tune, and

to reproduce it in sound or in writing if he happens to be a skilled musician. A second premise is that much can be learned about the structural relationships in any ordinary piece of music from a study of its orthographic representation. Take, for example, the musical cliché notated in Fig. 1.

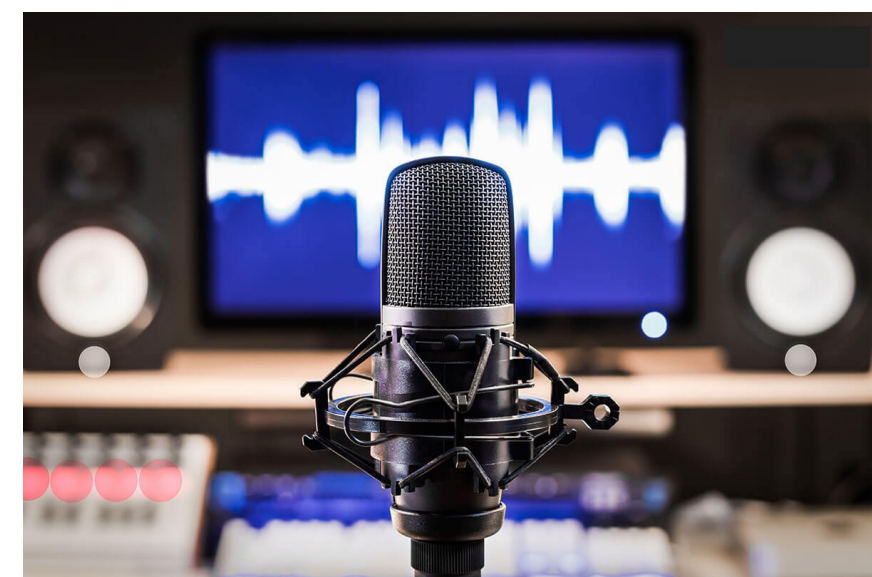
Fig. 1



Automated Music Transcription today

Conversion of a recorded music performance into a music score

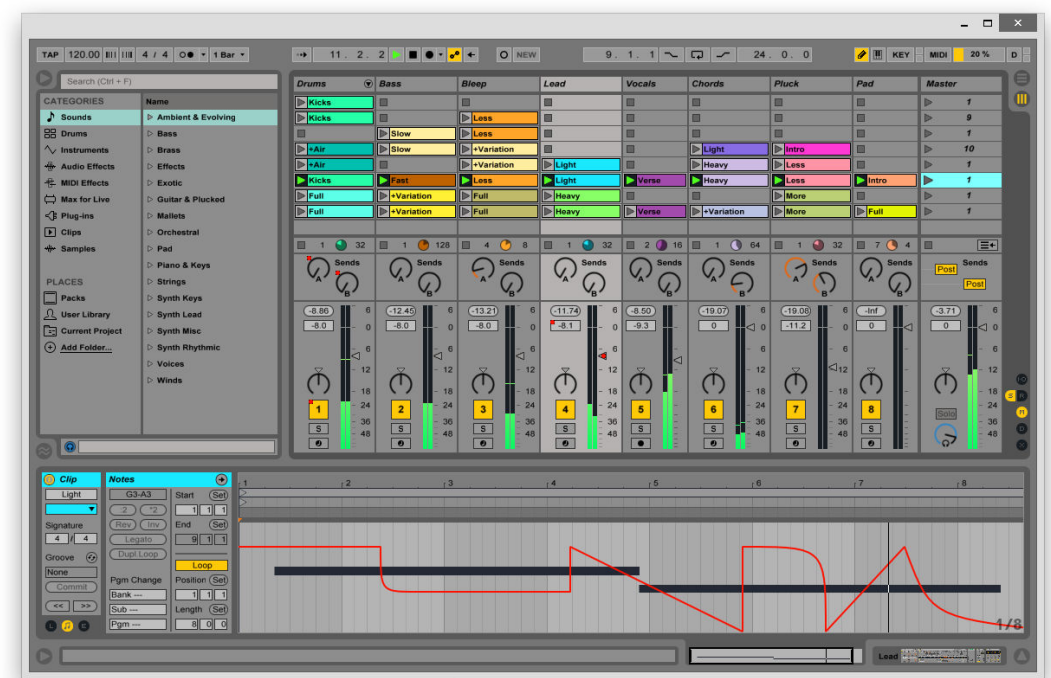
Audio recording



MIDI device
(score edition)



Algorithmic composition
DAW



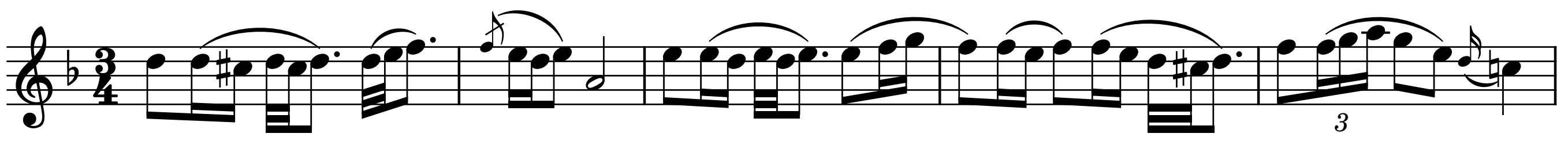
audio Music Information Retrieval

- fundamental freq. estimation
- onset detection
- beat tracking ...



symbolic Music Information Retrieval

- rhythm quantization
- tempo tracking
- score engraving...

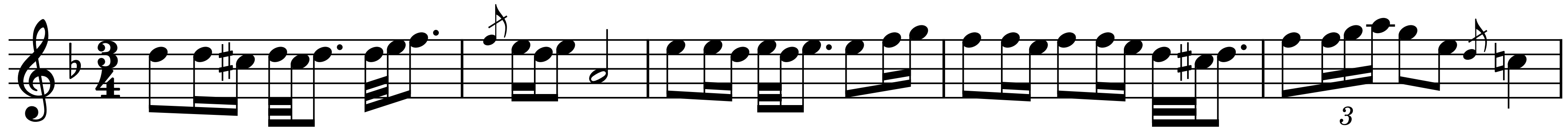


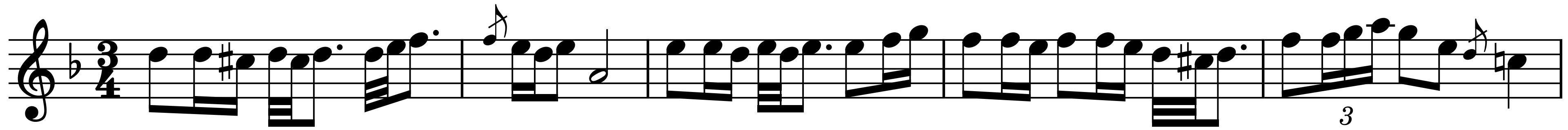
piano roll (MIDI file)

- unquantized onsets, durations
- quantized pitches

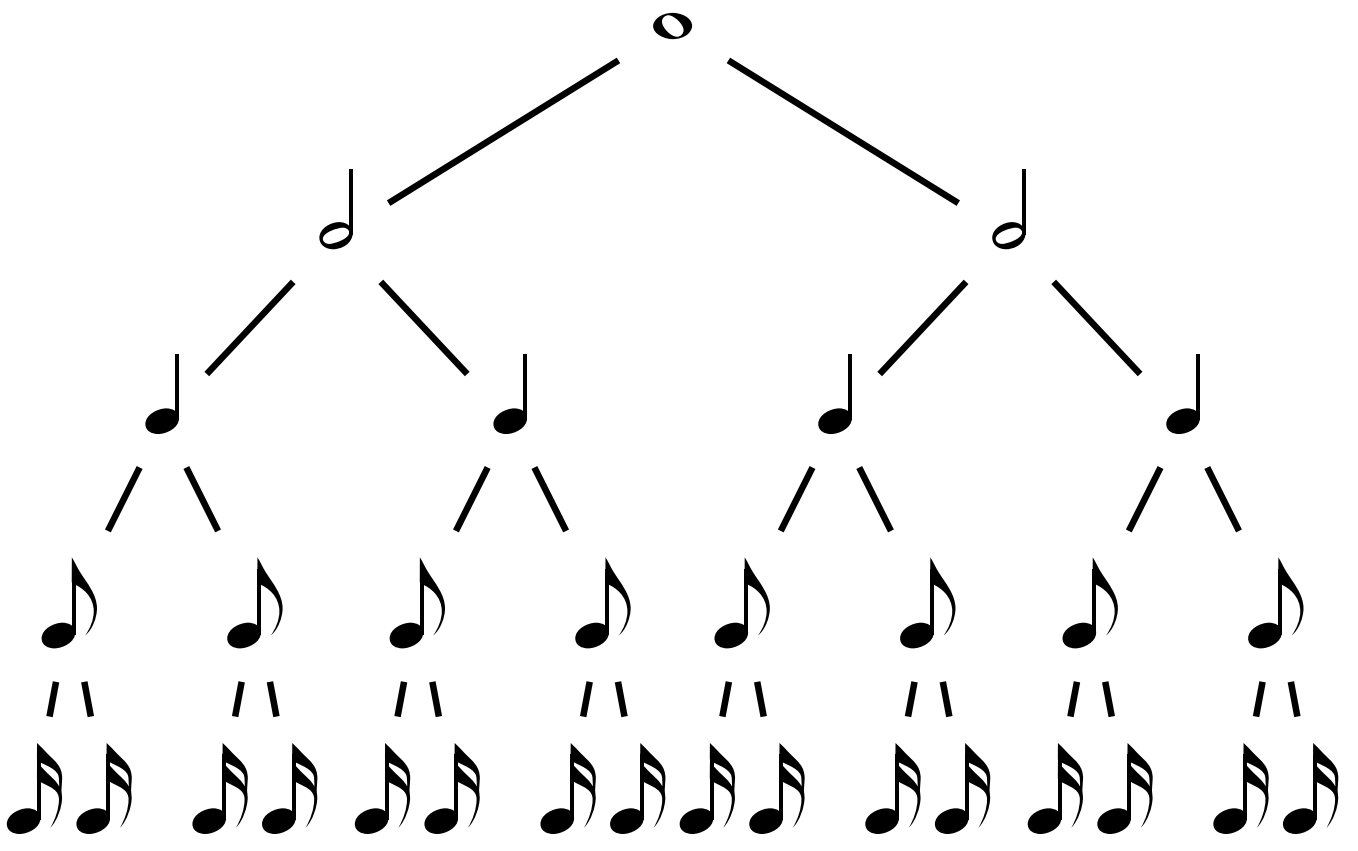
music score

(XML file)

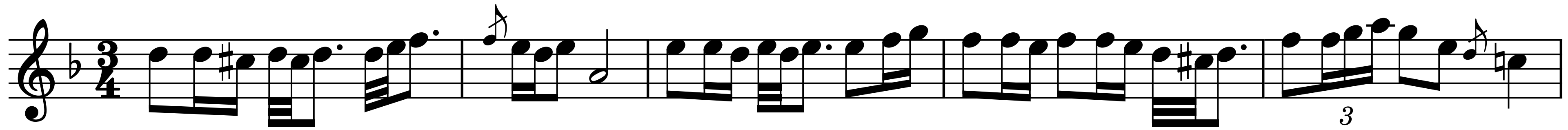




hierarchical
note
durations



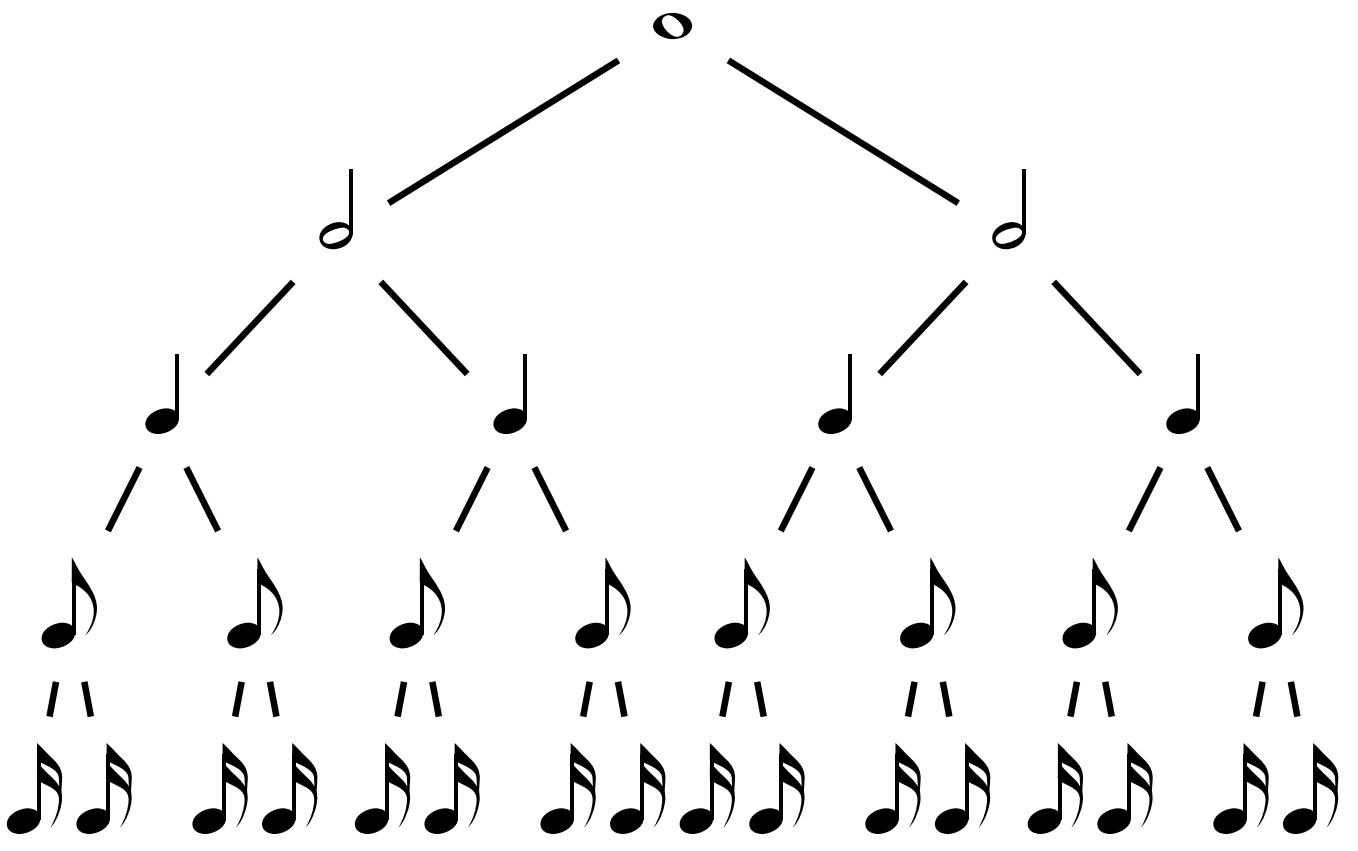
beamed



unbeamed



hierarchical
note
durations



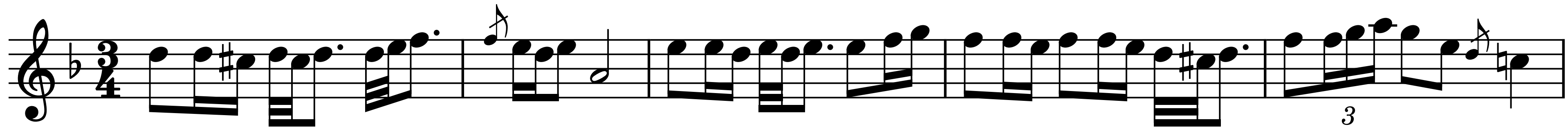
beamed

Musical notation for the 'beamed' version of the Polonaise in D minor. The piece is in 3/4 time and D minor. The notation shows a melodic line with eighth and sixteenth notes, many of which are beamed together. A triplet of eighth notes is indicated by a '3' below the notes in the final measure.

unbeamed

Musical notation for the 'unbeamed' version of the Polonaise in D minor. The notation shows the same melodic line as the beamed version, but with individual notes separated by stems. A triplet of eighth notes is indicated by a bracket and a '3' below the notes in the final measure.

beamed



unbeamed



grouping notes with measure bars and beams

- eases readability (player reads in a real-time context)
- highlight the metric structure hierarchy of strong / weak beats

Polonaise in D minor from Notebook for Anna Magdalena Bach BWV Anh II 128

metric structure

bar	1			2		3		4		5									
beat	1.1	1.2	1.3	2.1	2.2	2.3	3.1	3.2	3.3	4.1	4.2	4.3	5.1	5.2	5.3				
subbeat	1.1.1	1.1.2		2.1.1	2.1.2		3.1.1	3.1.2	3.3.1	3.3.2	4.1.1	4.1.2	4.2.1	4.2.1		5.1.1	5.1.2	5.2.1	5.2.2

beamed

unbeamed

grouping notes with measure bars and beams

- eases readability (player reads in a real-time context)
- highlight the metric structure hierarchy of strong / weak beats

Common Western Music Notation

Polonaise in D minor from Notebook for Anna Magdalena Bach BWV Anh II 128

metric structure

bar	1	2	3	4	5																			
beat	1.1	1.2	1.3	2.1	2.2	2.3	3.1	3.2	3.3	4.1	4.2	4.3	5.1	5.2	5.3									
subbeat	1.1.1	1.1.2		2.1.1	2.1.2		3.1.1	3.1.2		3.3.1	3.3.2		4.1.1	4.1.2	4.2.1	4.2.1		5.1.1	5.1.2		5.2.1	5.2.2		

beamed

A musical staff in treble clef with a key signature of one flat and a 3/4 time signature. The notes are beamed together in groups corresponding to the metric structure above. A triplet of eighth notes is marked with a '3' below it in the fifth bar.

unbeamed

The same musical staff as above, but the notes are not beamed together. A triplet of eighth notes is marked with a '3' below it in the fifth bar.

Common Western Music Notation

Polonaise in D minor from Notebook for Anna Magdalena Bach BWV Anh II 128

metric
structure

bar	1			2		3		4		5											
beat	1.1	1.2	1.3	2.1	2.2	2.3	3.1	3.2	3.3	4.1	4.2	4.3	5.1	5.2	5.3						
subbeat	1.1.1	1.1.2		2.1.1	2.1.2		3.1.1	3.1.2		3.3.1	3.3.2		4.1.1	4.1.2	4.2.1	4.2.1	5.1.1	5.1.2	5.2.1	5.2.2	

Common Western Music Notation

Polonaise in D minor from Notebook for Anna Magdalena Bach BWV Anh II 128

metric
structure

bar	1	2		3		4			5					
beat	1.1	1.2	1.3	2.1	2.2 2.3	3.1	3.2	3.3	4.1	4.2	4.3	5.1	5.2	5.3
subbeat	1.1.1 1.1.2			2.1.1 2.1.2		3.1.1 3.1.2		3.3.1 3.3.2	4.1.1 4.1.2 4.2.1 4.2.1			5.1.1 5.1.2	5.2.1 5.2.2	

The first staff of music shows the beginning of the piece in 3/4 time. It consists of five bars. The first bar contains a quarter note (F4), an eighth note (F#4), and a quarter note (G4). The second bar contains a quarter note (A4), an eighth note (Bb4), and a quarter note (B4). The third bar contains a quarter note (C5), an eighth note (B4), and a quarter note (B4). The fourth bar contains a quarter note (A4), an eighth note (G4), and a quarter note (G4). The fifth bar contains a quarter note (F#4), an eighth note (F4), and a triplet of eighth notes (E4, D4, C4).

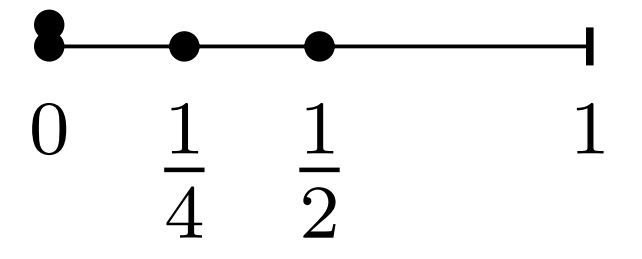
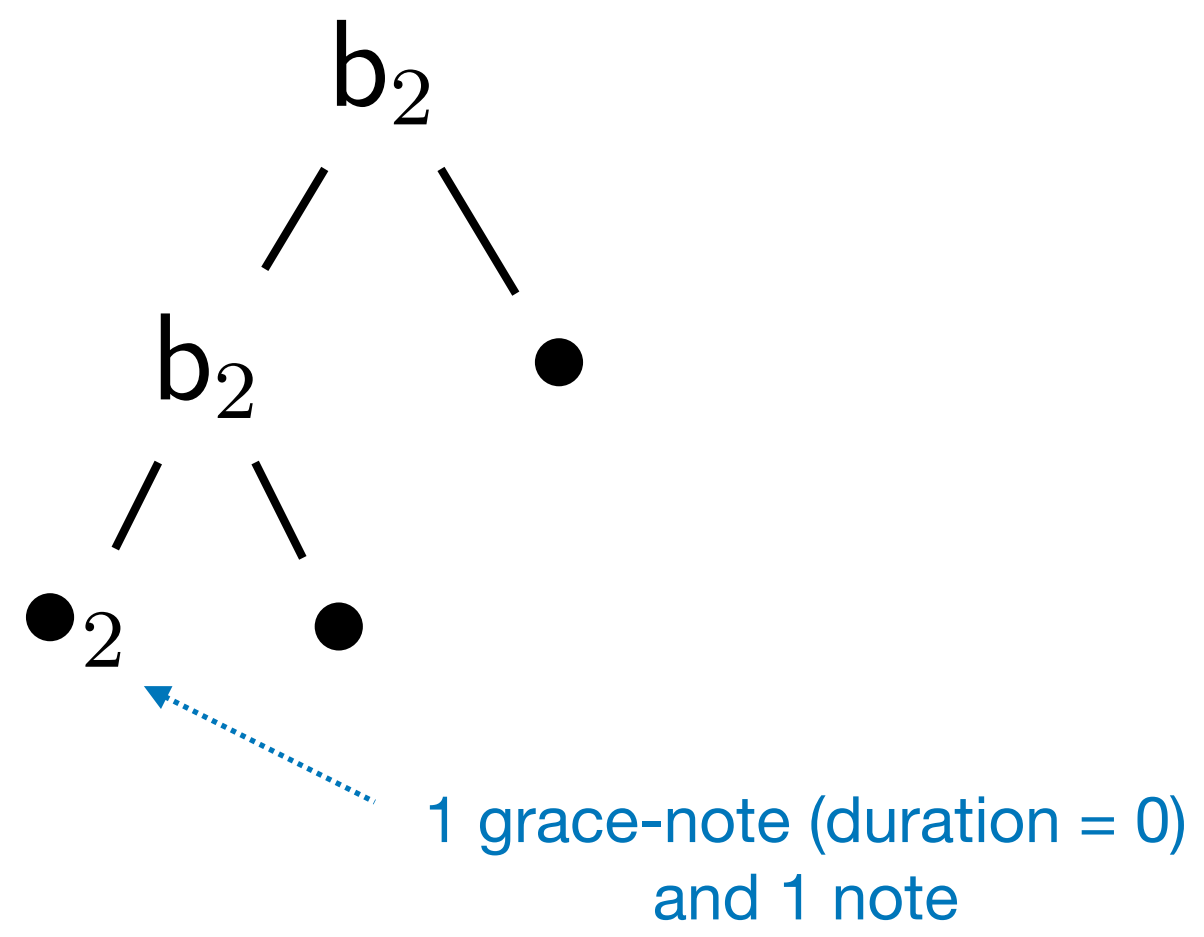
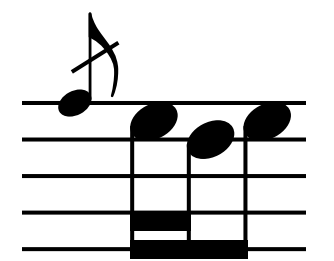
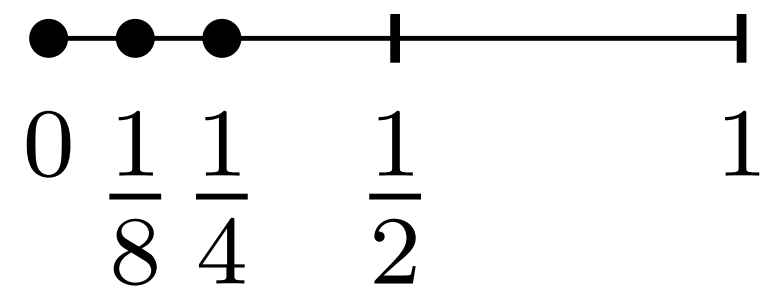
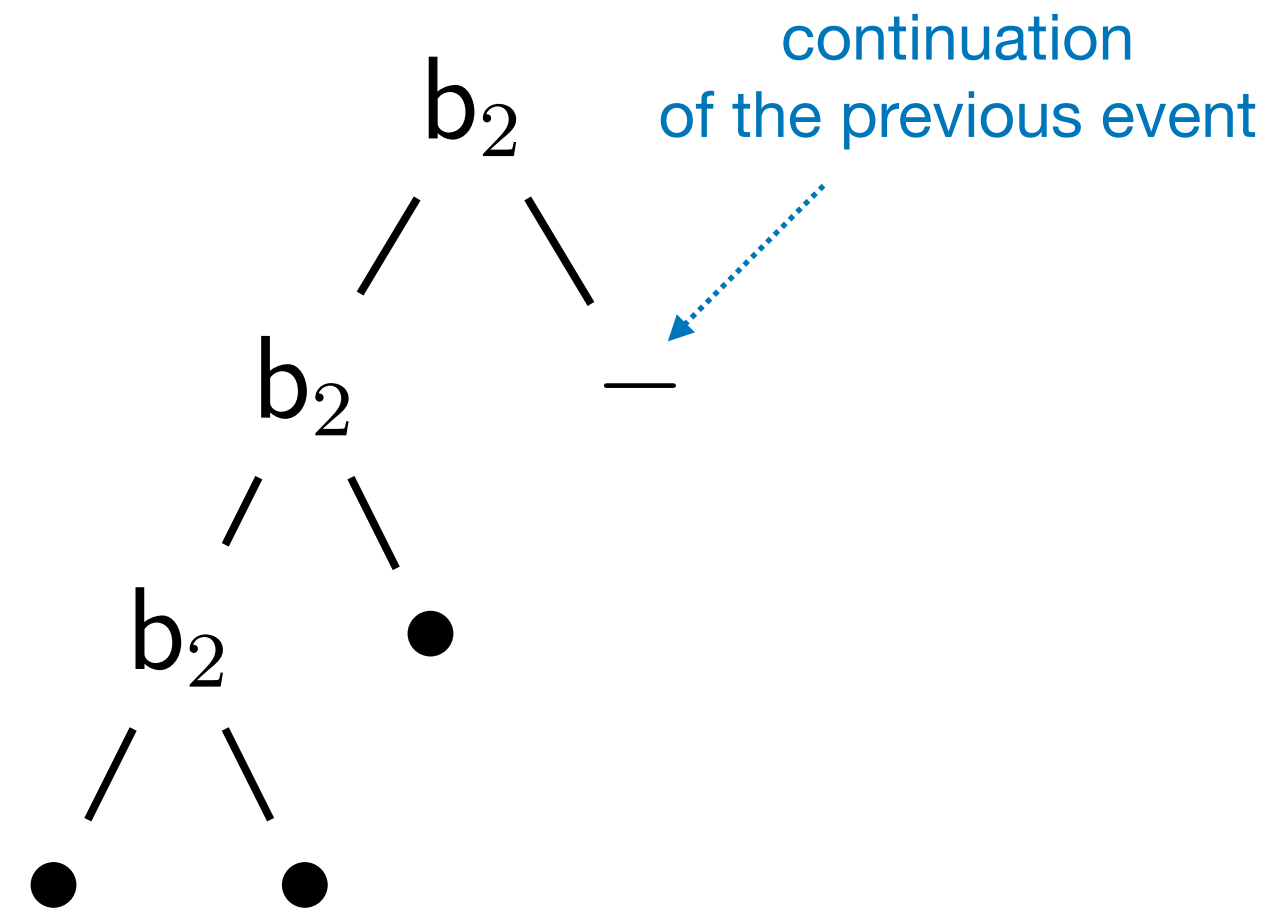
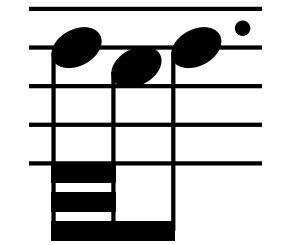
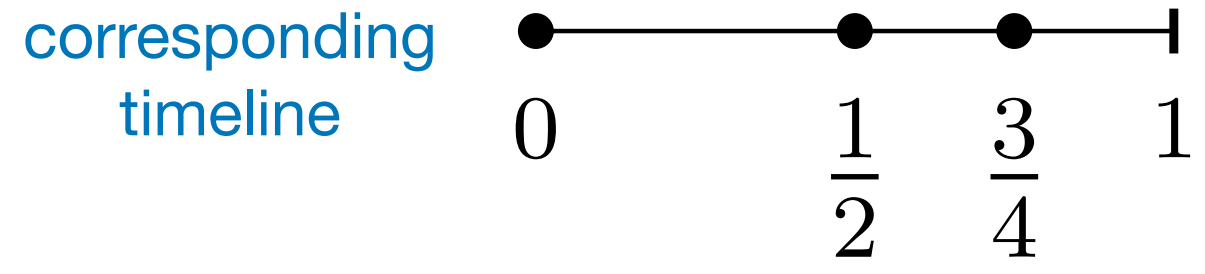
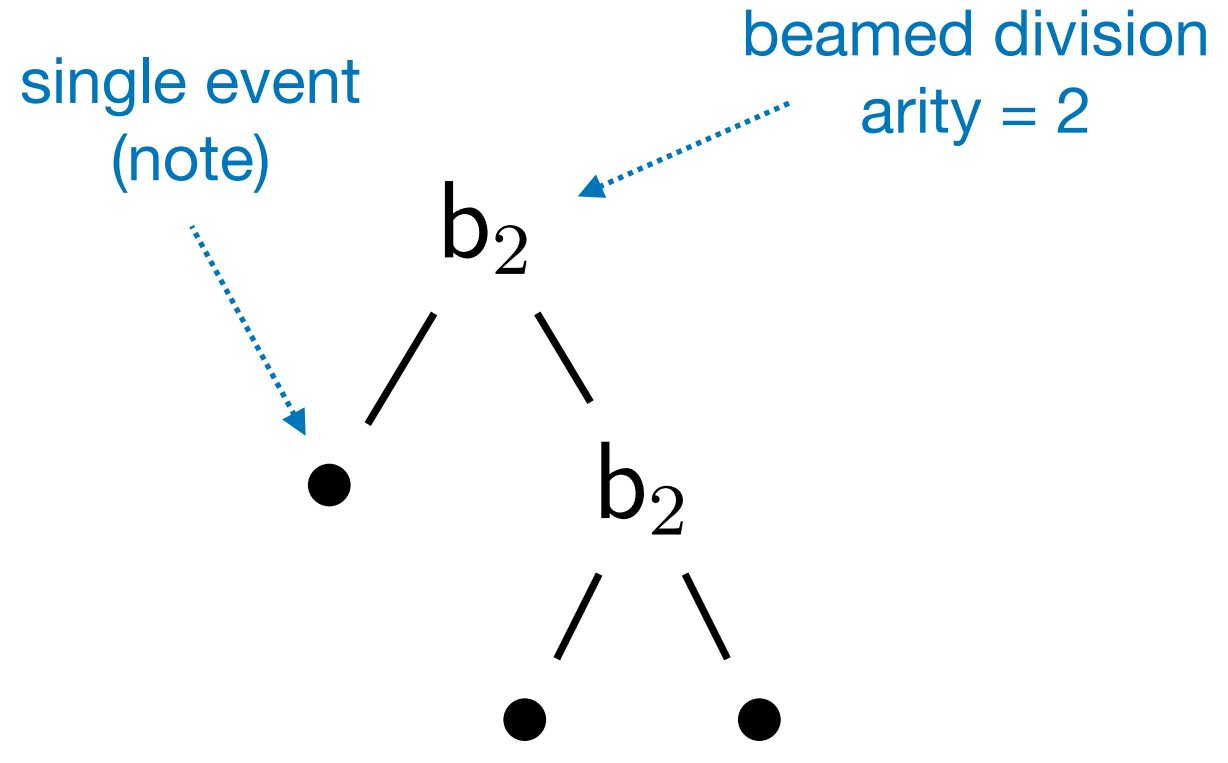
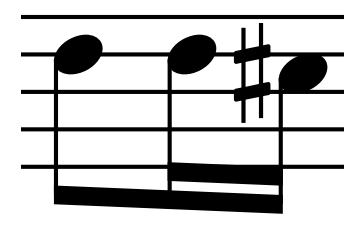
The second staff of music shows the beginning of the piece in 3/4 time. It consists of five bars. The first bar contains a quarter note (F4), an eighth note (F#4), and a quarter note (G4). The second bar contains a quarter note (A4), an eighth note (Bb4), and a quarter note (B4). The third bar contains a quarter note (C5), an eighth note (B4), and a quarter note (B4). The fourth bar contains a quarter note (A4), an eighth note (G4), and a quarter note (G4). The fifth bar contains a quarter note (F#4), an eighth note (F4), and a triplet of eighth notes (E4, D4, C4).

durations: $\frac{1}{2} \frac{1}{4} \frac{1}{4}$ $\frac{1}{16} \frac{1}{16} \frac{3}{4}$ $\frac{1}{16} \frac{1}{16} \frac{3}{4}$ 0 $\frac{1}{2} \frac{1}{4} \frac{1}{4}$ 2 $\frac{1}{2} \frac{1}{4} \frac{1}{4}$ $\frac{1}{16} \frac{1}{16} \frac{3}{4}$ $\frac{1}{2} \frac{1}{4} \frac{1}{4}$ $\frac{1}{2} \frac{1}{4} \frac{1}{4}$ $\frac{1}{16} \frac{1}{16} \frac{3}{4}$ $\frac{1}{2} \frac{1}{6} \frac{1}{6} \frac{1}{6}$ $\frac{1}{2} \frac{1}{2}$ 0 1

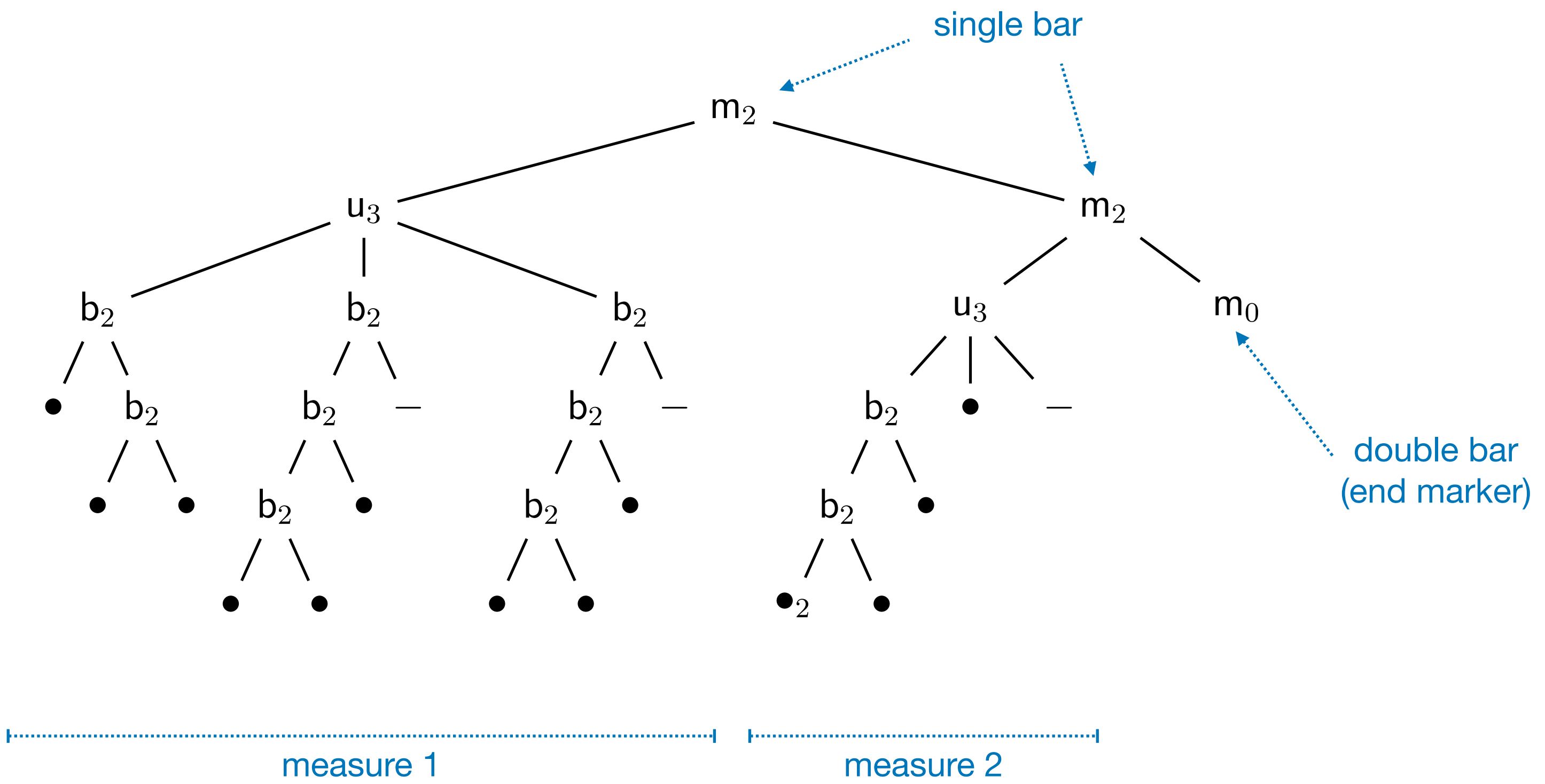
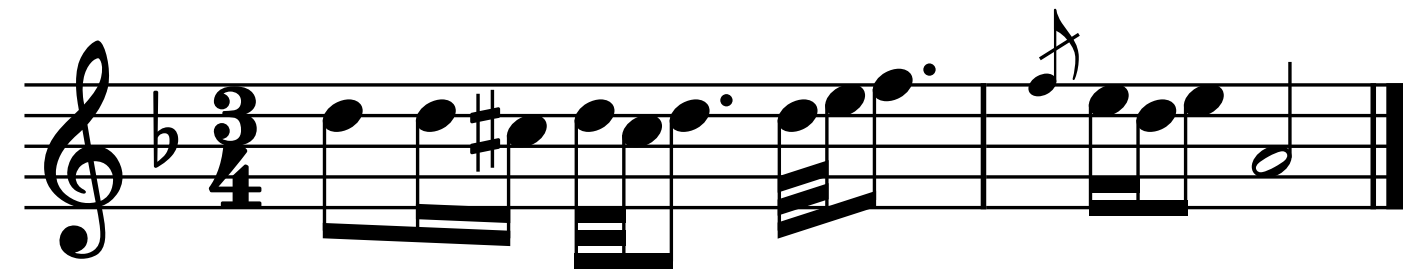
Tree-structured Representation of Music Notation

Tree representation of the proportional rhythmic notation with hierarchical encoding of durations: *“the (duration) data is in the structure”*

- the tree leaves contain the events
- the branching define durations, by partitioning of time intervals



Tree-structured Representation of Music Notation



Regular Tree Language (of Music Notation)

defined by a **Regular Tree Grammar**:

- non-terminal symbols: q, q_0, q_1, \dots
- terminal symbols (constants): \bullet (1 note), \bullet_2 (1 grace-note + 1 note), $-$ (continuation)
- production rules:

$$q \rightarrow m_2(q_0, q) \mid m_0$$

$$q_0 \rightarrow u_3(q_1, q_1, q_1) \mid \bullet \quad \text{measure}$$

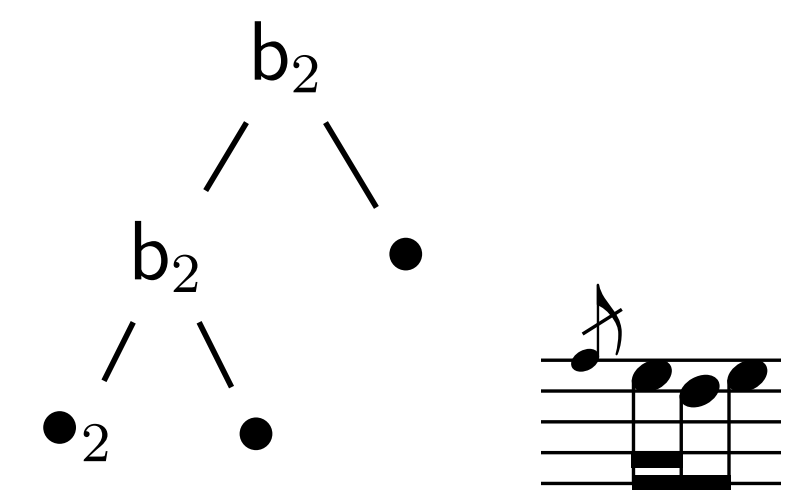
$$q_1 \rightarrow b_2(q'_2, q_2) \mid \bullet \mid \bullet_2 \mid - \quad \text{beat} = \text{♪}$$

$$q'_2 \rightarrow b_2(q'_3, q_3) \mid \bullet \mid \bullet_2 \mid - \quad q_2 \rightarrow b_2(q_3, q_3) \mid \bullet \mid - \quad \text{sub-beat} = \text{8th-note} = \text{♪}$$

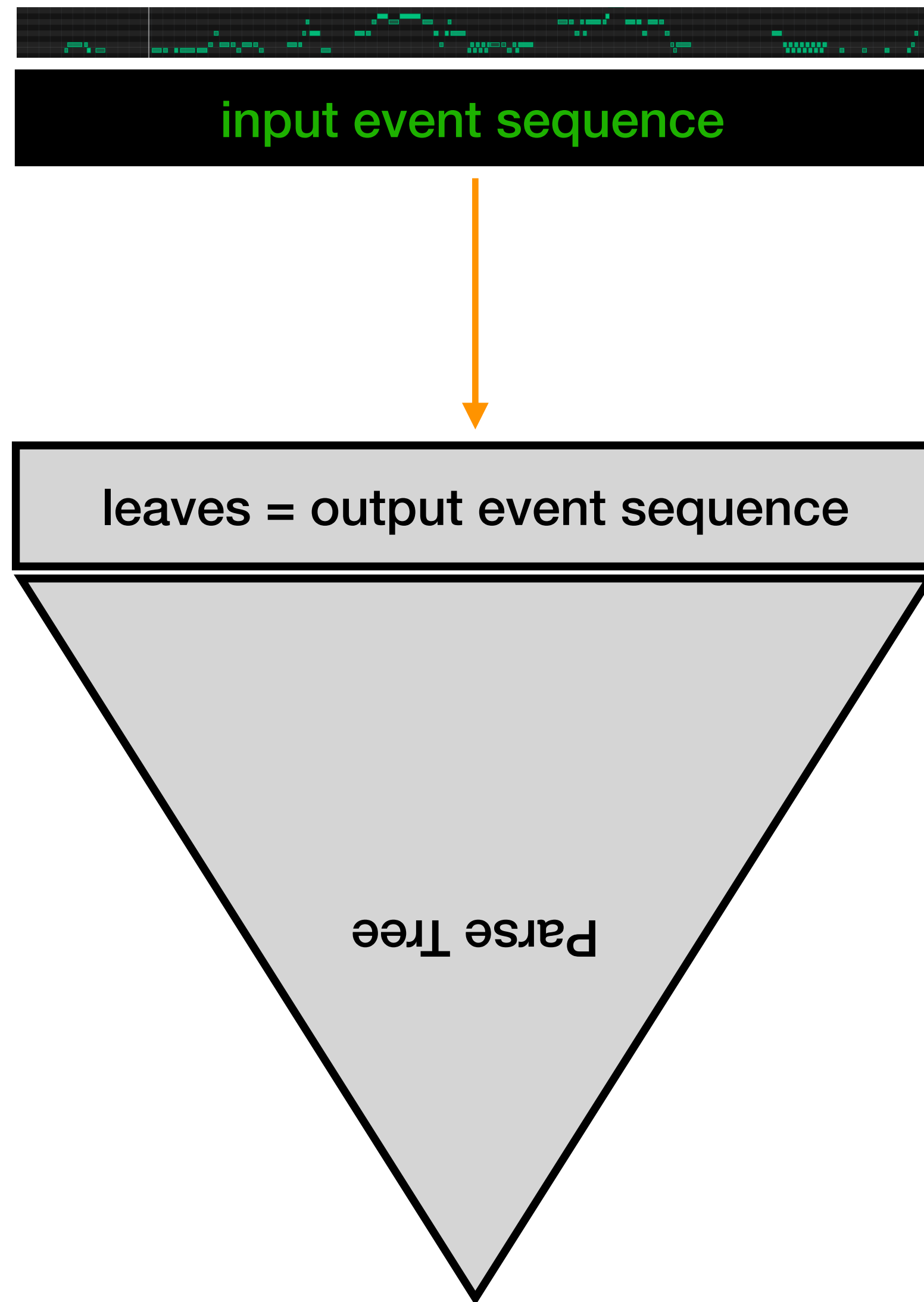
$$q'_3 \rightarrow \bullet \mid \bullet_2 \mid - \quad q_3 \rightarrow \bullet \mid - \quad \text{sub-sub-beat} = \text{16th note} = \text{♪}$$

derivations (leftmost)

$$q_1 \rightarrow b_2(q'_2, q_2) \rightarrow b_2(b_2(q'_3, q_3), q_2) \rightarrow b_2(b_2(\bullet_2, q_3), q_2) \rightarrow b_2(b_2(\bullet_2, \bullet), q_2) \rightarrow b_2(b_2(\bullet_2, \bullet), \bullet)$$



$$q \rightarrow m_2(q_0, q) \rightarrow m_2(u_3(q_1, q_1, q_1), q) \rightarrow m_2(u_3(b_2(q'_2, q_2), q_1, q_1), q) \rightarrow m_2(u_3(b_2(\bullet, q_2), q_1, q_1), q) \rightarrow \dots$$

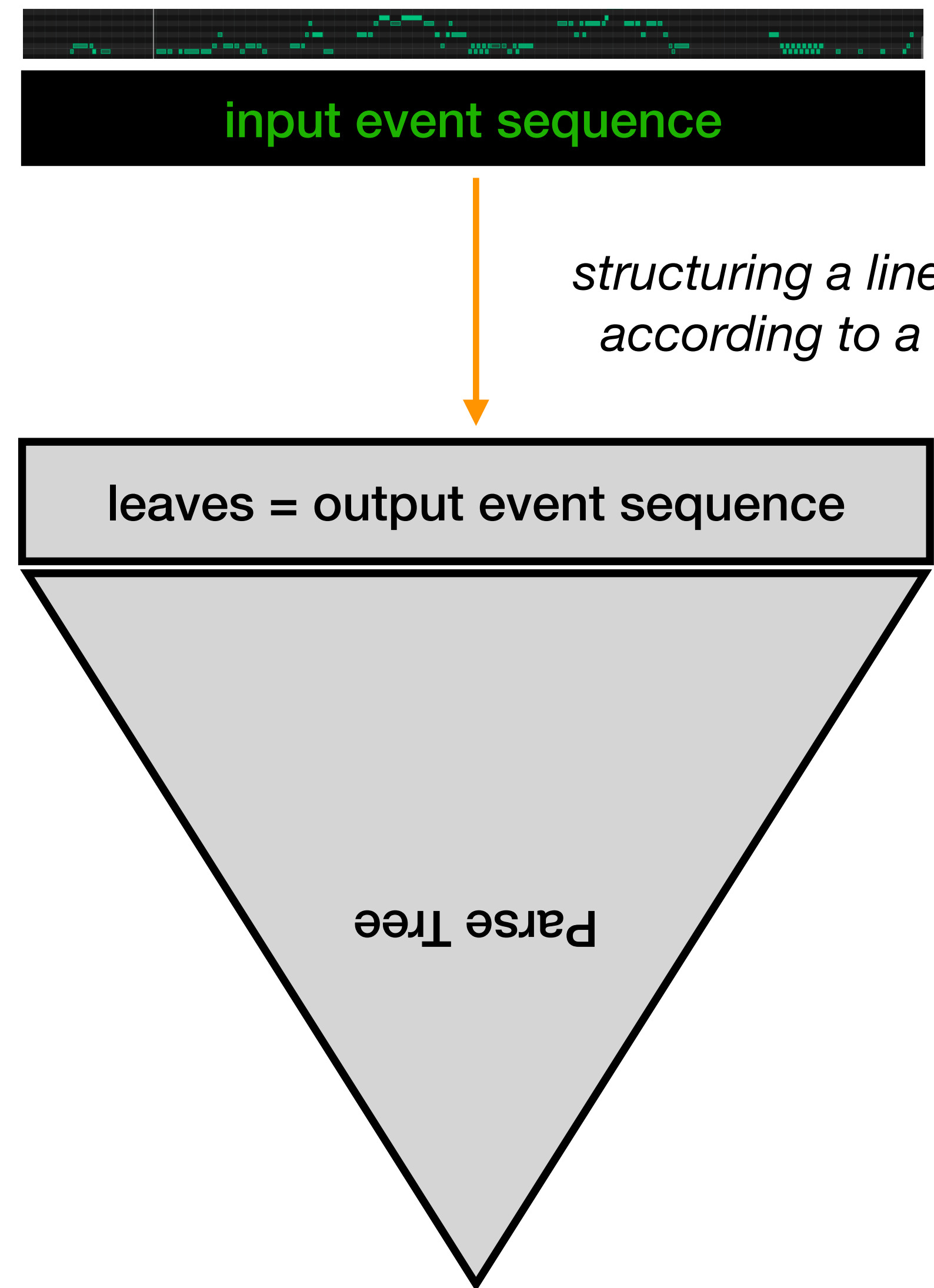


piano roll

= sequence of timestamped input events

tree-structured representation
of an output **music score**

conforming to a
prior language (expected notation)



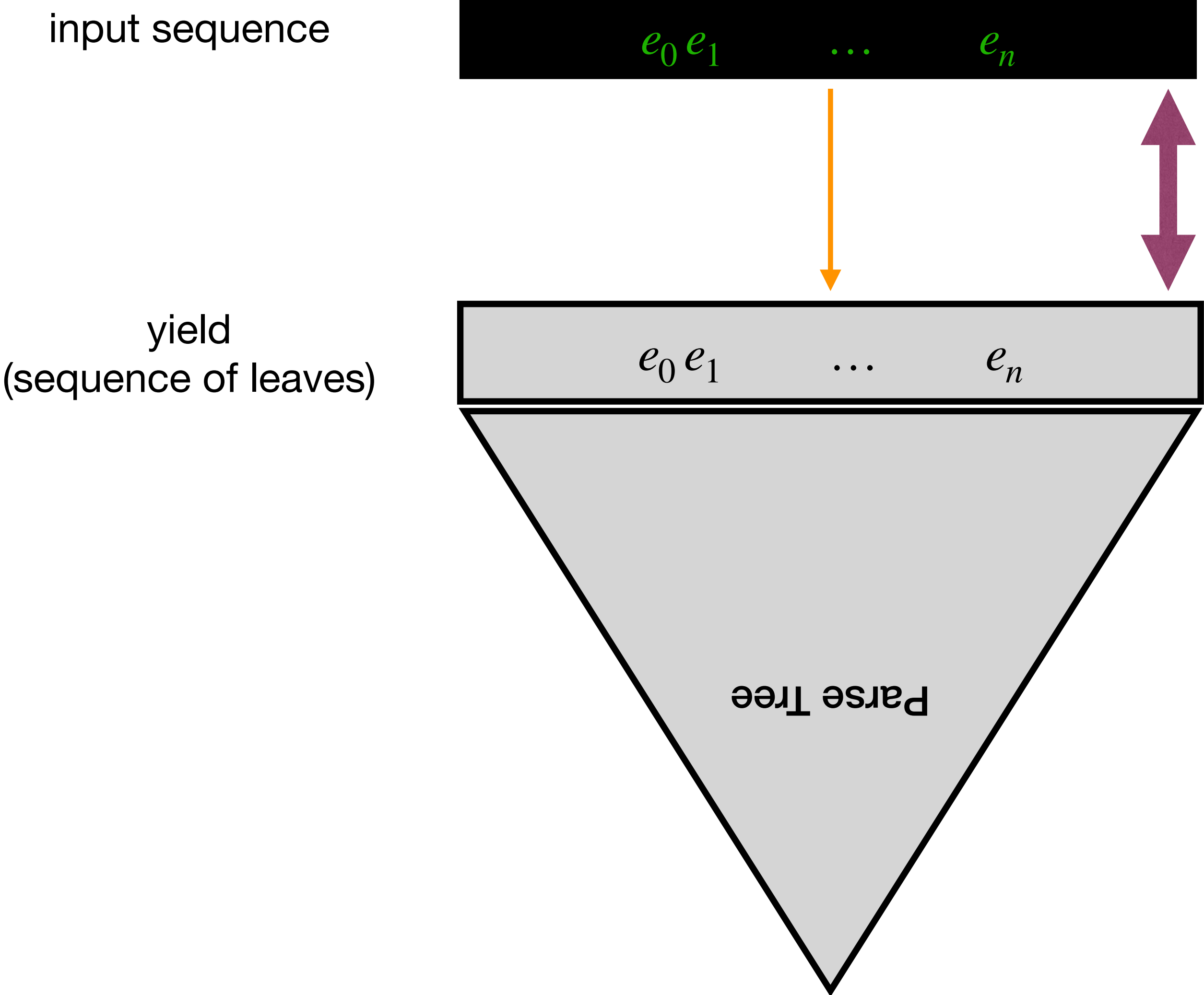
piano roll
= sequence of timestamped input events

structuring a linear representation according to a language model = **parsing**

tree-structured representation of an output **music score**
conforming to a prior language (expected notation)

Conventional Parsing

terminal symbols: e_0, \dots in a finite alphabet



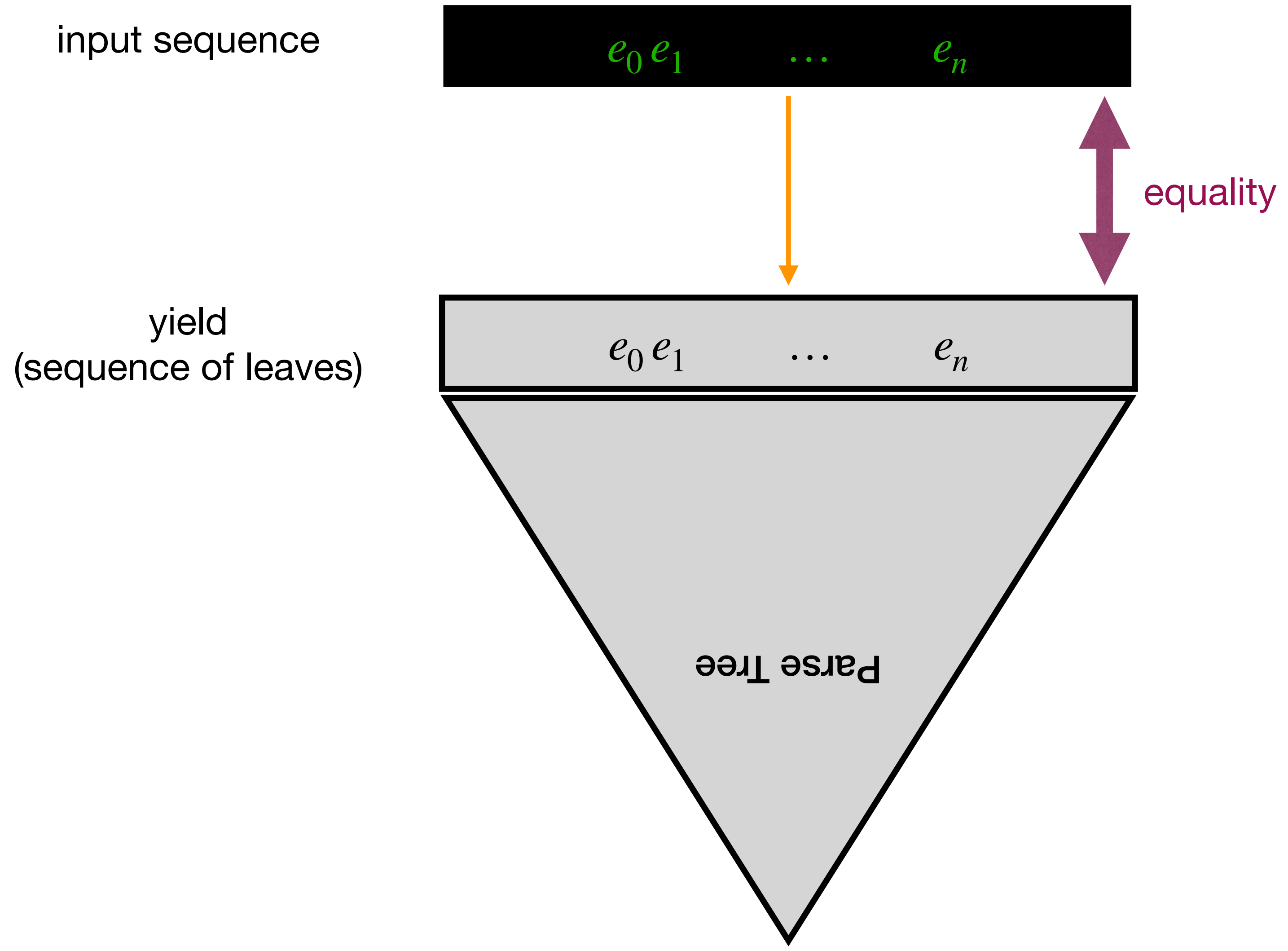
equality

parse-tree = representation of a leftmost derivation of $e_0 e_1 \dots e_n$ by a prior **CF-grammar** \mathcal{G} with production rules: $q_0 \rightarrow q_1 q_2$ or $q_0 \rightarrow e$ (non-terminal symbols: q_0, q_1, \dots)

Decision problem: (**membership**)
does there exist a parse tree (leftmost derivation) of \mathcal{G} that yields $e_0 e_1 \dots e_n$?

Weighted Parsing

Returning a parse tree of \mathcal{G} that yields $e_0 e_1 \dots e_n$



With an ambiguous prior CF-grammar \mathcal{G} there might exist several parse trees (exponentially many).

in order to choose one (or some) parse trees, rank them according to their **weight values**, computed by **Weighted Tree Grammar**

Weighted Regular Tree Grammar \mathcal{G} :

- non-terminal symbols: q, q_0, q_1, \dots
- terminal symbols (constants): \bullet (1 note), \bullet_2 (1 grace-note + 1 note), $-$ (continuation)
- every production rule is assigned a weight value (e.g. cost to read):

$q \xrightarrow{0} m_2(q_0, q)$	$q \xrightarrow{0} m_0$					
$q_0 \xrightarrow{0.1} u_3(q_1, q_1, q_1)$	$q_0 \xrightarrow{1} \bullet$					measure
$q_1 \xrightarrow{0.1} b_2(q'_2, q_2)$	$q_1 \xrightarrow{1} \bullet$	$q_1 \xrightarrow{1.9} \bullet_2$	$q_1 \xrightarrow{1} -$			beat =
$q'_2 \xrightarrow{0.1} b_2(q'_3, q_3)$	$q'_2 \xrightarrow{1} \bullet$	$q'_2 \xrightarrow{2.25} \bullet_2$	$q'_2 \xrightarrow{1} -$			sub-beat = 8th-note =
$q_2 \xrightarrow{0.1} b_2(q_3, q_3)$	$q_2 \xrightarrow{1} \bullet$	$q_2 \xrightarrow{1} -$				
	$q'_3 \xrightarrow{1} \bullet$	$q'_3 \xrightarrow{3.25} \bullet_2$	$q'_3 \xrightarrow{1} -$	$q_3 \xrightarrow{1} \bullet$	$q_3 \xrightarrow{1} -$	sub-sub-beat = 16th note =

derivation (leftmost):

$$d : q_1 \xrightarrow{0.1} b_2(q'_2, q_2) \xrightarrow{0.1} b_2(b_2(q'_3, q_3), q_2) \xrightarrow{3.25} b_2(b_2(\bullet_2, q_3), q_2) \xrightarrow{1} b_2(b_2(\bullet_2, \bullet), q_2) \xrightarrow{1} b_2(b_2(\bullet_2, \bullet), \bullet)$$

cost of derivation: $\text{weight}(d) = 0.1 + 0.1 + 3.25 + 1 + 1$

learning weight values from corpus statistics - [Francesco Foscarin](#)

In general, the weight values are taken in a **commutative Semiring** $\langle \mathbb{S}, \oplus, \otimes, \mathbb{0}, \mathbb{1} \rangle$

- \oplus and \otimes are associative and commutative, with neutral elements $\mathbb{0}$ and $\mathbb{1}$
- \otimes distributes over \oplus : $x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$
- $\mathbb{0}$ is absorbing for \otimes : $\mathbb{0} \otimes x = \mathbb{0}$

	domain	\oplus	\otimes	$\mathbb{0}$	$\mathbb{1}$
Boolean	$\{\perp, \top\}$	\vee	\wedge	\perp	\top
Viterbi	$[0,1] \subset \mathbb{R}$	max	\times	0	1
Tropical min-plus	$\mathbb{R}_+ \cup \{+\infty\}$	min	+	$+\infty$	0

Moreover, \oplus is assumed to extend to **infinite sums**.

\otimes is for composition of rule's weights in derivations and \oplus is for optimal choice:

For a Weighted Regular Tree Grammar \mathcal{G}

$$\text{weight}_{\mathcal{G}}(d : q \xrightarrow{w_1} \dots \xrightarrow{w_n} t) = \bigotimes_{i=1}^n w_i \quad \text{and} \quad \text{weight}_{\mathcal{G}}(q, t) = \bigoplus_{d:q \xrightarrow{+} t} \text{weight}_{\mathcal{G}}(d)$$

or recursively:

$$\text{weight}_{\mathcal{G}}(q, a(t_1, \dots, t_n)) = \bigoplus_{q \xrightarrow{w} a(q_1, \dots, q_n) \in \mathcal{G}} \left(w \otimes \bigotimes_{i=1}^n \text{weight}_{\mathcal{G}}(q_i, t_i) \right)$$

\mathbb{S} is assumed :

- **idempotent** $x \oplus x = x$

that induces a partial **ordering**: $x \leq_{\oplus} y$ iff $x \oplus y = x$

- **total** : $\forall x, y \in \mathbb{S}$, either $x \oplus y = x$ or $x \oplus y = y$ i.e. \leq_{\oplus} is total

- **bounded** : $\mathbb{1} \oplus x = \mathbb{1}$, or equivalently: $\forall x, y \in \mathbb{S}, x \leq_{\oplus} x \otimes y$

i.e. combining elements with \otimes always increases their weight,

see the *non-negative weights* condition for Dijkstra's shortest path algorithm

k -best parsing : enumeration of the k best weighted trees wrt \leq_{\oplus} for \mathcal{G} and a non-terminal q , in PTIME.

similar to best path search in hyper-graphs (Dynamic Programming)

- Viterbi algorithm in acyclic case

- Knuth generalization of the Dijkstra algorithm in the general case

	domain	\oplus	\otimes	\ominus	$\mathbb{1}$
Boolean	$\{\perp, \top\}$	\vee	\wedge	\perp	\top
Viterbi	$[0,1] \subset \mathbb{R}$	max	\times	0	1
Tropical min-plus	$\mathbb{R}_+ \cup \{+\infty\}$	min	+	$+\infty$	0

in the context of music transcription, the symbols in the sequence in input are timestamped



input sequence

e_0	e_1	...	e_m
τ_0	τ_1	...	τ_m

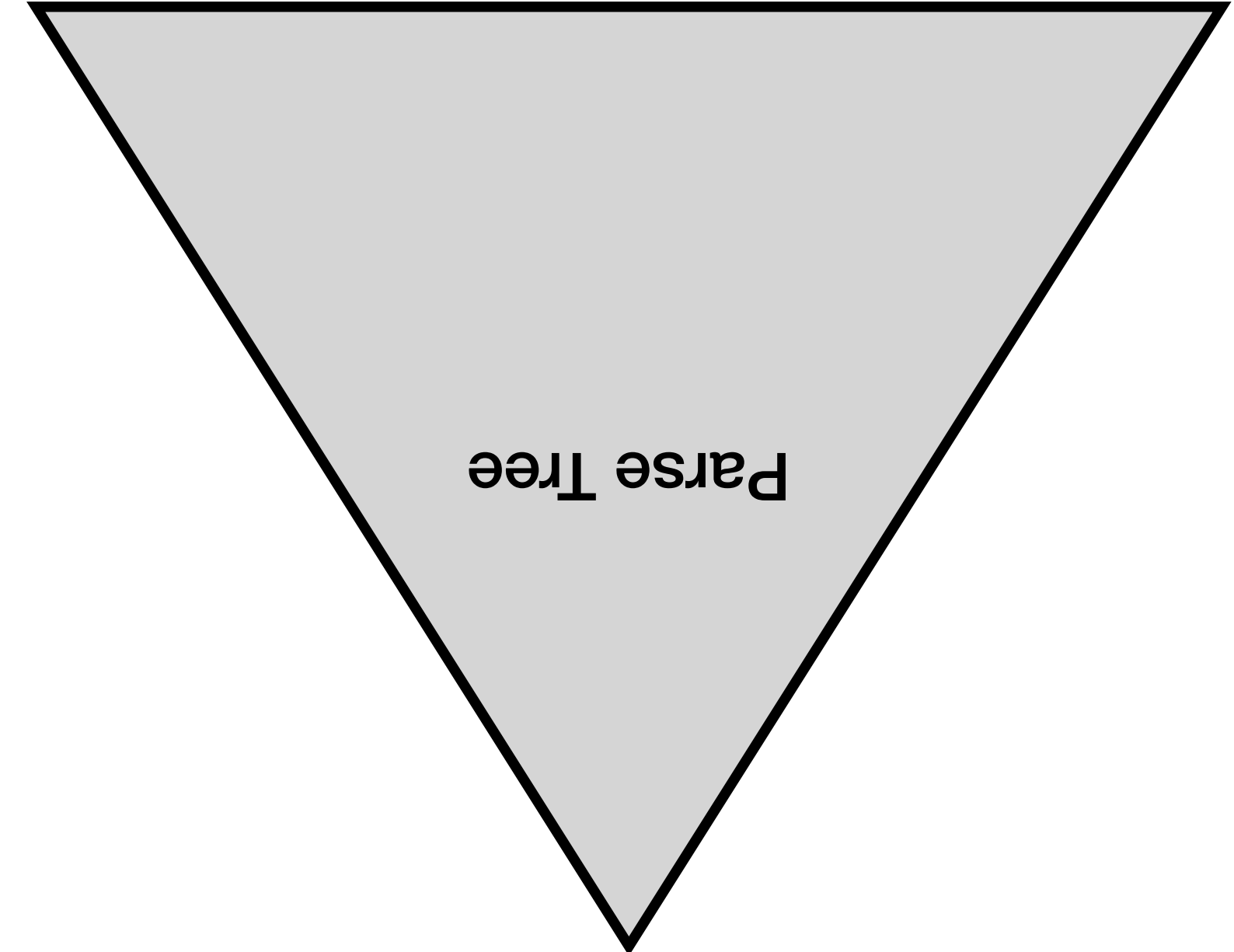
ranking of solutions with:

a measure of input / output fitness
= cost of IO alignment

yield
= sequence of leaves
decorated with dates θ
(computed with the durations
encoded in the tree structure)

a_0	a_1	...	a_n
θ_0	θ_1	...	θ_n

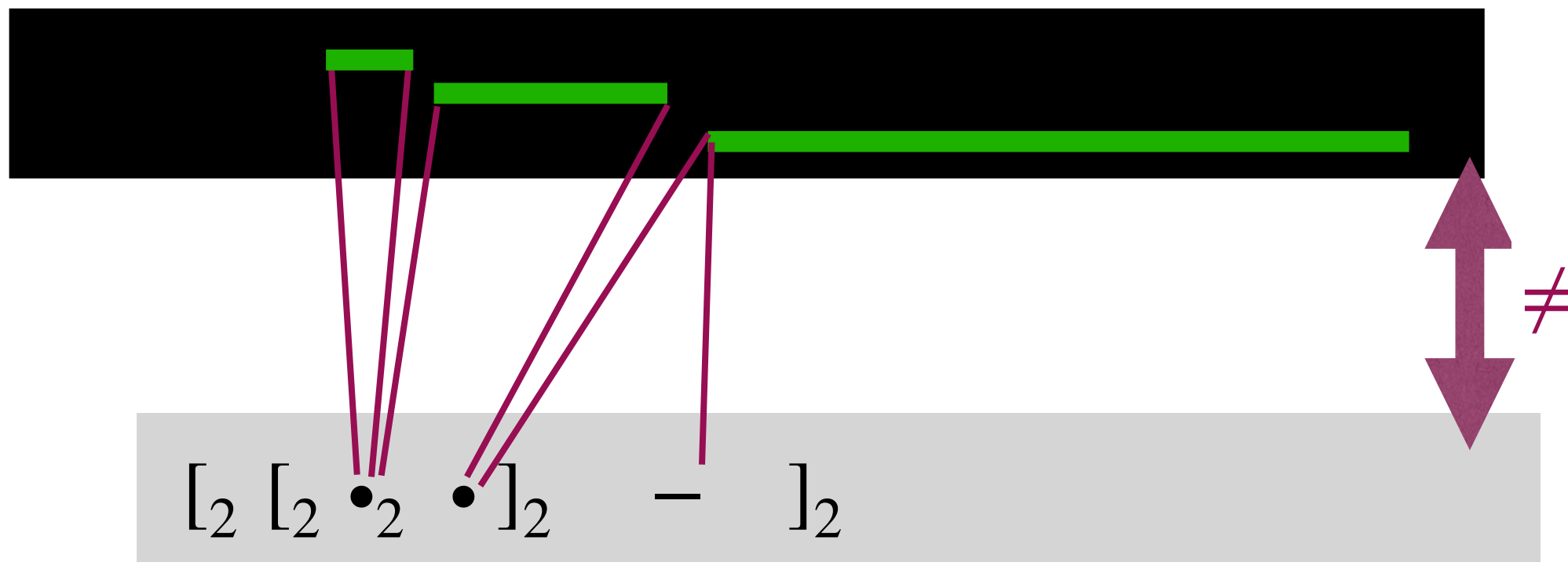
\otimes
a measure of cost-to-read
(weight value computed by **Weighted Tree Grammar**)



measure of input/output fitness

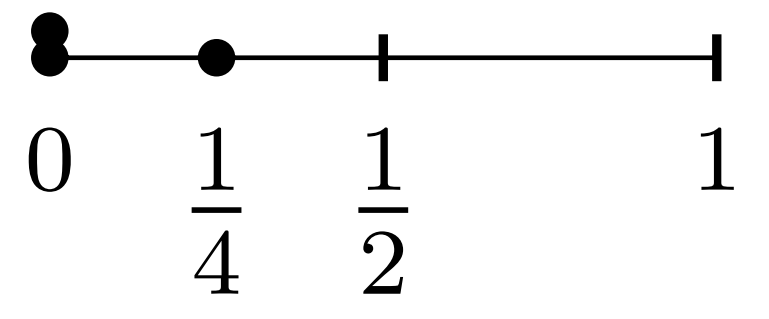
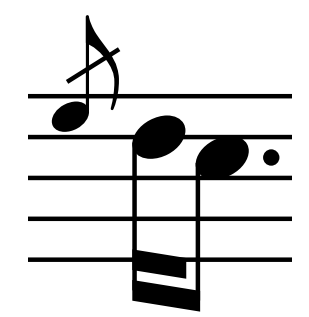
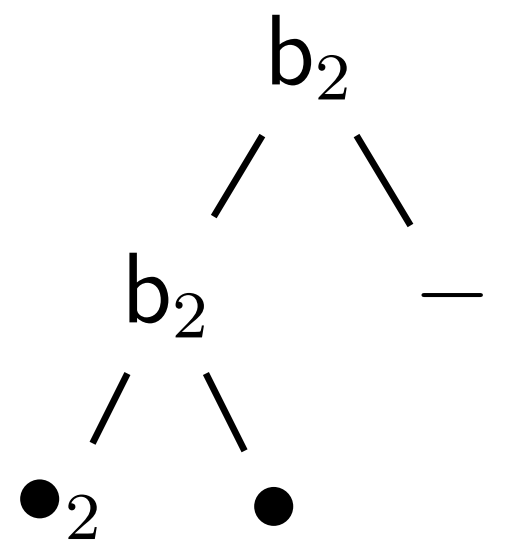
E_{on}	E_{off}	D_{on}	D_{off}	C_{on}	C_{off}
0.11	0.19	0.22	0.48	0.53	1.08

input sequence

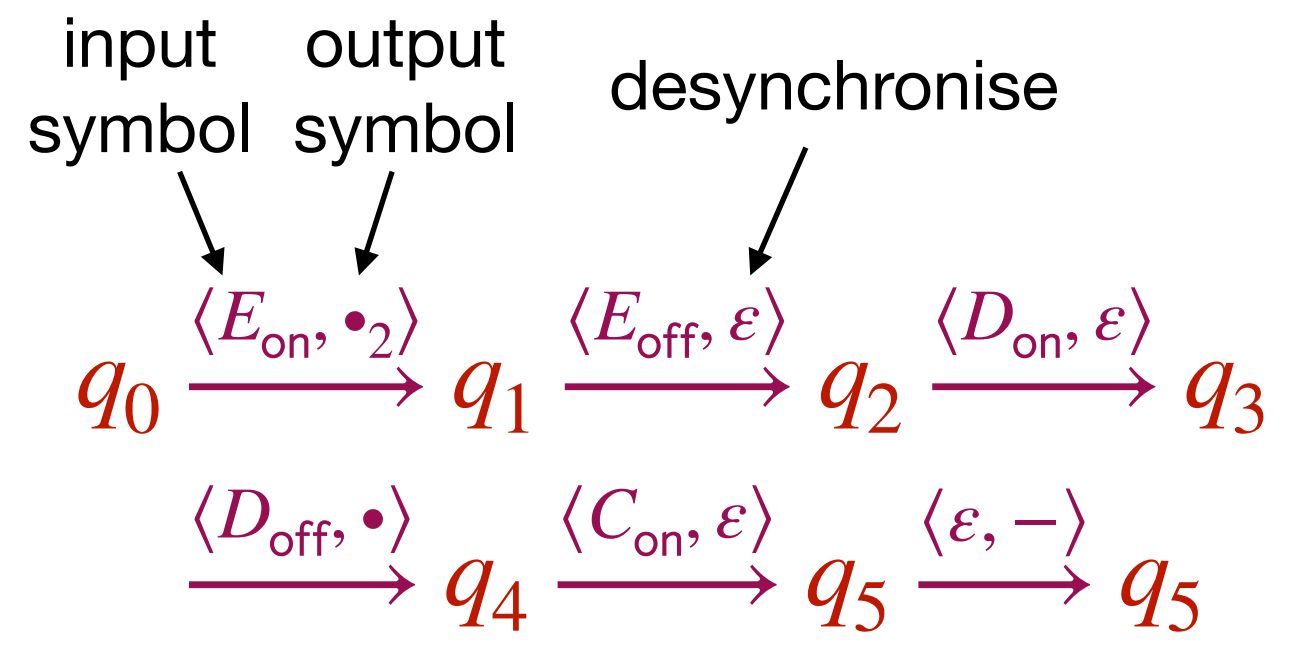


linearisation of the output tree

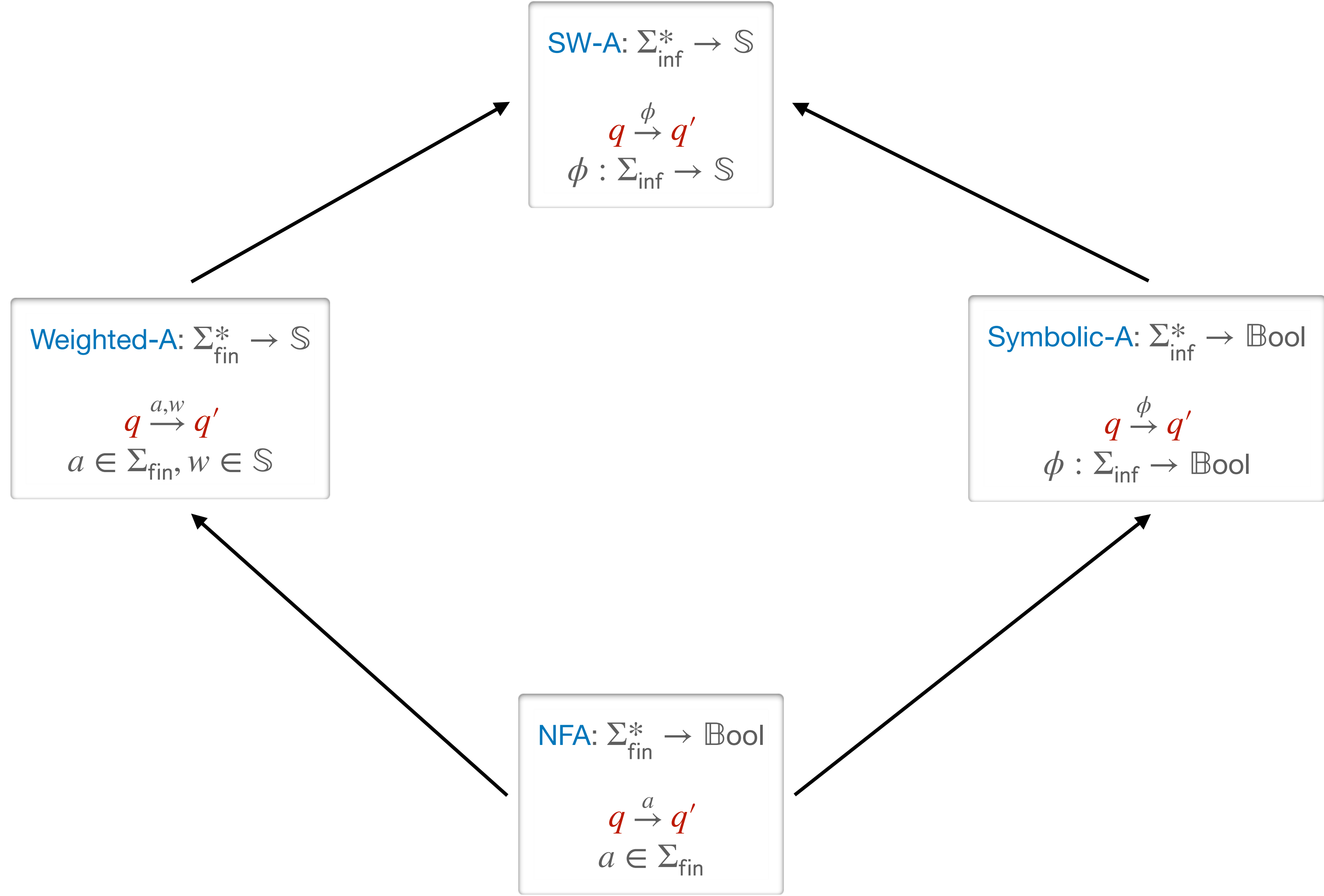
$[_2 [_2 \bullet_2 \bullet_2]_2 -]_2$



cost of IO alignment
computed by a
word-to-word Transducer
(stateful definition of an edit-distance)



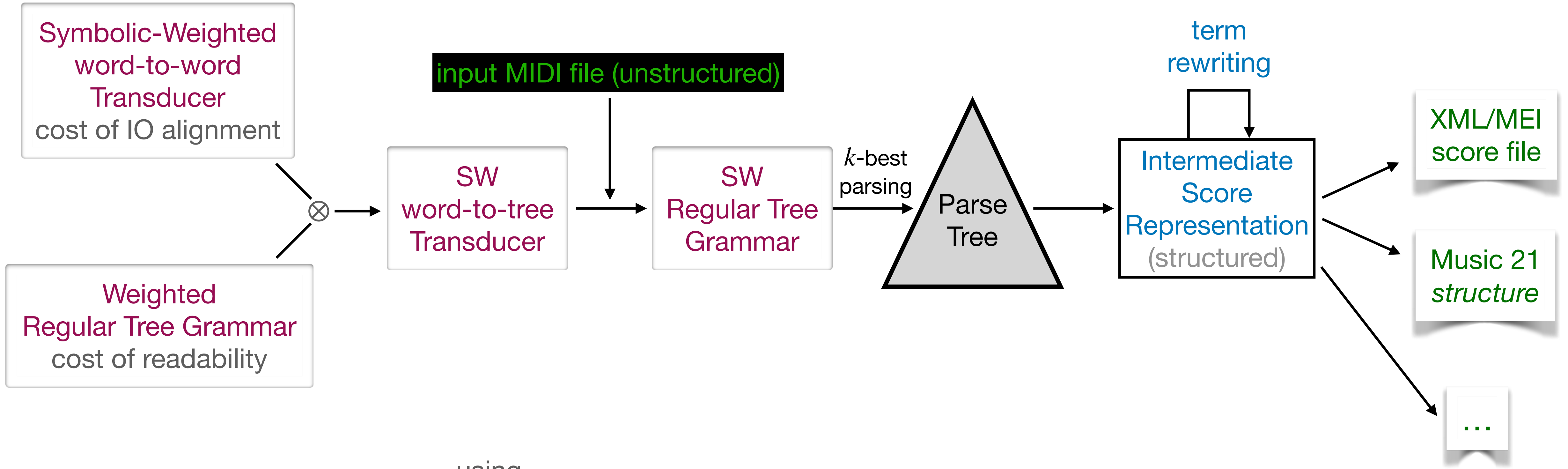
in the context of music transcription, the symbols are timestamped \rightarrow infinite alphabet Σ_{inf}



Droste, M., Kuich, W., Vogler
Handbook of WA, 2009

Veanes et al.
CAV 2017, CACM 2021

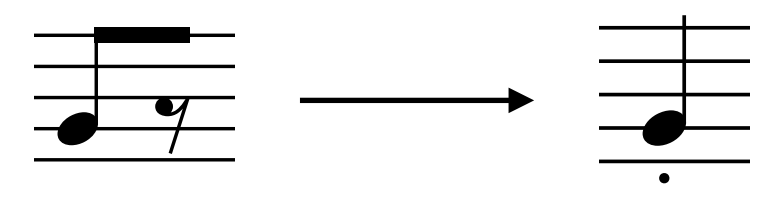
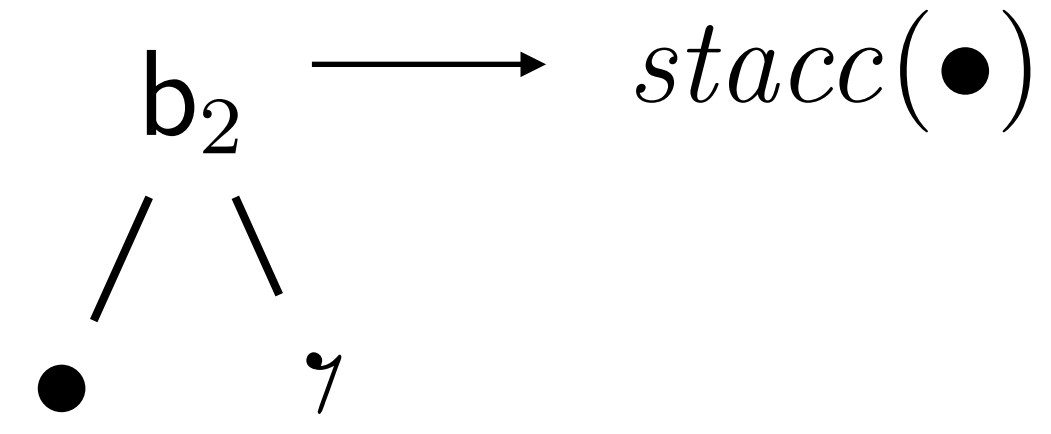
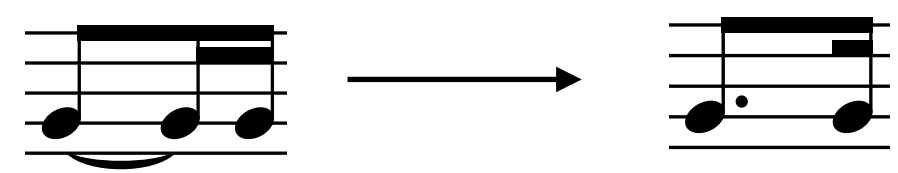
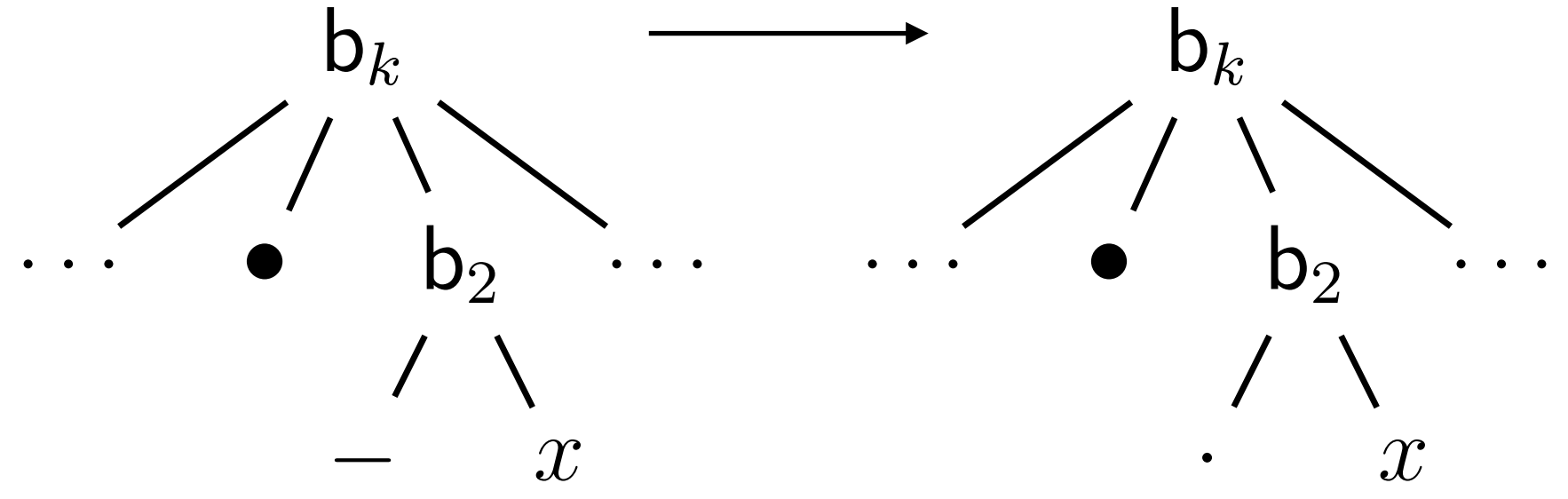
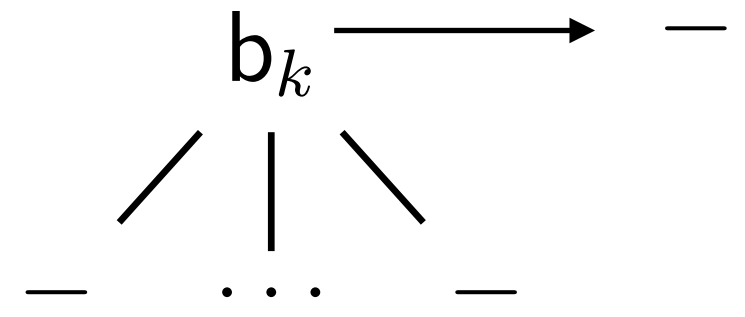
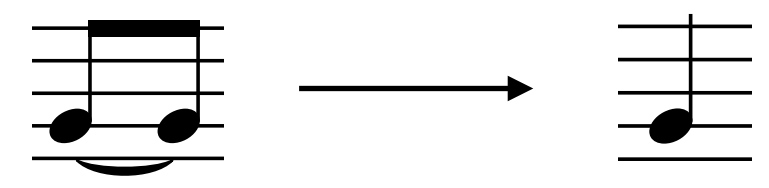
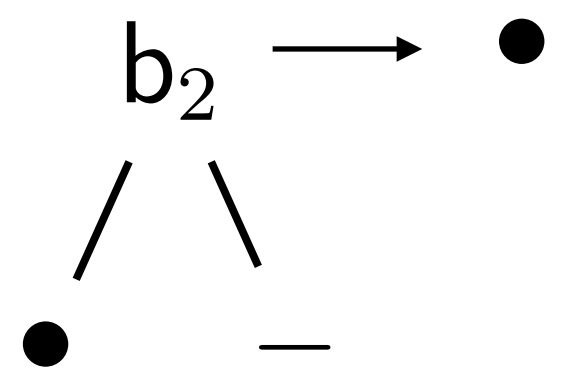
Automated Music Transcription: qparse framework



using
SW Visibly Pushdown Automata
as intermediate model

Term Rewriting Rules

for the transformation of the intermediate score representation



questions: rewrite strategies (e.g. IO or OI), conflicts...

implementation of

- the above transcription by parsing framework
- + other subtasks: pitch-spelling, key estimation, beat tracking...

qparse: 75 Kloc C++

<https://gitlab.inria.fr/qparse/qparselib>
<https://qparse.gitlabpages.inria.fr>

- command lines tools:
monoparse, drumparse, grammar-learning, engraving
- Python binding - [Lydia Rodrigez-de la Nava](#)
automatic evaluation
- online port, real-time - [Leyla Villaroel](#)

Monophonic transcription

monophonic : one note at a time

Good results for complex cases (tuplets, ornaments, silences, small durations...)

Polonaise in D minor from Notebook for Anna Magdalena
Bach BWV Anh II 128

original score

transcription of MIDI recording by [qparse](#)

Monophonic transcription

monophonic : one note at a time

Good results for complex cases (tuplets, ornaments, silences, small durations...)

Polonaise in D minor from Notebook for Anna Magdalena
Bach BWV Anh II 128

original score

Moderato

6

11

17

transcription of MIDI recording by [Finale](#)

5

6

9

14

6

Beethoven, Trio for violin, cello and piano, op.70 n.2 (2d mov)

original score

Allegretto
Violon
p dolce

The original score for the Violon part is in 2/4 time, marked Allegretto and *p dolce*. It consists of three staves of music. The first staff begins with a treble clef and a 2/4 time signature. The music features a mix of eighth and sixteenth notes, with some triplet markings. The second staff includes a fermata over a triplet and a *8va* marking. The third staff contains a trill (*tr*) and further triplet markings. Green horizontal lines are drawn under the first and second measures of the third staff.

transcription
of MIDI recording
by [qparse](#)

The transcription of the MIDI recording by qparse is shown in two staves. The first staff is a single-line transcription of the first two staves of the original score. The second staff is a single-line transcription of the third staff of the original score. Red circles highlight three specific areas of discrepancy: the final measure of the first staff, the first measure of the second staff, and the second measure of the second staff.

Monophonic transcription

Beethoven, Trio for violin, cello and piano, op.70 n.2 (2d mov)

original score

transcription of MIDI recording by **Finale**

- options:
- mixed rhythms,
 - triplets
 - smallest note = 32nd
- The time signature and the tempo are given.

OMRized version

Les surprises de l'amour

Musical score for the first system (measures 41-45) of the OMRized version. It features three staves: Treble, Bass, and Bass. The music is in 3/4 time with a key signature of one flat. The notation includes various note values, rests, and accidentals. A yellow highlight is present under a note in the Bass staff at measure 44.

Musical score for the second system (measures 46-51) of the OMRized version. It features three staves: Treble, Bass, and Bass. The music continues with similar notation. A red highlight is present under a note in the Bass staff at measure 51.

Musical score for the third system (measures 52-56) of the OMRized version. It features three staves: Treble, Bass, and Bass. The notation includes various note values, rests, and accidentals. Yellow highlights are present under notes in the Treble and Bass staves at measures 52, 53, 54, and 55.

Manual correction (ground truth)

Les surprises de l'amour

Ouverture
Adagio

Musical score for the first system (measures 41-45) of the Manual correction version. It features three staves: Pr violon, 2e violon, and Basses. The music is in 3/4 time with a key signature of one flat. The notation includes various note values, rests, and accidentals. Dynamic markings 'Doux' and 'Fort' are present. Trill ornaments are marked above notes in the Pr violon and 2e violon staves. A yellow highlight is present under a note in the Basses staff at measure 44.

Musical score for the second system (measures 46-50) of the Manual correction version. It features three staves: Pr violon, 2e violon, and Basses. The notation includes various note values, rests, and accidentals. Dynamic markings 'Doux', 'Fort', and 'F' are present. Trill ornaments are marked above notes in the Pr violon and 2e violon staves. A yellow highlight is present under a note in the Basses staff at measure 49.

Musical score for the third system (measures 51-56) of the Manual correction version. It features three staves: Pr violon, 2e violon, and Basses. The notation includes various note values, rests, and accidentals. Dynamic markings 'Doux' are present. Trill ornaments are marked above notes in the Pr violon and 2e violon staves. A yellow highlight is present under a note in the Basses staff at measure 51.

Lamarque-Goudard dataset (w. Francesco Foscarin, Teysir Baoueb)

- 283 monophonic extracts of classical repertoire inspired by a rhythm learning method
- ~ 20 measures per extract
- progressive difficulty cover a very large spectrum of rhythmic features
- score files (XML) and MIDI performances for [evaluation](#) and calibration of transcription tools



FiloBass by John-Xavier Riley (QMUL, C4DM) project "Dig That Lick"

- jazz bass lines, acc. of saxophone
- 48 tracks, 24 recorded hours of melodies and improvisations
- qparse as backend of an audio-MIDI transcription

-2-

The image displays six staves of musical notation, likely bass lines, in a key signature of three flats (B-flat major or D-flat minor). The staves are numbered 74, 80, 86, 92, 98, 104, and 110. The notation includes various rhythmic patterns, including triplets and sixteenth notes. The first staff (74) shows a sequence of eighth and sixteenth notes. The second staff (80) shows a sequence of eighth notes. The third staff (86) shows a sequence of eighth notes with some accidentals. The fourth staff (92) shows a sequence of eighth notes with some accidentals. The fifth staff (98) shows a sequence of eighth notes with some accidentals. The sixth staff (104) shows a sequence of eighth notes with some accidentals and a triplet. The seventh staff (110) shows a sequence of eighth notes with some accidentals and a triplet.

Groove MIDI Dataset

- by Google Magenta
- 13.6 hours, 1150 MIDI files, ~ 22000 measures recorded by professional drummers on a electronic drum kit
- audio (wav) files synthesized from (and aligned to) MIDI files for evaluation of audio-to-MIDI drum transcription
- no score files!



scoring the GMD with qparse

Martin Digard (INALCO)

- all score files (XML) produced from the MIDI files with the same generic tree grammar (4/4 measure)
- polyphonic case-study, simpler than piano
- specific drumming constraints (hands ≤ 2 , feet ≤ 2)
- processing errors from MIDI sensors

A musical score for a drum set in 4/4 time. The score consists of 29 measures, numbered 1 through 29. The notation uses a standard drum set notation with various symbols for snare, tom, floor tom, kick, and cymbal. The score is written on a single staff with a treble clef and a 4/4 time signature. The notation includes various rhythmic patterns, such as eighth and sixteenth notes, and rests. There are also some triplets and accents indicated in the score.

- Dataset **ASAP** - [Francesco Foscarin, Andrew Mc Leod](#)
MIDI and audio recording from Yamaha piano competition
+ XML scores
+ alignments
+ beat tracking annotations
- voice separation - [Lydia Rodrigez-de la Nava](#), evaluation [Augustin Bouquillard](#)
and for piano guitar transcription.
integration in transcription:
 - before parsing, or
 - after parsing (on intermediate model), or
 - joint with parsing.