

Music Content Descriptors

18 janvier 2022

Florent

<https://github.com/polifonia-project/facets-search-engine>

Draft sur les descripteurs de contenu musical utilisés dans des documents précédents (articles orientés langages et code/doc de qparse) et évolutions plus récentes.

L'objectif de ces descripteurs était à l'origine différent de la recherche/indexation: génération ou transformation de partition... Mais le modèle est relativement ouvert et peut-être adapté à ce cas.

principe

représentation structurée de contenu musical écrit, avec dans un même modèle

- des événements musicaux
- des relations entre ces événements, en particulier des relations temporelles.

labeled ranked trees

The representation of a music score is a ranked tree labeled with symbols in a finite ranked alphabet.

musical events

they are represented by nullary symbols (atoms)

- note
- rest
- continuation of event (tie, dot)
- blank (duration without event)
- ... ?

These events have no duration attributes, unlike MusicXML or MEI counterparts. The duration and start time of each event are deducible from the embedding temporal operators (see description below).

The grace notes and chords can be constructed using the temporal operators below.

temporal operators

We consider musical time values in \mathbb{Q} , with a unit (MTU) = 1 measure (bar).

The following n -ary operators of Σ ($n > 0$) can be interpreted on temporal durations (MTU domain $D = \mathbb{Q} \cup \{+\infty\}$ or more generally on time intervals (left-closed, right-open).

- n -uplet: division in n equal durations

$$d_n : D^n \rightarrow D, \frac{p}{q.n}, \dots, \frac{p}{q.n} \mapsto \frac{p}{q}$$

$$\text{or } d_n : I^n \rightarrow I, [a, a + \frac{b-a}{n}[, [a + \frac{b-a}{n}, a + \frac{2(b-a)}{n}[\dots [a + \frac{(n-1)(b-a)}{n}, b[\rightarrow [a, b[$$

- bar-split (binary): 1st arg. has duration 1 bar, 2d is the rest

$$b : D^2 \rightarrow D, 1, +\infty \mapsto +\infty$$

$$\text{or } I^2 \rightarrow I, [a, a + 1[, [a + 1, +\infty[\rightarrow [a, +\infty[$$

- ornament (binary): first argument has duration 0

$$o : D^2 \rightarrow D, \frac{0}{q}, \frac{p}{q} \mapsto \frac{p}{q}$$

$$\text{or } o : I^2 \rightarrow I, [a, a[, [a, b[\rightarrow [a, b[$$

- fork : concurrent simultaneous voices

$$f : D^2 \rightarrow D, \frac{p}{q}, \frac{p}{q} \rightarrow \frac{p}{q}$$

$$\text{or } f : I^2 \rightarrow I, [a, b[, [a, b[\rightarrow [a, b[$$

- chord = special case of fork (?)

example: one 4/4 measure $t_0 = d_2(\dots)$.

For completeness, we may consider args with **multiplicity**, for the summation of siblings.

example (N = note): $d_4(N, N : 2, N)$ to represent a syncope, where the second argument has a double duration compared to the others.

For more examples, see

- MCM'19 <https://hal.inria.fr/hal-01988990> version ancienne et partielle de scole model (arbres de rythme, de syntaxe, serialization...)
- <https://gitlab.inria.fr/qparse/paper-sdags> exemples avec multiplicités

other optional operators

for performance indications, control, engraving stuff, annotations (lyrics, fingering...) etc.

All these symbols are unary (markups). Similarly as for musical events, the temporal position of these markups follows from the embedding time operators.

- articulation: *staccato* etc
- dynamics: *p*, *f*...
- ornament: *mordent*, *gruppetto*, *trill*...

- bar (special mark, in addition to above *split*)
- *jump...*
about control indications, possibly paired, cf. R. Dannenberg et al. ICMC 2012.
- fermata
- breath
- key change of time signature, key, clef...
- tempo
- ...

example: two 4/4 measures ($B_2 = \text{double-bar}$): $b(t_0, b(t_0, B_2))$.

example: two 2/4 measures with dynamics ($N = \text{note}$, $B = \text{blank}$, $P = \text{piano}$):

$$f(b(d_2(N, N), d_2(N, N)), d_2(d_2(B, P), d_2(B, B)))$$

We might also want to consider unary symbols in order to anchor explicitly an annotation to an event.

exemple: *fermata*.

"spanning" annotations (annotation with ha duration, like notes):

- dynamic: "spin" *cresc* or *dim*...
- slur
- trill (spanning)
- vibrato
- glissando
- volta
- octave change (+8 etc)
- pedal...

example: spin *cresc* during the 2 first beats of a 4/4/ bar ($F = \text{forte}$):

$$f(d_2(d_2(N, N), d_2(N, N)), d_2(CRESC, d_2(F, B)))$$

non-ambiguity

Ideally, we need to be able to build a unique representation of a given score, to act as a reference. In other words, we need a canonical representative for equivalent tree representation, where equivalence follows from equations like below.

$$d_2(d_3(x_1, x_2, x_3), d_3(x_4, x_5, x_6)) = d_3(d_2(x_1, x_2), d_2(x_3, x_4), d_2(x_5, x_6))$$

$$f(b(x_1, y_1), b(x_2, y_2)) = b(f(x_1, x_2), f(y_1, y_2))$$

Some equations are simplifying, like (C = continuation):

$$d_2(N, C) = N, \quad d_2(C, C) = C$$

for more details, see

- <https://hal.inria.fr/hal-01403982> (equivalent rhythm trees)
- <https://hal.inria.fr/hal-01138642>

For disambiguation, we may also limit the tuples to prime numbers.

$$d_4(x_1, x_2, x_3, x_4) = d_2(d_2(x_1, x_2), d_2(x_3, x_4))$$

linearization

for the linearization of these descriptors, see the nested words of Alur et al.

[Nested Words and Visibly Pushdown Languages](#)

page maintained by [Rajeev Alur](#) summary and biblio on newsted words and VPL

some docs in this repo:

- [NestedWords.md](#) (definitions of Alur et al)
- [RhythmNestedWords.md](#) (an example with rhythm trees)

languages

Several well studied formalisms (grammars and automata on tree or nested-words) exist for characterizing languages of syntactically well formed score representations.

(analogous of XML schemas ?)

Assuming a fixed language description (grammar or automaton), a score representation may be given directly as a run (of automaton) or derivation (of grammar), for simplicity.

exact search

Sequences of events, for building an index (e.g. n-gram based) can be extracted from nested-word representation, by projection.

Is it interesting to add the temporal markups (parentheses) in the index?

approximated search

Does the nested-word representation allow for tractable computation of a restricted Montgeau-Sankoff edit distance?