MIDI to Score Automated Drum Transcription

{ Martin Digard, Florent Jacquemard and Lydia Rodriguez-de la Nava } INRIA, France

Overview

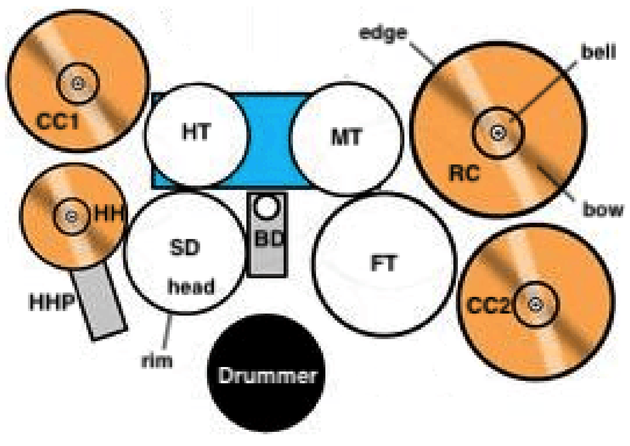
We propose our C++ and python library qparse, which is able to build a whole dataset of scores from a set of MIDI files.

We also propose an application on automated drum transcription, which is a special case of transcription of polyphonic instruments.

Drum Notation

Drums have a notation specific to each of the elements, and also the way that the drummer has to hit the drum or cymbal.

Elements of the drumkit



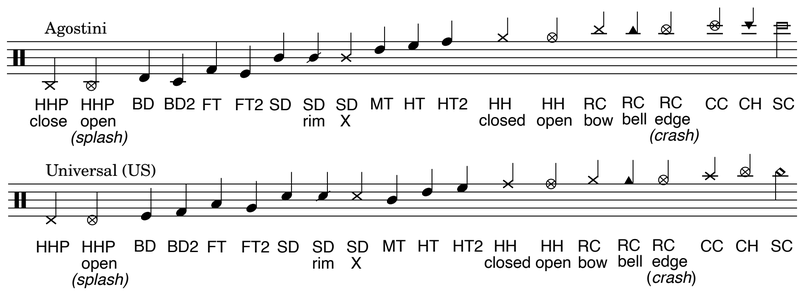


Figure 1: Pitches and note heads denoting the drum kit elements and modes for two differents notations: US and Agostini.

- Bass Drum (BD, also called *kick*),
- Snare Drum (SD),
- 3 toms, from left to right: High (HT), Medium (MT) and Floor (FT),
- Hi-Hat (HH): the two cymbals that can be joined (*open*) or disjoined (*closed*) using a pedal,
- *ride* (RC), hit on the *bow* (the body) or the *bell* (the top part) of the cymbal,
- *crash* (CC) or *splash* (SC), usually hit on the *edge* to produce an explosive sound.

Dynamics

- *accent*: indicates that a particular technique ought to be used, according to the related pitch,
- *ghost notes*: notes with low dynamics, but played firmly.

Ornaments

- *flam*: figure made of one grace note, played with lower intensity and slightly ahead the main note,
- *drag*: two grace notes are played with the same hand, ahead of a main note played by the other hand.

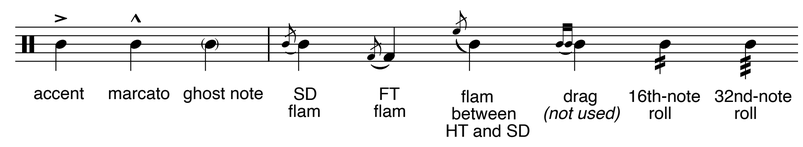
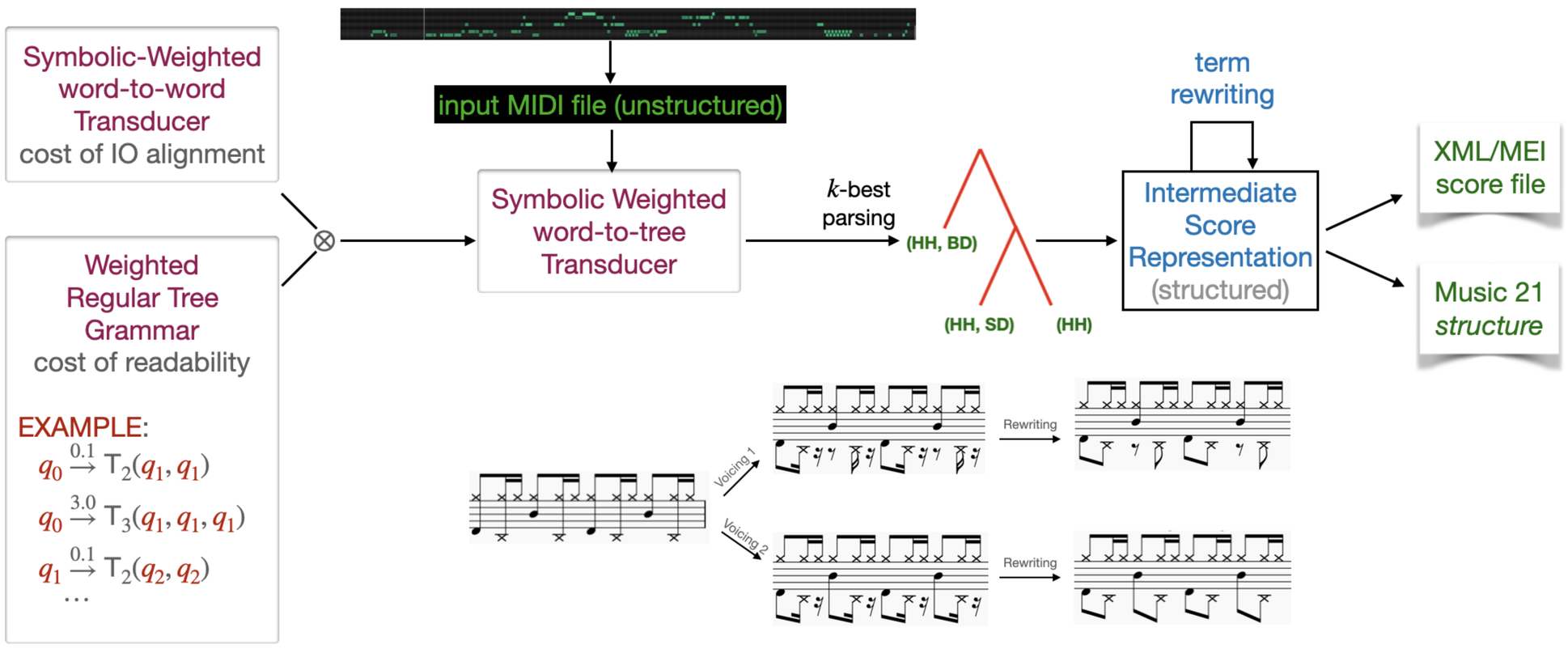


Figure 2: Accents, ghost notes, and flams.

Method: Transcription using Parsing Techniques [1]



Evaluation

In our experiment, we used the Magenta's Groove MIDI Dataset [2]:

- proposes 1,150 audio files (∼22,000 measures of drumming),
- diverse styles (e.g. funk, rock, jazz...),
- recorded from professional drummers performing on an electronic drum kit ROLAND T-11,
- recorded with a metronome, so no need for beat detection task.

Here is an example with one of the files from the Groove dataset. We have no ground truth to make a quantitative evaluation, so we can only procede to qualitative comparison between our result, and usual score writing softwares.

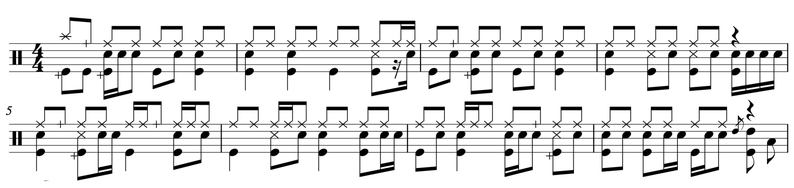


Figure 3: Transcription by our method.

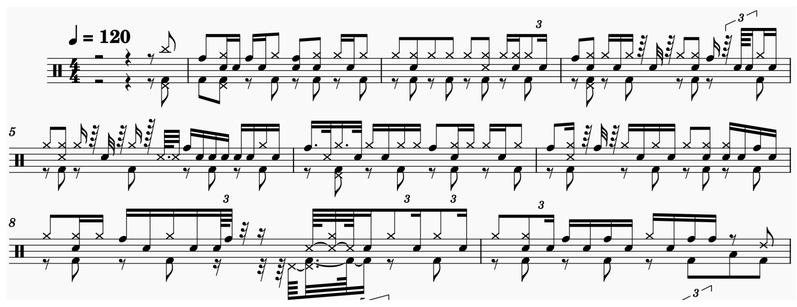


Figure 4: Transcription by MuseScore.

References

[1] Manfred Droste, Werner Kuich, and Heiko Vogler. *Handbook of weighted automata*. Springer Science & Business Media, 2009.

[2] Jon Gillick, Adam Roberts, Jesse Engel, Douglas Eck, and David Bamman. Learning to groove with inverse sequence transformations. In *International Conference on Machine Learning (ICML)*, 2019.

[3] Florent Jacquemard and Lydia Rodriguez-de La Nava. Symbolic Weighted Language Models, Quantitative Parsing and Automated Music Transcription. Springer, 2022.

Links

link to qparse on gitlab:
https://gitlab.inria.fr/qparse/qparselib

See our paper on our parsing technique [3].