



HAL
open science

MIDI To Score Automated Drum Transcription

Martin Digard, Florent Jacquemard, Lydia Rodriguez-de la Nava

► **To cite this version:**

Martin Digard, Florent Jacquemard, Lydia Rodriguez-de la Nava. MIDI To Score Automated Drum Transcription. DLfM 2022 - 9th International Conference on Digital Libraries for Musicology, Jul 2022, Prague, Czech Republic. hal-03847232

HAL Id: hal-03847232

<https://hal.science/hal-03847232>

Submitted on 10 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OVERVIEW

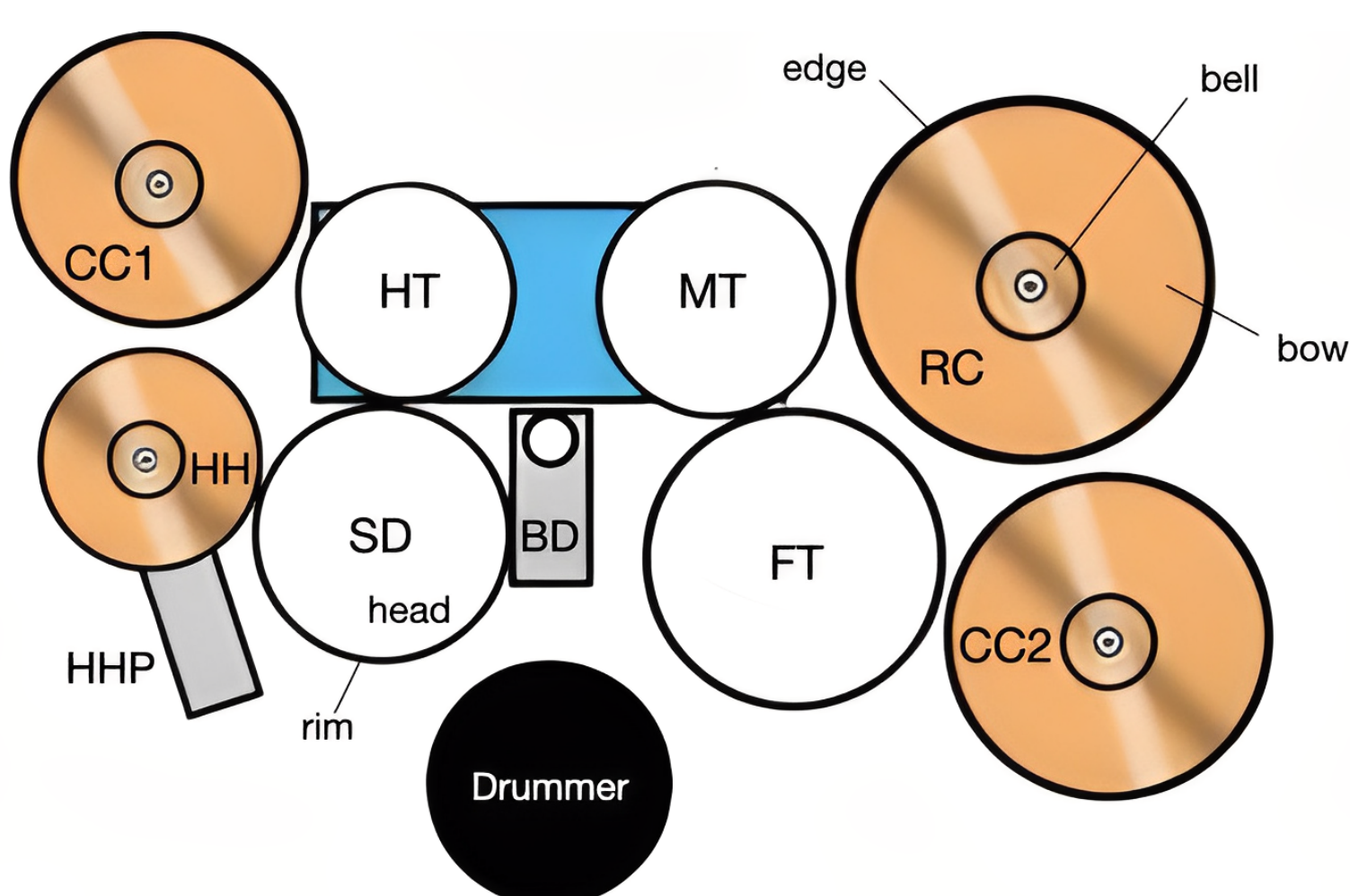
We propose our method for automated drum transcription, which is a special case of music transcription of polyphonic instruments.

We implemented this method in `qparse`, a C++ and python library able to build a whole dataset of music scores from a set of MIDI recorded performances.

DRUM NOTATION

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

Elements of the drumkit

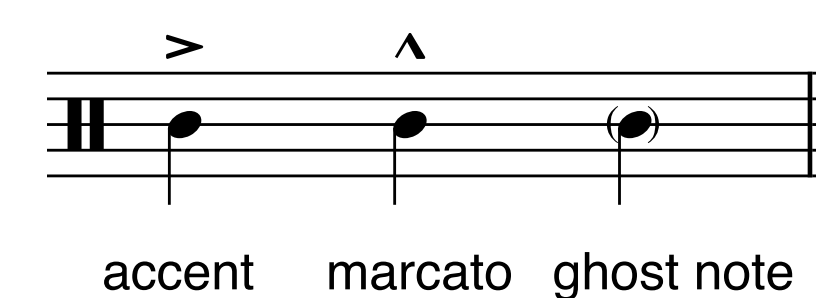


Pitches and note heads denoting the drum kit elements and modes for two different notations: US and Agostini.

- 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.

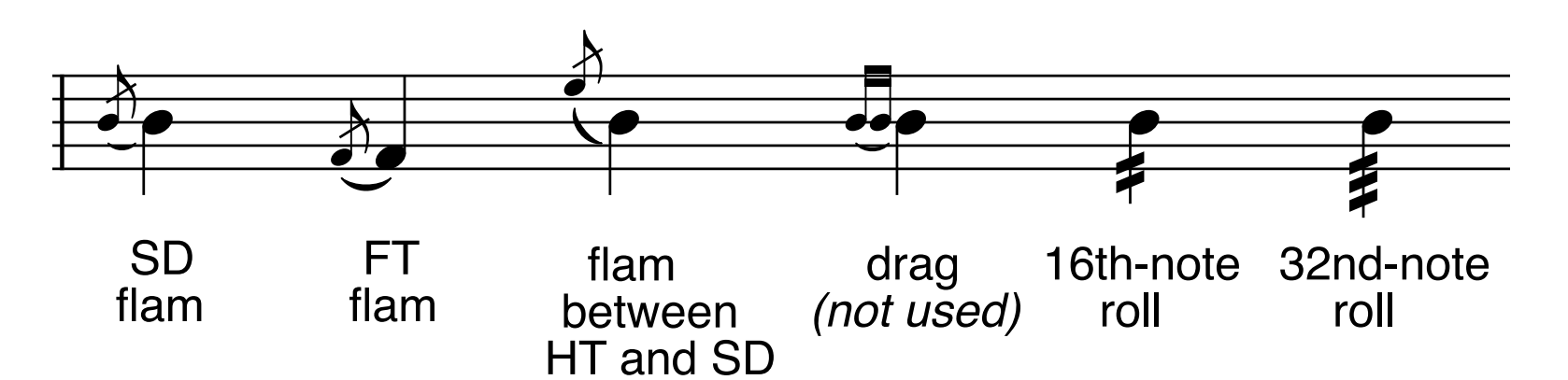
Dynamics

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



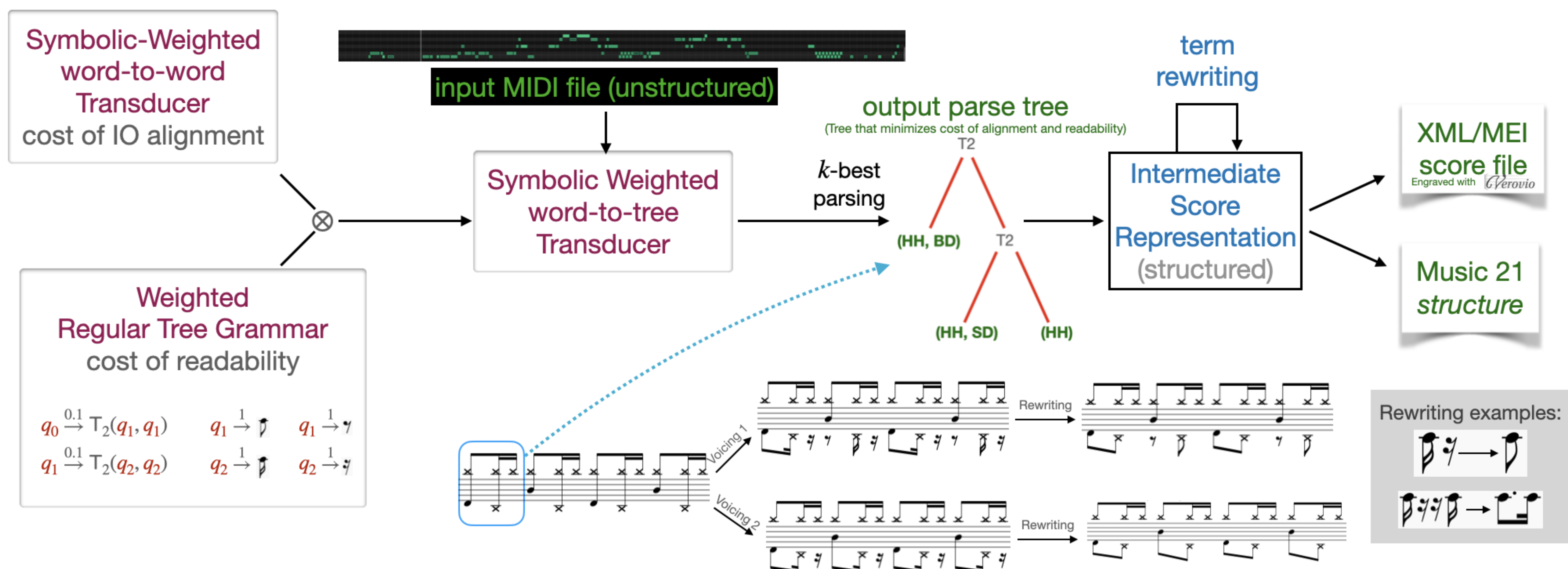
Ornaments

- *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.



TRANSCRIPTION USING WEIGHTED PARSING TECHNIQUES

In order to transcribe a MIDI file into drum notation, we use weighted parsing techniques [1]. At the beginning, a transducer calculates the cost of aligning the MIDI and the output score, and detects MIDI errors. Moreover, a weighted tree automata specifies the language of the preferred output notation [2]. Together, they perform the transcription as the search for an optimal parse tree. To make the output easy to read, we then proceed with some post-processing on the parse tree: separating voices (one tree per voice), and rewriting for minimal notation.

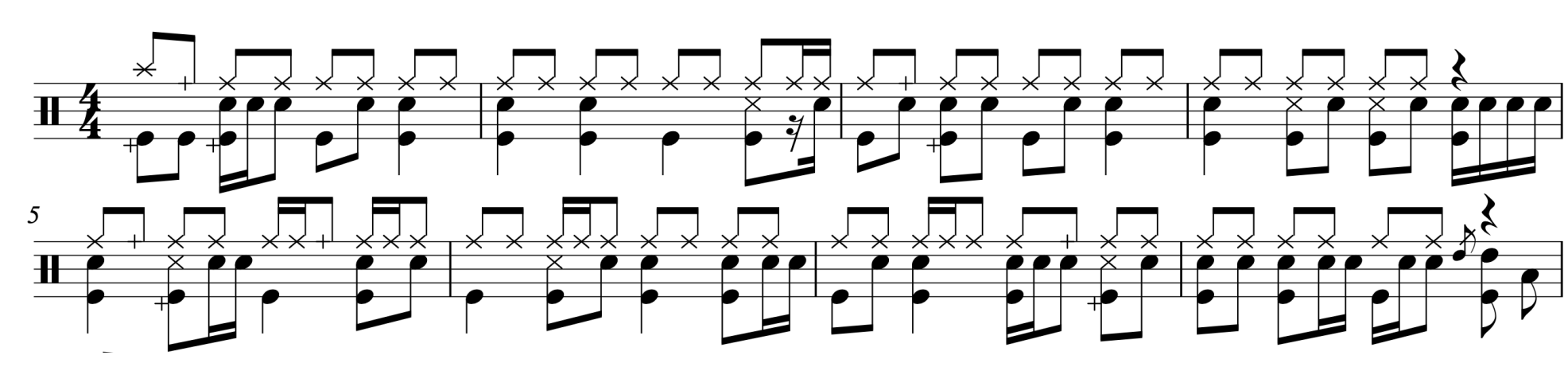


EVALUATION

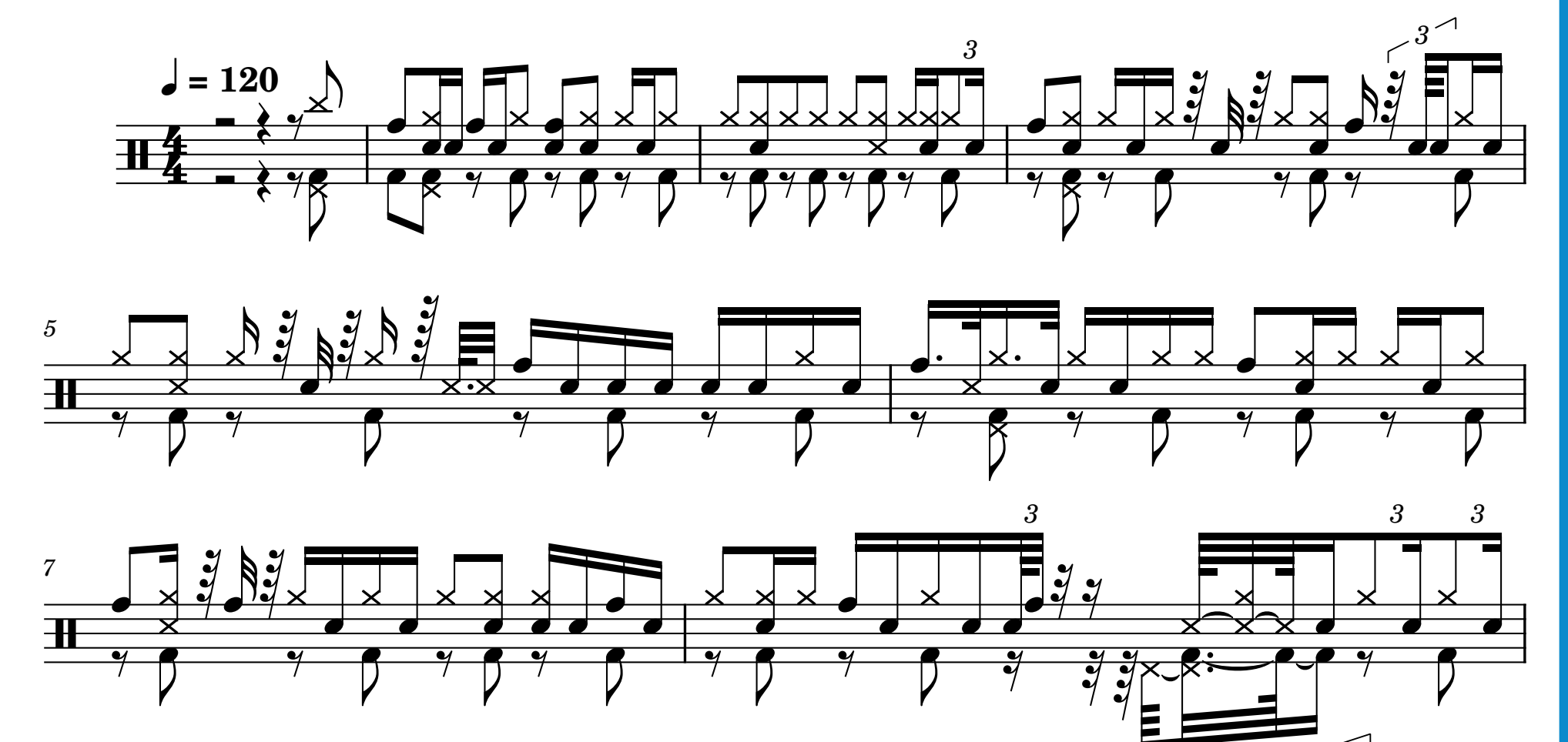
In our experiment, we used the Magenta's Groove MIDI Dataset [3]:

- 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: no need for a preliminary beat detection task.

Here is an example with one of the files from the Google's Groove MIDI Dataset. We have no ground truth to make a quantitative evaluation, so we can only proceed to qualitative comparison between our result, and MIDI output by score typesetting softwares.



Transcription by our method.



Transcription by MuseScore.

REFERENCES

- [1] Manfred Droste, Werner Kuich, and Heiko Vogler. *Handbook of weighted automata*. Springer Science & Business Media, 2009.
- [2] Florent Jacquemard and Lydia Rodriguez-de La Nava. *Symbolic Weighted Language Models, Quantitative Parsing and Automated Music Transcription*. Springer, 2022.
- [3] 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.

LINKS

Link to `qparse` on gitlab:
<https://gitlab.inria.fr/qparse/qparselib>

See more examples, and learn more about `qparse`:
<https://qparse.gitlabpages.inria.fr/>