

----- REVIEW 1 -----

PAPER: 25

TITLE: Encoding Matters

AUTHORS: Nestor Napoles Lopez, Gabriel Vigliensoni and Ichiro Fujinaga

Contribution: 2 (good)

Relevance: 2 (excellent)

Readability and paper organisation: 0 (minor improvements required)

Overall recommendation: 1 (accept)

Accept as short paper (full papers only): no

Accept as Unlocking Musicology Challenge paper (full and short papers only): yes

----- Overall recommendation -----

I congratulate the authors on taking on what is certainly an important issue to the MIR community in general and to the DLfM readership specifically.

Relevance

The paper's relevance is strong, though as a small point the language could perhaps be more closely pitched to the readership which I suspect is more expert in this area than this paper's exposition suggests.

Contribution

The contribution of this paper is less assured. Readers will be all too aware of the low quality of most online symbolic music corpora, and would no doubt be keen to develop a more systematic understanding of the relative strength / weaknesses of different corpora or file formats, but that is explicitly beyond the scope of the study. The abstract emphasises how these discrepancies 'can change the results of music analysis tools', and the paper points out the possibility of corpus encoding auto-correction, though no examples or solutions are provided in either case. As such, the provision of code for comparing scores is the strongest material contribution to the field.

Error or editorial?

Sections 3.2 and 3.3 should acknowledge that this division is not too clear cut. For instance, there are cases where the notation softwares can force errors of the kind here labelled as human, e.g. in handling tremolo symbols, one cannot always reproduce printed editions within the wellformedness constraints of file formats. Put another way, 'errors' of that kind are not always incorrect; this is an area of musical notation that is highly subject to editorial variance. Without knowing our transcribes sources, we cannot presume to comment on that. 'Discrepancy' is a well chosen word in this case; 'error' is not. As a small final point, the titles are awkward here: either compress to just 'human encoders' or explained to include 'Error related to ...'.

Readability and paper organisation

Mostly the writing and organisation is of good quality, though there is some inelegant/non-native language e.g. 'in most of the cases', 'In the other hand', and in the meandering final sentence.

----- REVIEW 2 -----

PAPER: 25

TITLE: Encoding Matters

AUTHORS: Nestor Napoles Lopez, Gabriel Vigliensoni and Ichiro Fujinaga

Contribution: 3 (excellent)

Relevance: 2 (excellent)

Readability and paper organisation: 0 (minor improvements required)

Overall recommendation: 2 (strong accept)

Accept as short paper (full papers only): no

Accept as Unlocking Musicology Challenge paper (full and short papers only): no

----- Overall recommendation -----

This paper convincingly demonstrates both the results of, and the value in, comparing

encodings of different symbolic music files. It is well-structured and makes excellent use of illustrative material to clarify the approach and findings. It is accessible to somebody like myself without a technical background, and the methodology also appears sound. The decision to anonymise both the music notation software and the translations (MEI and Humdrum) is particularly valuable, as it ensures that the focus remains on differences in encodings.

The paper concentrates exclusively on a single, very specific example, and therefore runs the risk of being relevant to a very narrow audience. Nevertheless, this approach is very successful, and these limitations are managed well. The tight focus allows the key issues to be explored in sufficient depth (without getting bogged down by a large amount of data and/or variants). In addition, the conclusion successfully points towards a much wider application of the findings, for example by identifying areas for auto-correction in the future. In addition, the limited focus makes this paper a good fit for the "short paper" category.

There were occasional slips in the English, which meant that the paper did not always read smoothly (for example, section 1, "in activities requiring to search or browse" and section 3.3, "In the other hand"). I would also suggest re-thinking the use of the word "tremolos" on p. 4 (section 3.3.2), as this is misleading. Tremolo normally refers to note values that are unmeasured (i.e. very, very fast), and is notated using three lines. The context of the discussion is repeated quavers, and the single line is a shorthand to denote them (see <http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/-omo-9781561592630-e-0000000023>). This is a small quibble and should be straightforward to revise (for example by replacing the word "tremolos" with "repeated quavers").

----- REVIEW 3 -----

PAPER: 25

TITLE: Encoding Matters

AUTHORS: Nestor Napoles Lopez, Gabriel Vigliensoni and Ichiro Fujinaga

Contribution: 2 (good)

Relevance: 2 (excellent)

Readability and paper organisation: 0 (minor improvements required)

Overall recommendation: 1 (accept)

Accept as short paper (full papers only): no

Accept as Unlocking Musicology Challenge paper (full and short papers only): no

----- Overall recommendation -----

This paper presents results from the comparison of generating encoded music from common music notation software as well as translations between three common encoding formats. While the results are not surprising, it is useful to see this done in a systematic ways. The authors also make useful suggestions to make export and translation more efficient, possibly automatically.

A couple of suggestions:

1. "it is typically possible—and easy—to generate a new rendering of the music score" I disagree, particularly with the "easy" part. The precision of automatic of rendering is often poor compared to a version typeset in a WYSIWYG environment. Let's say that it really depends on what the new rendering is for.

1. "but they can be encoded differently as to consider them equivalent." This sentence is fragmentary/unclear.

3.1 "this plot has been obtained by using music21" if this has been done for encodings translated to MEI too, then I worry about the quality of the data: music21 has issues with accidentals in MEI as they are not explicitly encoded if defined at the key signature. I would recommend to verify this and address it at the presentation and add a sentence about it before publication.