Music Notation Representation in the TMI

Paper presented at the HARMONICA forum, Athens, 14 March 1997

Frans Wiering
Information Science
Utrecht University

Introduction

The *Thesaurus musicarum italicarum* (TMI) is an initiative of the Department of Computer and Humanities at Utrecht University, the aim of which is to publish a cohesive electronic corpus of Italian music treatises from the second half of the sixteenth to the early seventeenth century. As a pilot, a CD-ROM will be published with the music treatises of Gioseffo Zarlino (1517-1590). The next stage will be to construct to a Web server with these and other treatises.

In the 'digital workplace' we have in mind as our goal, there will be a complete electronic facsimile of each document, but more importantly an edition in which, for example, references and quotations, names, places and terms, variants, uncertainties and editorial emendations are duly recorded. This is done by providing the documents with SGML-markup, in accordance with the guidelines of the *Text Encoding Initiative*. The result is not so much a printable document, as a ***document information system*** from which knowledge can be extracted through sophisticated or not so sophisticated queries.

The problem

These treatises contain many music examples. Essentially, the problem we face with these is, how we can present them in a ***musical document information system*** that can be used in similar ways as a textual system. Some aspects of this problem are:

1. TMI is about source information, i.e. about ***music notation***, not music as **sound** or as **logical information.**
2. The notation in the sources may contain elements that do not occur in Common Music Notation (CMN), for example mensural signs, proportions, and ligatures; or be of an altogether different kind (lute and keyboard tablatures).
3. Source information (damage, gaps, uncertainties, variants) and editorial emendations must be recorded, and more generally, it must possible to enrich the electronic transcription ***while leaving the original transcription intact.***
4. It should be possible to manipulate the digitized music notation, for instance by clef change, transposition, reduction, (re)barring, transcription to other type of notation.
5. It must be possible to query the music notation. Simple examples of such queries include:
    - where in the piece/corpus is a major third followed by a tritone?
    - where in the piece/corpus is a note written a second too high or low by mistake?

Apart from software which can perform the outlined and other tasks — which is a problem in itself that I will not discuss — an (internal) music representation is needed that can record the necessary information.

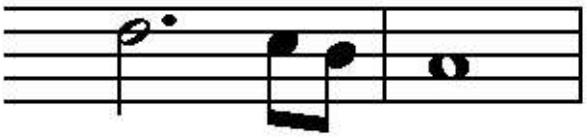
Most music representations are designed either for music printing (NIFF etc.) of for sound synthesis (MIDI). Only a minority offer other kinds of access. These are old-fashioned encoding systems for CMN such as DARMS and Plaine and Easy Code, or analytical representations such as ESAC.

In these systems, encoded music is represented as an unstructured 'musical ASCII', in which it impossible to record the structured information needed for an information retrieval system. On the other hand, SMDL is structured but does not address ***notation***, but rather the ***logical content*** of music.

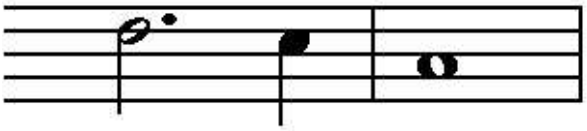
Towards a solution

Currently we are using as a temporary solution a combination of *Note Processor* DARMS and TEI markup. The example below illustrates this.

Source A:



Source B:



The example displays the variants of the same passage in a hypothetical composition known from two sources. In *Note Processor* DARMS, the fragments would be encoded as follows:

Source A:

```
7H. 6E( 5E) / 4W /
```

Source B:

```
7H. 6Q / 4W /
```

They can be integrated into one document by using TEI-markup (the tags between '<' and '>'), as in the following:

```
7H. <app><rdg wit=A> 6E( 5E)</rdg>
         <rdg wit=A> 6Q</rdg>
    </app>
/ 4W /
```

The **<app>** tag indicates an 'apparatus entry' consisting of several readings, each of which is suurounded by **<rdg>** tags the attribute wit indicating the 'witness' for each particular reading.

While essentially this is what is needed, this particular method has serious shortcomings because:

1. compatibility issues arise, notably concerning synchronisation of staves;
2. DARMS lumps together several notation symbols in one **token**;
3. standard DARMS is for CMN only, not for other kinds of notation;
4. the TEI tagset is not wholly adequate for music notation.

So in the end the only option seems to be to develop a new representation of music notation for the TMI. It will be in two parts:

1. A markup-system inspired by on the one hand TEI (source and editorial markup) and the other HyTime/SMDL (synchronisation) providing the general structure.
2. A 'musical ASCII' encoding system for each type of notation that occurs in the documents.

Conclusion

For this forum, I would like to emphasise that digitized music is more than an intermediate representation for the printing or playing of music. It possesses an information content that can be retrieved for scholarly purposes, and no doubt for many other purposes as well. retrieval presupposes adequate music representations which so far have hardly not emerged. The development of such representations must be encouraged.