A standard format proposal for hierarchical analyses and representations

David Rizo
Department of Software and Computing Systems, University of Alicante
Instituto Superior de Enseñanzas Artísticas de la Comunidad Valenciana
Ctra. San Vicente s/n. E-03690
San Vicente del Raspeig, Alicante, Spain
drizo@dlsi.ua.es

Alan Marsden
Lancaster Institute for the Contemporary Arts, Lancaster University
Bailrigg
Lancaster LA1 4YW, UK
a.marsden@lancaster.ac.uk

ABSTRACT

In the realm of digital musicology, standardizations efforts to date have mostly concentrated on the representation of music. Analyses of music are increasingly being generated or communicated by digital means. We demonstrate that the same arguments for the desirability of standardization in the representation of music apply also to the representation of analyses of music: proper preservation, sharing of data, and facilitation of digital processing. We concentrate here on analyses which can be described as hierarchical and show that this covers a broad range of existing analytical formats. We propose an extension of MEI (Music Encoding Initiative) to allow the encoding of analyses unambiguously associated with and aligned to a representation of the music analysed, making use of existing mechanisms within MEI's parent TEI (Text Encoding Initiative) for the representation of trees and graphs.

Keywords

Encodings, standards, music analysis, music representations

1. INTRODUCTION: DESIRABILITY OF AN ANALYSIS ENCODING STANDARD

Musical scholarship in the digital age requires not just digital representations of music, but also digital representations of data associated with that music. For some time, it has been recognised that standards are required for recording bibliographic and related metadata about music.[1] Less attention has been paid to the digital representation of analyses of pieces of music, despite the fact that a great deal of musical scholarship depends on the recording and communication of such commentary on pieces of music. This is evident of course in specialist academic publications dedicated to music analysis, but communication of analysis also has an important place in historical research, in pedagogy, and even in such humble publications as concert programme notes.

Standard text is generally a crucial part of music analyses, often augmented by specialist terminology, and frequently accompanied by some form of diagram or other graphical representation. There are well established ways of representing texts in digital form, but not for the representation of analytical diagrams. In some cases the diagrams are very like music notation, with just the addition of textual labels, brackets or a few other symbols (see Figure 1). In other cases they are more elaborate, and might not make direct use of musical notation at all. It is possible to digitise these analytical diagrams as images, just as it is possible to convert a score to a digital image, but with exactly the same disadvantages: the essential content of the diagram remains hidden in the image and is not exposed for explicit and unambiguous processing.

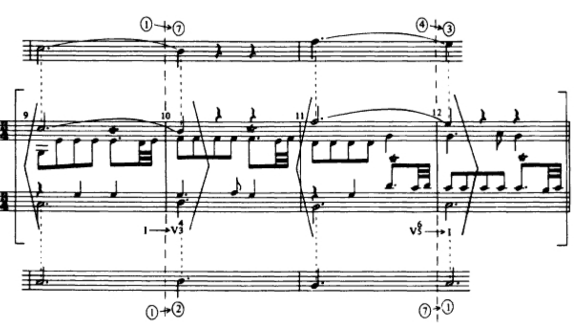


Figure 1: A Classic Turn of Phrase: Music and the Psychology of Convention by Robert O. Gjerdingen.

Recent developments in digital musicology make the development of standards for the representation of analytical information important. Much research in the field of Music Information Retrieval relies on collections of data for the development and testing of algorithms. Systems involving machine learning (which have now become standard fare in MIR) depend for their effectiveness on large sets of reliable data. To properly compare the performance of different systems designed to carry out similar tasks requires those systems to be supplied with the same sets of data. It is therefore becoming increasingly common for researchers to make sets of digital music-analytical data available. Examples include segmentation and chord-labelling data for Beatles songs and other popu-

DLfM August 12, 2016, New York, USA
ACM ISBN .
DOI:

[1] See for example the Variations projects at the Indiana University, http://www.dlib.indiana.edu/projects/variations3/

lar music,[2] segmentation data arising from the SALAMI project,[3] analyses in the style of Lerdahl and Jackendoff [5], Schenkerian analyses from text books and the like [15], and chord labelling and phrase structures in theme and variations.[4] While all of these give a clear definition of the structure of their data, they each use different representation schemes, bringing with them the common difficulties in reusing the data and potential for misinterpretations.

Advances in technology always have an impact on music, and we have seen a transformation in the ways in which music is produced, recorded and communicated in the past few decades. Advances which envisage yet more sophisticated digital processing of music are envisaged (automatic composition, variation, arrangement, etc.). In the recent past the concepts by which music-processing has been organised have been derived from the world of sound recording ('tracks', 'mixes' and the like). Composers and performers, however, tend to deal with concepts such as 'voice', 'phrase', 'chord' and so on, which are closer to the elements of music analysis. Future, properly 'musical', software will need to operate with this kind of concept and so data of that nature will need to be made available. In this respect, our concerns align with those who aim to give adequate computational definitions of these high-level musicological concepts [8].

Therefore, the same reasons which require standards for the digital representation of music also apply to the digital representation of music analyses: to preserve not just the images of analyses but also the knowledge embodied in them, to allow the sharing of data among researchers, and to facilitate automatic processing of that data. There is no limit to forms which analytical commentary on music may take, so it is unrealistic to seek standards for the representation of all kinds of analysis. In this project, we concentrate on those classes of analysis which are hierarchical, generally represented in the forms of trees and graphs.

2. HIERARCHICAL ANALYSIS AND REPRESENTATION

Hierarchies have had an important role both in musical analysis and, more recently, in representations of music in tasks such as computer assisted composition or music information retrieval. By 'hierarchical analysis' or 'hierarchical representation', we mean an analysis or representation which presents the structure of the music on different levels, in which higher levels correspond to larger spans of the music. Relations between different elements are shown, but only between neighbouring elements, either above or below in the hierarchy, or side-to-side from one 'branch' to another. An analysis which is based on the relationships of distant segments of music, such as a 'paradigmatic analysis' in the style of Nattiez [22], is not hierarchical according to our definition. While hierarchical analyses and representations are often presented as trees, it is not essential for them to be so, as will become clear from the examples below.

The two most famously hierarchical kinds of music analysis are those of Schenker and of Lerdahl and Jackendoff. The well known Generative Theory of Tonal Music (GTTM) [17] proposes several viewpoints for analyzing tonal music based on generative linguistic theory, i.e., modeling analysis using formal grammars (see figures 2, 10, 11, 12, and 14). Although the theory is presented as a set of rules, and the results are presented in tree diagrams, the analyses are not so systematic as this might suggest. For discussion of some of the difficulties concerning polyphony and uncertainties over what information is maintained from one level to the next, see [20]. Other issues are discussed in [5] which also demonstrates some of the difficulties over systematic treatment of Lerdahl and Jackendoff's 'preference rules'.

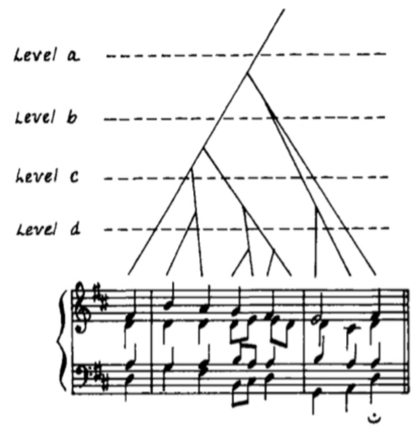
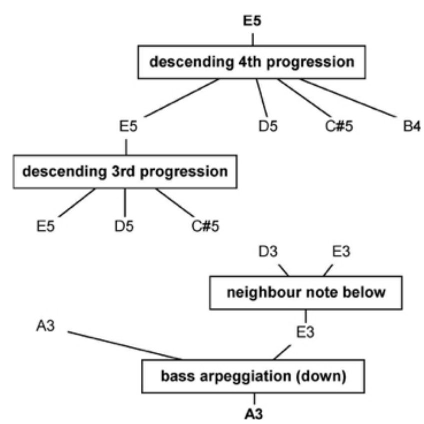


Figure 2: Generative theory of tonal music (GTTM) reduction at different levels (extracted from [17], figure 5.8)

(a) Schenker's analysis of the first phrase of Mozart's piano sonata in A, K.331 (Schenker, 1935, Fig. 157)



(b) Tree representation derived from Schenker's analysis

Figure 3: Marsden's tree representation Schenker's analysis as proposed in [20] (figures 5 and 6).

While the analyses of Heinrich Schenker were not presented in trees, levels of hierarchy are explicit, and the analyses can (with the loss of some information) be converted to trees, as demonstrated in [20] (see figure 3). Further systematisations of Schenkerian analysis are found in [19, 21] and [15]. The latter presents an analysis not as a tree but as a *Maximal outerplanar graph* (MOP), a representation for Schenkerian analyses introduced by Yust [27] (see

[2] http://isophonics.net/content/reference-annotations

[3] http://ddmal.music.mcgill.ca/research/salami/annotations

[4] http://u.osu.edu/tavern/

Figure 16).[5]

Less obviously, traditional tonal-harmonic analysis of music is also hierarchical: spans of music are identified as being in a key, which are identified as subordinate to longer spans which might be goverened by a different, related, key. This is most evident in the 'keyscale plots' presented by Sapp [25], whose triangular shape shows the structural similarities to the reduction trees of Lerdahl and Jackendoff (see Figure 4).

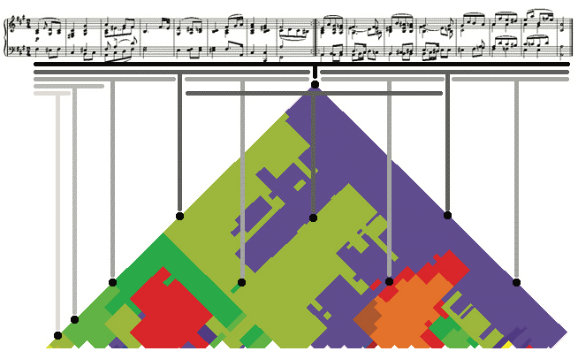


Figure 4: Sapp's keyscape plot (extracted from [25], figure 3.14)

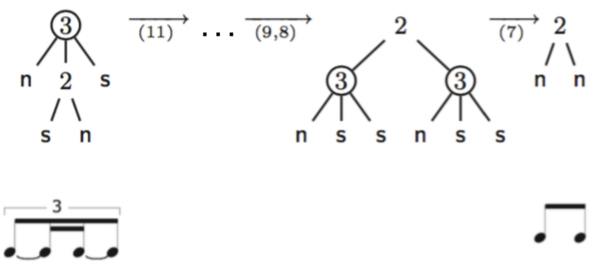
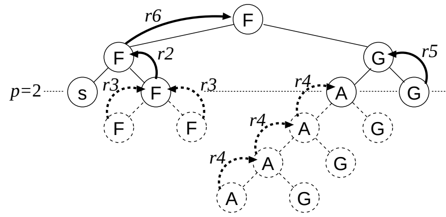


Figure 5: Tree rewrite sequence proposed by Jacquemard et al. [13] (figure 7)

The way music content is represented determines what an algorithm can do with it for analysis, classification, retrieval or composition (see [10]), and in this domain also hierarchical representations have proven useful. For automatized composition, Högberg et al., in the system *Wind in the willows*, proposed the use of tree transducers [9]. The same hierarchical structure type was used by Jacquemard et al. [13] for elegantly representing and transforming rhythm notation (see Figure 5). With the objective of measuring the similarity of music, *metric tree* structures were introduced by Rizo [24] (see Figure 6) and extended with k-testables by Bernabeu et al. [1]. With the same purpose, Pinto and Tagliolato [23] proposed a graph representation of melodic sequences (see Figure 7). Finally, several syntactic analysis systems have been proposed that produce parse trees such as the work by Gilbert and Conklin [3] for melody reduction, the effort by Bod [2] for formalizing the perception of phrase structures (see Figure 8), or the syntactic perspective of simple rhythm structures by Lee [16].

[5] Yust argues that a graph better represents elaborations which depend on the melodic interval from one note to another; for the same reasons a graph representation was proposed in [19] but subsequently abandoned as excessively complex in computational terms.

(a) Source melody.



(b) Propagation and prunning.

Figure 6: Rizo's metric trees propagation and prunning [24] (figure 3.28).

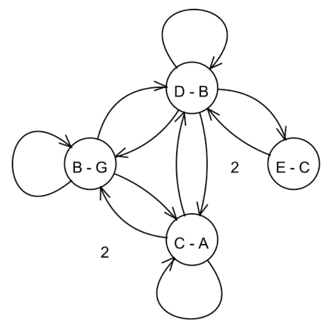


Figure 7: Pinto's graph encoding of a sequence. Figure 6 in [23]

3. EXISTING HIERARCHICAL REPRESENTATIONS

Many of the approaches for hierarchical analysis and representations introduced above have points in common that could foster interesting synergies, but these are currently difficult to achieve due to the different representations and encodings of, nevertheless, similar concepts. A first humble attempt to normalize some representations was made by Rizo [24] (see figures 3.11 and 3.14 in that work) where standardized tree representations were offered to model different melodic reduction strategies such as the GTTM [17] and Marsden's elaboration trees [20]. However, no proposal for encoding datasets of this kind of tree structures was offered. Thus, the field of potentially useful standardization of many different ad-hoc encodings of similar musical-analysis and intermediate representation remains open.

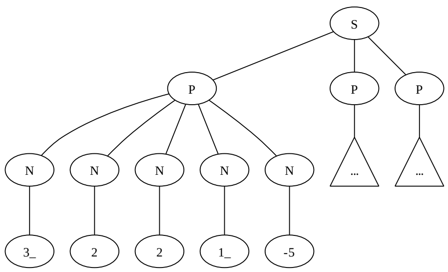
To illustrate the issues, we discuss the case of the SCHENKER41 database of analyses by Kirlin [14]. This work of Kirlin's is remarkable and valuable. Each analysed work is encoded in two files: the music itself is encoded in a MusicXML file [4], and the analysis is encoded in a text file with extension `.analysis` (see Figure 16). No grammar to read the format of these files is given, and in the paper [14] or the section 4.2 of his PhD thesis [15], only the encoding of the notes based on scientific pitch and the prolongation operation are described. Once the developer opens the `.analysis` file, he/she has to figure out the meaning of some of the elements not described in the paper. In [15], Kirlin says that "The text file description is more relaxed than the MOP representation to allow for easy human creation of analyses." But SCHENKER41 is proposed as a corpus for use in other music analy-

```
(3_221_-5)(-533221_-5)(13335432)(13335432_)(3_221_-5_)
```

(a) Bracket representation for folksong K0029, "Schlaf Kindlein feste"

```
S(P(N(3_)N(2)N(2)N(1_)N(-5))P(...)P(...)  ...  )
```

(b) Labeled bracket representation for (first five notes of) the structure in figure above



(c) Graphical representation of parenthesized notation above as found in [24] (figure 3.6)

Figure 8: Example of data used to learn Bod's grammar in [2] (figures 7 and 8).

sis applications, which means that its format must also be readable by other computer software.

Another notable effort is that of the corpus introduced by Hamanaka et al. [6] containing 300 musical pieces analyzed using the GTTM paradigm.[6] For encoding music, the MusicXML [4] format was used. The different GTTM analyses (grouping, metrical, time-span, prolongational, and harmonic) were encoded using several ad-hoc XML grammars (see Figure 9).

```
▼<tstree>
  ▼<ts timespan="21.0" leftend="0.0" rightend="21.0">
    ▼<head>
      ▼<chord duration="3.0" velocity="90">
          <note id="P1-7-1"/>
        </chord>
      </head>
    ▼<at>
      ▼<temp difference="0.0" stable="unknown">
          <pred temp="-inf"/>
          <succ temp="+inf"/>
        </temp>
      </at>
    ▼<primary>
      ▼<ts timespan="9.0" leftend="12.0" rightend="21.0">
        ▼<head>
          ▼<chord duration="3.0" velocity="90">
              <note id="P1-7-1"/>
            </chord>
          </head>
        ▼<at>
          ▼<temp difference="0.0" stable="unknown">
              <pred temp="-inf"/>
              <succ temp="+inf"/>
            </temp>
          </at>
        ▼<primary>
          ▼<ts timespan="3.25" leftend="17.75" rightend="21.0">
```

Figure 9: GTTM Time-span encoding extract Hamanaka et al. [6]

All these encodings share the same drawbacks. When music content is kept in a file different from that of the intermediate or analysis representation, tricks to unambiguously link analysis to music are devised that in many cases complicate the parsing of the format (see Figure 16). In cases where musical content and analysis is recorded in the same file, usually the original musical content is modified to accommodate the specifics of the representation (see Figure 6 where long notes are split), and, even worse, only minimal content is kept, discarding what the representation does not need but could be useful in the future.

The main problem that arises from having so many different encodings comes from the fact that in order to represent similar concepts different parsers have to be built with nuances that are not always cross-compatible.

In the following sections a standard encoding proposal is offered that tries to overcome these problems.

4. REQUIREMENTS OF THE ENCODING

Our current proposal has been designed observing much of the work presented in previous sections. In general the encoding introduced in this work uses tree and graph structures to encode the different analysis types taking into account that analysis and music content will be recorded in the same file. As stated above, it would be presumptuous to try to encode any kind of hierarchical analysis. Thus, although just some of the most representative types of analysis are taken into account, some extension mechanisms are devised to accommodate new structures with little effort.

Many segmentation analyses and elaboration reductions are usually represented using nested slurs (see Figure 10). As those slurs can be represented hierarchically, tree structures will be used for encoding this kind of analysis.

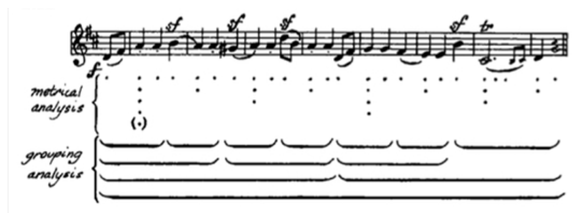


Figure 10: Metrical and grouping analysis proposed in GTTM [17] (figure 2.12)

GTTM explicitly prohibits the crossing of branches and avoids gaps between them. The current encoding is not only designed to represent current GTTM, but new possible approaches. Besides, if the proposed format is used to encode examples in books where these prohibitions have to be drawn, we cannot discard the possibility of crossing branches and gaps.

Apart from the hierarchical structure itself, one of the most important element in those structures is the node. It encodes most of the semantics of an analysis. The meaning of a node varies, depending on the approach. Some examples are Marsden's elaborations in Figure 3, the *left* and *right* branching in GTTM (see Figure 11 *a* and *b* respectively) that denote which note is more important, or the music content embedded in Rizo's metric trees (see Figure 6) or Pinto's graphs (Figure 7) and Yust's MOP (Figure 16). Sometimes the original graphical representations have to be interpreted as being node information. In GTTM, for example, cadences that are represented by using ellipses that connect branches (see Figure 12) can be encoded as node labels. In any case, nodes must be able to contain both analysis information and links to the referred musical events.

Another essential feature of any analysis encoding is the ability to represent simultaneous alternative analyses of the same musical event or segment. Figure 12 shows different reduction approaches to the same extract of music. In addition to concurrent analysis, annotations must be allowed to record alterntive explanations, whether originating from an analyst or generated by a com-

[6]http://gttm.jp/gttm/

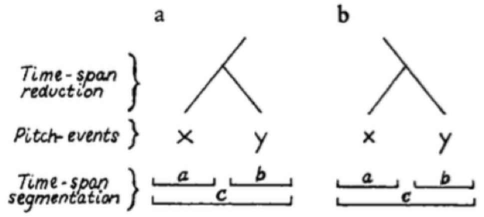


Figure 11: Time span segmentation in GTTM [17] (figure 6.6)

puter algorithm. It is not uncommon for analysts to disagree about the interpretation of a particular passage, or even for a single analyst to recognise that there are alternative interpretations. A large quantity of the literature about music analysis is concerned with the discussion of such cases, so it is essential to be able to represent alternative analyses in a way which allows accurate comparison.

Sometimes not everything in a representation can be strictly encoded in a tree. In some cases links are made between the branches of a tree (causing the structure to approach a graph), as proposed in [20] and implemented in [21]. The 'retained cadence' concept of Lerdahl and Jackendoff can also be considered as a case of such linking (see the ellipse in Figure 12).

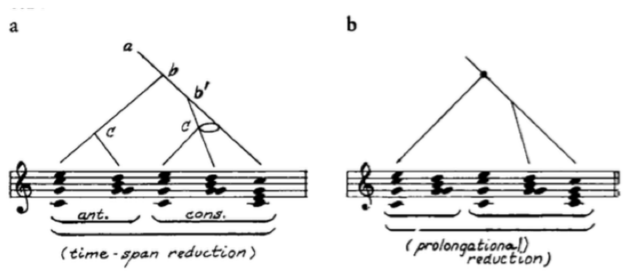


Figure 12: Two different types of reductions of the same musical segment in GTTM [17] (figure 8.14)

It can be argued that all hierarchies can be encoded using directed graphs, and tree structures can be avoided for the sake of unification of data types. We have discarded this approach because the abstract data type tree has many specific properties that facilitate efficient processing and so it is worth retaining. An example is simple branch pruning operations (see Figures 2 and 6) which are simple to perform on a tree but complex if done on a graph.

On the other hand, some analyses and intermediate representations require full graph encodings as previously indicated (Figures 7 and 16). These could be encoded as trees by unfolding the graphs and using the mechanism for representing linking between branches as mentioned above, but the semantics of the graph itself would be lost. We therefore consider it important to be able to use a graph representation when appropriate.

Finally, following the principle of encoding music and analysis together, the result of analytical processes such as reductions (see Figure 14), must be recorded in the same file as well.

5. MUSIC ENCODING INITIATIVE CUSTOMIZATION

Several formats have been proposed so far to standardize musical content [26]. However, to the best of our knowledge only one work [11] has pursued the standardization of analytical processes. However this work is focused on a different scope and lacks the embedding of music and analysis together.

The XML format named Music Encoding Initiative (MEI) [7], based on the successful experience of the Text Encoding Initiative (TEI) [12], and able to encode music of different notations, such as mensural and common western notation, together with metadata for bibliographic cataloguing, seems the most promising format to be used for our proposal.

MEI has an analysis module available. It has two parts, the first one focused on relating elements inside the musical score, the second one allowing the encoding of event-based analytical information such as melodic and pitch-class analysis. Being founded on an XML format, MEI is inherently hierarchical. If the hierarchical analysis matches the beginning and ending of work divisions, voices or measures (see Figure 13), the XML structure itself could be used to encode analyses. However, this is not always the case (see Figure 1). It is therefore necessary to make use of some other mechanism to overcome the limits of the XML format structure.

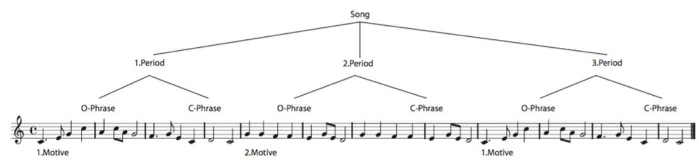


Figure 13: Form analysis from [18].

The MEI initiative is conceived to accommodate any kind of musical content, for which it has been equipped with a extension and customization mechanism [7]. This device enables the incorporation of our needs, taking advantage of all MEI infrastructures. In particular, we have reused all musical content elements from MEI and some of its analysis module that will be used to encode the musical content of nodes and the results of analytical processes such as reductions (e.g., the bottom system in Figure 14). MEI does not have components for explicitly encoding tree or graph structures but its parent format, TEI, includes such structures in the *net* module that we will adapt to fulfil our needs.

All requirements expressed in the previous section have been adopted into our proposal. The result is a self-contained music and analysis format. As we cannot give here a comprehensive explanation of all its elements, two representative examples are offered.

The listing in Figure 15 shows an extract of the anacrusis measure of the GTTM reduction in Figure 14. Only essential elements for understanding the example have been kept in the XML. Both music and analysis are encoded in the same file: the music extract is shown in the listing inside a MEI `<measure>` element; the analysis is encoded using an `eTree` element, adapted from the *net* module of TEI. In this example, the tree analysis contains a GTTM prolongation reduction as shown in the `type` attribute of the `eTree` element. `eTree` contains a `label` with information relevant to the semantics of the root node of the subtree it is representing. In this example only *elaboration* nodes are shown, but many other kinds may be represented depending on the analysis paradigm. Tree leaves are represented using `eLeaf` elements, which formally are nodes as well, containing a `label` element too. In our example, the leaf is representing a time slice that spans the anacrusis measure. This time span is represented using a `timespan` element introduced by us that contains attributes for setting the measure the segment starts (`startid`) and metrical positions to show the beginning (`tstamp`) and ending (`tstamp2`) of the time span following the MEI guidelines. Some other combinations of MEI attributes that unambiguously identify timespans or individual events have been used in other encodings. In addition, this label contains a link to the notes in the musical part of the encoding using the `sameas` attribute from the analysis module and `xml:id` attributes that anchor the target notes. Finally, a product of the analysis, in this case a reduction or music summarization, is emitted and recorded using the `summary` element inside the same tree label.

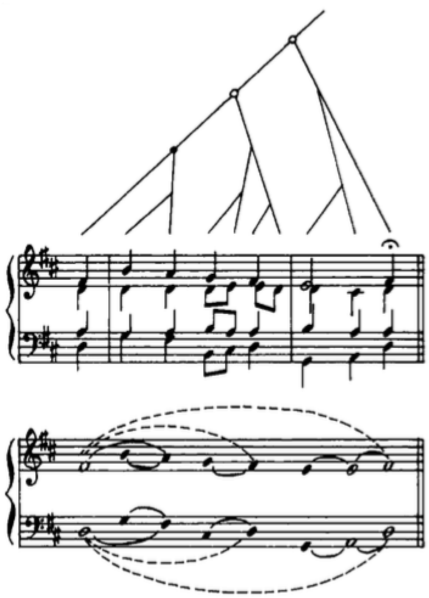


Figure 14: Bottom part contains the result of the GTTM reduction depicted in the tree above (extracted from [17], figure 8.31)

Another example, this time encoding a graph, is shown in the listing in Figure 17. The same principles hold for encoding the music content. Now the structural elements of the analysis, all of them adapted from the *net* module of the TEI format, are the `graph`, `node`, and `arc` elements. The `node` encoding follows the same principles as the `eLeaf` in the case of trees. The `arc` represent the edges between nodes and contain the semantics of the edge, in this case the type of MOP elaboration, in the `label` element.

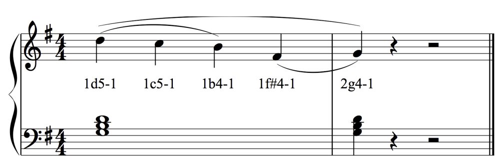
6. CONCLUSIONS

If format standards did not exist digital libraries would not be possible, and all possible research based on data and its further application to methods useful for society. This condition, while obvious, is not always met. There are formats for encoding music content, either notated or played, metadata describing content or the source of music documents, and some standards allowing the encoding of critical apparatus. However, few standards exists yet for the encoding of the analysis of musical works. It seems that in the near future, methods in digital musicology will not only require standard formats to encode music and metadata, but also standard representations of the output of analytical processes, both as a means for sharing results and as an intermediate step for building more complex systems that have as input the yield of music analysis methods.

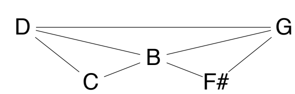
In this work we have introduced a proposal for a standard for the encoding of hierarchical analyses and intermediate representations of music that fulfils these needs.

```
<mei meiversion="2013">
    <!-- ... removed for saving space ... -->
    <measure xml:id="meas0" type="anacrusis">
        <staff n="1">
            <layer n="1">
                <note xml:id="n1" dur="4" oct="4" pname="f" accid="s"/>
            </layer>
            <layer n="2">
                <note xml:id="n6" dur="4" oct="4" pname="d"/>
            </layer>
        </staff>
        <staff n="2">
            <layer n="3">
                <note xml:id="n13" dur="4" oct="3" pname="a"/>
            </layer>
            <layer n="4">
                <note xml:id="n20" dur="4" oct="3" pname="d"/>
            </layer>
        </staff>
    </measure>
    <!-- ... removed for saving space ... -->
    <eTree type="gttm-prolongations-reduction">
        <label type="elaboration" subtype="left"/>
        <eTree>
            <label type="elaboration" subtype="right"/>
            <eLeaf xml:id="timeSlice0">
                <label>
                    <timespan startid="#meas0" tstamp="4" tstamp2="0m+5"/>
                    <note sameas="#n1"/>
                    <note sameas="#n6"/>
                    <note sameas="#n13"/>
                    <note sameas="#n20"/>
                    <summary>
                        <note sameas="#n1"/>
                        <note sameas="#n20"/>
                    </summary>
                </label>
            </eLeaf>
            <!-- ... removed for saving space ... -->
```

Figure 15: Extract of encoding of analysis in Figure 14 (anacrusis measure)



(a) Melodic sequence with note identifiers.



(b) MOP representation of this music.

```
1d5-1 (1c5-1) 1b4-1
(1f#4-1) 2g4-1
1d5-1 (1b4-1) 2g4-1
```

(c) An encoding of this MOP representation.

Figure 16: MOP encoding as found in SCHENKER41 dataset (figures extracted from [14], figures 1, 3, and 2).

```xml
<mei meiversion="2013">
    <!-- ... removed for saving space ... -->
    <measure xml:id="meas1" n="1">
        <staff n="1">
            <layer n="1">
                <note xml:id="n1" dur="4" layer="1" oct="5" pname="d"/>
                <note xml:id="n2" dur="4" layer="1" oct="5" pname="c"/>
                <note xml:id="n3" dur="4" layer="1" oct="4" pname="b"/>
                <note xml:id="n4" dur="4" layer="1" oct="4" pname="f" accid="s"/>
            </layer>
        </staff>
        <!-- ... removed for saving space ... -->
    </measure>
    <!-- ... removed for saving space ... -->
    <graph type="MOP">
        <node xml:id="vD"> <!-- v = vertex -->
            <label>
                <note sameas="#n1"/>
            </label>
        </node>
        <node xml:id="vC">
            <label>
                <note sameas="#n2"/>
            </label>
        </node>
        <node xml:id="vB">
            <label>
                <note sameas="#n3"/>
            </label>
        </node>
        <arc from="vD" to="vC">
            <label type="elaboration" subtype="elaboratedNode" value="vC"/>
        </arc>
        <arc from="vB" to="vC">
            <label type="elaboration" subtype="elaboratedNode" value="vC"/>
        </arc>
        <!-- ... removed for saving space ... -->
```

Figure 17: Extract of encoding of Figure 16 (first three notes)

Our proposal has recently been presented at the Music Encoding Initiative Conference 2016 in Montréal. As a proposal, it is being steadily improved and the final elements and attributes may slightly change in the near future, specifically, a more semantic approach to tag and attribute structures may be adopted. Nevertheless, the core idea and principles are expected to be maintained.

Complete details, examples, and the customization files and grammars can be found online and updated at http://grfia.dlsi.ua.es/cm/worklines/mei-ha.

7. ACKNOWLEDGEMENTS

We would like to thank Eleanor Selfridge-Field, Don Byrd, and Tom Collins for contributions during the MEC'2016 conference which have informed this work, and to thank Ichiro Fujinaga and Perry Roland for their encouragement to continue it.

This work has been supported by the Spanish Ministerio de Economía y Competitividad through project TIMuL (No. TIN2013-48152-C2-1-R, supported by UE FEDER funds).

References

[1] Bernabeu, J. F., Calera-Rubio, J., Iñesta, J. M., and Rizo, D. (2011). Melodic Identification Using Probabilistic Tree Automata. *Journal of New Music Research*, 40(2):93–103.

[2] Bod, R. (2002). A General Parsing Model for Music and Language. *Music and Artificial Intelligence*, pages 77–90.

[3] Gilbert, E. and Conklin, D. (2007). A probabilistic context-free grammar for melodic reduction. *International Workshop on Artificial Intelligence and Music at IJCAI-07. Twentieth International Joint Conference on Artificial Intelligence, Hyderabad, India*.

[4] Good, M. and Actor, G. (2003). Using MusicXML for file interchange. *Web Delivering of Music, International Conference on*, 0:153.

[5] Hamanaka, M., Hirata, K., and Tojo, S. (2006). Implementing "A Generative Theory of Tonal Music". *Journal of New Music Research*, 35(4):249–277.

[6] Hamanaka, M., Hirata, K., and Tojo, S. (2014). Musical Structural Analysis Database Based on GTTM. In *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*, pages 325–330.

[7] Hankinson, A., Roland, P., and Fujinaga, I. (2011). The Music Encoding Initiative as a Document-Encoding Framework. *ISMIR*.

[8] Harley, N. and Wiggins, G. A. (2015). An ontology for abstract, hierarchical music representation. ISMIR. late-breaking paper, http://ismir2015.uma.es/LBD/LBD31.pdf.

[9] Högberg, J. (2005). Wind in the Willows - Generating Music by Means of Tree Transducers. In *CIAA*, pages 153–162.

[10] Höthker, K., Hörnel, D., and Anagnostopoulou, C. (2001). Investigating the Influence of Representations and Algorithms in Music Classification. *Computers and the Humanities*, 35(1):65–79.

[11] Humphrey, E. J., Salamon, J., Nieto, O., Forsyth, J., Bittner, R. M., and Bello, J. P. (2014). JAMS: A JSON Annotated Music Specification for Reproducible MIR Research. In *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*, pages 591–596.

[12] Ide, N. M. and Sperberg-McQueen, C. M. (1995). The TEI: History, goals, and future. *Computers and the Humanities*, 29(1):5–15.

[13] Jacquemard, F., Donat-Bouillud, P., and Bresson, J. (2015). A Structural Theory of Rhythm Notation Based on Tree Representations and Term Rewriting. In *Advances in Artificial Intelligence*, pages 3–15. Springer International Publishing, Cham.

[14] Kirlin, P. B. (2014a). A data set for computational studies of Schenkerian analysis. In *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*, pages 213–218.

[15] Kirlin, P. B. (2014b). *A probabilistic model of hierarchical music analysis*. PhD thesis, University of Massachusetts Amhers, Massachusetts.

[16] Lee, C. S. (1985). The rhythmic interpretation of simple musical sequences: Towards a perceptual model. In West, R., Howell, P., and Cross, I., editors, *Musical Structure and Cognition*, pages 53–69. Academic Press, London.

[17] Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press.

[18] Linster, C. (1992). On analyzing and representing musical rhythm. In Balaban, M., Ebcioğlu, K., and Laske, O., editors, *Understanding music with AI: perspectives on music cognition*, pages 414–427. AAAI Press/MIT Press, Menlo Park, California, USA.

[19] Marsden, A. (2001). Representing Melodic Patterns as Networks of Elaborations. *Computers and the Humanities*, 35(1):37–54.

[20] Marsden, A. (2005). Generative structural representation of tonal music. *Journal of New Music Research*, 34(4):409–428.

[21] Marsden, A. (2010). Schenkerian Analysis by Computer: A Proof of Concept. *Journal of New Music Research*, 39(3):269–289.

[22] Nattiez, J.-J. (1975). *Fondements d'une sémiologie de la musique*. Union Générale d'Éditions.

[23] Pinto, A. and Tagliolato, P. (2008). A generalized graph-spectral approach to melodic modeling and retrieval. In *MIR '08: Proceeding of the 1st ACM international conference on Multimedia information retrieval*, pages 89–96, New York, NY, USA. ACM.

[24] Rizo, D. (2010). *Symbolic music comparison with tree data structures*. PhD thesis, Universidad de Alicante.

[25] Sapp, C. S. (2011). *Computational methods for the analysis of musical structure*. PhD thesis, Stanford University.

[26] Selfridge-Field, E. (1997). *Beyond MIDI: The handbook of musical codes*. Beyond MIDI: The handbook of musical codes. MIT Press, Cambridge, Massachusetts, USA.

[27] Yust, J. (2009). The Geometry of Melodic, Harmonic, and Metrical Hierarchy. In *Mathematics and Computation in Music*, pages 180–192. Springer Berlin Heidelberg, Berlin, Heidelberg.