The Implications of Digital Music Libraries for Music Theory

Society for Music Theory 2003, Madison

Eric Isaacson and Brent Yorgason

Indiana University

- In the 1960s, Allen Forte wrote a computer program that could perform simple segmentation tasks, and demonstrated it using works from the standard atonal repertory.
- In the early 1970s, Steven Smoliar developed a LISP-based system to study how Schenkerian theory could be implemented using a linguistically-motivated augmented transition network.
- In the late 70s, Aleck Brinkman developed a computer-based system for tracing motives and their transformations through the many preludes in Bach's *Orgelbüchlein*.
- In the mid-80s, John Maxwell developed a rule-based expert system to analyze keyboard dance movements by J. S. Bach. His analyses focused on three movements from the French Suites.
- In the late 80s, Robert Gjerdingen devised a neural network system that learned to recognize and anticipate various harmonic and contrapuntal features in several of Mozart's juvenilia for piano.
- Also in the late 80s, John Roeder created a knowledge-based system in Prolog to perform automated score segmentation tasks in atonal works.
- Around 1990, Brinkman and Martha Mesiti developed a system to graphically display computer-encoded pieces, including music of Bach, Bartok, Schoenberg, and Webern.

The studies I've cited, which are all listed in the reference list on your handout by the way, have in common the fact that they are examples, models even, of computer-mediated musical analysis that uncover interesting aspects of musical structure. They share another notable feature: The scores on which they are based were coded by their authors for the specific project at hand. And were never to be seen again.

Encoding scores for computer-based musical analysis is time-consuming—so much so that one rarely wants to encode more music than necessary, and is even tempted to pick shorter examples to simplify the process. Computer-based studies involving large numbers of works, or works that are long or have complex instrumentations, are very rarely done. If researchers are all encoding their own scores yet sharing them with no one, there is certainly the potential for duplicated effort, or what is perhaps just as bad, unduplicated research. When we do analysis by hand and are looking for examples to use, we often browse a large number of works, to find ones that illustrate a particular problem in certain ways. Those doing computer-based studies would also benefit tremendously if there were a large collection of works in the form they needed for their research. In fact, those doing "traditional" research (whatever that might be!) could benefit from being able to search a library of scores for specific features. In other words, we need a digital library of computer-encoded scores—a collective database of music in symbolic form, from which all interested researchers can draw.

This is not to say that there are no such collections. [SLIDE] In fact, many such collections are available. This slide, taken from an online list maintained by Don Byrd, shows a number of collections of computer-encoded scores in various formats. Among the largest of these, most are simply encodings of melodic incipits, such as the Themefinder database, which is searchable online, and the RISM collection (which has graphical images over 500,000 incipits from music since 1600—which are not, however, content searchable, but see your corner music library). Perhaps the two most interesting collections for our purposes are the MuseData database and Huron's Humdrum repository. [SLIDE] The MuseData database has been under development for a number of years at the Center for Computer-Assisted Research in the Humanities at Stanford, under the direction of Walter Hewlett and Eleanor Selfridge-Field. It contains more than 4000 movements or pieces from 939 different works by 11 composers, mainly from the late Baroque through Beethoven, and including a number of major works, such as all (or nearly all) of Beethoven's symphonies. It's a tremendous collection, but clearly has a long way to go before it can be considered complete. [SLIDE] Also, the MuseData format isn't well-suited to content-based search and retrieval. Here is an excerpt from the first violin part of a Correli trio sonata in the MuseData format. [EXLAIN]

[SLIDE] David Huron's Humdrum toolkit is a wonderful resource consisting of an extensive set of Unix-based programs to manipulate and search musical information. Part of a Humdrum file is shown here. It includes bibliographic data in the lines beginning with three exclamation points, and other types of comments with single or double exclamation points. The musical data itself is represented in a tabular form in which rows represents successive time points and each column contains some type of musical information. In this case, the columns represent the four instrumental parts of a trio sonata (the figures for the continuo part are not included in this encoding). The data here are in what Huron calls the *kern* format, which packs information on duration and pitch, as well as such things as beams, ties, slurs, articulation, and fermatas into a single data token. Although this is the most commonly used data type, Humdrum includes more than 60 pre-defined data types, covering a number of other pitch representations, lyrics and other types of textual information, dynamics, formal labels, MIDI, and on and on. Users are free to create their own types as well. The Humdrum tools, which are designed to work in conjunction with many of the powerful text-manipulation tools in Unix, perform many useful tasks, such as converting data from one form to another, extracting parts or sections of pieces, analyzing the data in various ways, looking for features or patterns in musical data, and so on.

In his lab at Ohio State, David Huron has accumulated over 30,000 scores in some format. Many of these are of western music, though there are also a number of folksong collections and music of other cultures. Many of the files were encoded to support a particular research question, so they may not include the same amount of bibliographic metadata, or have all of the musical features encoded, which may limit them to certain uses. Most of the MuseData collection is also available in Humdrum format as well. For copyright reasons, much of Huron's collection of encoded music is not available to other researchers.

The MuseData and Humdrum libraries, as well as others listed on the earlier screen, are valuable collections of music. Individually each is well-suited for a particular purpose. Taken as a meta-collection, however, there remain barriers to their general use. First, they represent different degrees of encoding. Some encode complete works, some encode only melodies, or melodic incipits, some include only MIDI information. Furthermore, they are all encoded in different, non-standard formats. Translation programs to convert one format to another, where they exist, almost invariably lose data. And, inasmuch as the creators have invested substantial resources into creating these collections, they tend to restrict or entirely prevent access to the collections. The labor involved with setting up and maintaining such a library is also prohibitively challenging for an individual researcher.

[SLIDE] All of us who are interested in applying information technologies to the study of music would benefit from a common repository of musical scores, whose content was in the public domain. Such a repository would need to include a set of tools to search the database and retrieve works that contained whatever information was of interest. And it would need to support the conversion of the data to whatever format was most useful for a given task.

Let me now describe a possible environment for such a collection. [SLIDE] Variations2 is a five-year research project, funded by the National Science Foundation, that is creating a second-generation digital library testbed system for music. The first generation is our current production digital library system, called VARIATIONS, which makes available over 9000 hours of CD-quality audio to our classrooms and library computer workstations. Variations2 is a multidisciplinary project, involving researchers from information science, law, computer science, university libraries, and information technology services. Obviously, much of what we will be showing here represents the work of many people, and not just Brent and I.

We will discuss the implications of this system for music theory shortly, but first we'll give a brief tour of some of the features that are currently implemented. We will have to skip over a number of details because of time limitations, but will be happy to demonstrate the system afterward. This is live demonstration—we are connected to the library directly (we hope!)

[SLIDE] As anyone who has searched a sizable online music library catalog knows, these catalogs were designed for finding books, not for looking for a score or recording of, say, a particular Schubert song. This is a screen from our campus library catalog. It shows results of an author search for "Schubert, Franz" which, not surprisingly, has resulted in an enormous number of records—over 2300—with helpful titles such as, "Piano Music. Selections." The problem is that librarians catalog *physical* objects like bound volumes and CDs, not the things contained *in* them, such as an individual *Lied.* Of course, we have all developed obscure strategies for getting around the system, but really, it's just a bad way to do things for music.

[SWITCH TO SEARCH WINDOW] The Variations2 catalog is built, not around the physical object, but around the notion of the *work* in its abstract form. This is the Variations2 search window. If we search here for Schubert, we get, not a list of *items* in the Variations2 library that *contain* works by Schubert, but a list of the *works themselves*. As we will see when the results appear, there are over 400 of them in the Variations2 library at this point and they are sorted by their uniform title. The search window allows the user to narrow the search by specifying a title and performer, the key of the work, and the media format (such as only recordings). The advanced search tab lets the user specify additional restrictions, such as publisher, or other contributors, such as an editor or librettist. Keyword searching is also supported. We can also browse the entire collection—a really cool feature that is so new I am seeing for the very first time myself right now!

Now, if we select one of those Schubert pieces, say, "Die Forelle," we get a list of all the instances of that work currently in the library. As you can see, there is one score and three recordings. Selecting the score opens the score viewer, which shows the contents of the physical volume containing the work along the left edge, with the music turned to the correct page. The scores are stored in the DjVu format, which supports zooming in and out without loss or distortion of things like staff lines, stems, barlines, and beams. We can navigate quickly a page at a time, or jump to an arbitrary page in the score. There is a bookmarking facility, which allows a user to mark a page in the score and then come back to it later. These locations are stored on the server and the users can access their bookmarks via the bookmarks menu during later sessions or from other computers.

If from the search window we select one of the recordings, a player application launches, with the contents of the CD or LP the recording came from, again cued up to the piece selected in the search window. (We plan to also support digital video in the Variations2 library, with similar functionality, in the future.) The player works as any media player, with the ability to change tracks, scan forward and backward, and so on. The audio is streamed, which provides for nearly real-time control over the location in the file. The bookmarking feature works for audio files as well. It is possible to edit the bookmarks, to change their text, and it is also possible to export a collection of bookmarks to an HTML file, which can be posted for students as a listening list, for instance.

During the cataloging process, we encode structural metadata representing key points in a work. This information is available on the Work tab in either the media player or score viewer. Although generally the work structure replicates the track structure of a recording, this mechanism will allow us to also store the locations of points *within* a movement, such as the exposition, development, and recap of a sonata form, or of individual variations in a variation set, and we will be able to jump to those points in a piece with just a mouse click.

[CLICK LINK IN POWERPOINT] Another feature we have implemented is the ability to synchronize the playback and display of audio and score files. In the search window, works that say "Listen and View" can be studied this way. To support this feature, the number of measures in a movement is stored as part of the metadata for the *work*. The time index for the start of each measure is stored as part of the metadata for *audio recording*, and the number of the first measure on each page is stored as part of the metadata for the *score*. These are then synchronized in what we call the Opus Window. As the audio file plays, the pages of the score will automatically turn (though this feature can be disabled). Changing pages in the score will cause the recording to jump to the correct place, while changing the location in the recording will cause the score to turn to the correct page.

We now return to the issue of encoded score files. Besides digitized audio and video and scanned score images, we plan to also include symbolically encoded score files in the Variations2 library. Having a library of encoded scores opens up a world of possibilities. For one, unlike scanned score images, they can be edited, so in a teaching situation for instance, one could quickly show the effects of changing a tonal answer to a real answer in a study of a fugue, or one could display a period from a Mozart sonata and ask the students to recompose the second phrase so that it modulates instead. Also, with digitally encoded scores and appropriate software tools along the lines of the Humdrum toolkit, for instance, one could do content-based retrieval—asking the library to search for 18th-century pieces that include the Neapolitan sixth chord. This is helpful for researchers. As David Huron and his collaborators and colleagues have demonstrated in impressive ways, when one has an extensive repertoire of works, a powerful set of tools, and the musical wisdom to use them well, one can ask sophisticated questions about those bodies of music, as well as about theories of music. But in the context of a digital library, content-based retrieval will also allow performers and teachers find pieces that have certain characteristics. For example, a tenor might want a piece with Italian lyrics, including the word "amore," in a major key, that has an heroic high C under a fermata. Not only will this hypothetical library provide that ability, but our library will *already* allow that singer to use the successful search results to call up sounds recordings of the piece in the library.

The decision to include encoded scores in a library raises a host of interesting practical issues, however. For example, what is the purpose of the encoded scores? Should they function as scholarly editions in which it is important to encode the actual location of each symbol on the page, to include editorial markings, together with an indication of the level of editorial interpretation? Or is it sufficient to encode the basic elements of the score, on the assumption that the edited version is already available in scanned form? Will the scores be played via MIDI and if so, should performance implications of musical symbols such as ornaments and grace notes be included as well? Who decides these things? How wide a range of music should be supported? Should we support neumes? Tablature? Pop chord symbols? Graphical notation? How are editorial decisions made in the encoding process where there is ambiguity in the source file (say in the placement of a dynamic)? What is the relative importance of the various features? What kinds of tools exist to work with the representation? Will they support the conversion to other representations? How do the needs of an academic library differ from the vast digital libraries maintained by companies in the recording industry, such as Sony? Or publishers, like Hal Leonard? One of the unique interests of the scholarly community, for example, is that we would like a musical score representation to support analytical information as well, such as the display and editing of Schenkerian analyses or roman numerals attached to chords.

Don Byrd, whose name you may know as creator of the Nightingale notation program for the Macintosh and who is a programmer on the Variations2 project, and I have considered many of these questions and have developed a requirements specification for digital music scores in an academic music library, which will be published in the December issue of *Computer Music Journal*. The specification classifies over 240 features, all representing some combination of graphical, performance, logical, and analytical information, into three categories: required, very desirable, and desirable. Our next step is to look at a number of existing representations to find the one that best matches these requirements. A discussion list has been set up with the OASIS standards organization to explore some of these questions in a serious way.

Our next question will be, how do we populate the library? Encoding scores is an immensely tedious process. Ideally, groups that are already working on this, together with new groups, would cooperate in developing a shared repository of scores. In the same way that OCLC uses a distributed model for library cataloging, the "many hands make light work" approach is simply critical if there is to be any hope of building such a library. (Substantial government help wouldn't hurt, either!) Technology is also coming to the rescue. Optical music recognition technology, perhaps most notably that developed by Ichiro Fujinaga and his former colleagues at Johns Hopkins University, continues to improve. This technology should speed up the encoding process considerably. MusicXML, a representation developed by a company called Recordare, provides interchange capabilities between different representations, including to and from Finale and Sibelius, Sibelius, and Humdrum. Other potentially helpful developments are emerging from ISMIR, the young music information retrieval discipline, which includes researchers from disciplines such as computer science, audio engineering, music, and information science. Members of the ISMIR community are working to develop a testbed of audio and score data, likely to be hosted at the University of Illinois, for researchers interested in this area.

Although the digital delivery of scanned scores and audio is already really pretty cool, as our user studies have shown, the real potential of the Digital Music Library comes when learners, teachers, and researchers can interact with those materials in new ways, can discover previously hidden things, and can communicate all of this innovatively. We will now describe some of the tools we are developing to enable this and what uses they might have for researchers.

One tool we have already implemented is a digital timeliner, which allows the user to create a formal diagram of a work or an excerpt. Brent will demonstrate this.

[BRENT: QUICK TIMELINER DEMO]

As Brent has shown, the timeline tool allows us to associate a temporal span with a formal description or a temporal location with a particular musical event. It provides a familiar, yet powerful tool for visualizing structural information about a musical work. The relative widths of the bubbles depict formal proportions. Color and labels can be used to relate or differentiate sections, while labels and annotations provide further information. Markers can be used to indicate the density and spacing of a particular event throughout a work.

Although the primary purpose of the timeliner is visual, the temporal information that is created in the process of making the diagram—information that can be extracted and analyzed—can also be used in a number of ways. For instance, it could be used to study a work's proportions or the timings of particular performances, or to compare the timing variations between several performances of a single work. Within a single work, we could use the data to look at greater and lesser golden sections, nested golden sections, or the use of Fibonacci sequences. More generally, one could examine whether comparable sections—such as the exposition and the recapitulation, or the variations on a theme—have comparable or disproportionate lengths. One could compare the durations of the lower-level sections with each of the other sections at different hierarchical levels, noting the ratios, perhaps highlighting any predetermined relations.

We might be interested in timing differences between several performances of a work. If studying tempo choices in performance, one could compare the overall durations of a work or of a work's major sections. If studying the timing or use of rubato in particular phrases, one could compare the relative durations of smaller segments. The same could be done with separate works that have similar forms. For example, Schubert songs or Scarlatti sonatas could have similar overall forms, but quite different internal proportions.

[SLIDE] This next slide is a purely visual prototype illustrating some of the functions still on the Variations2 drawing board. These are features that would relate directly to library content, in some cases audio, in some cases scanned score images, and in some cases encoded score files. [click] We plan to include a notation editor that supports both traditional and analytic notation. Here, a score and the Schenkerian analysis of it are displayed, synchronized horizontally. [click] We will provide tools to support the addition of roman numerals and other symbols, so that one can include analytic notation with a minimum of fuss. [click] A "scrubber" is used to track playback. [click] Playback is supported with an integrated media player. [click] Standard shape drawing tools to mark up the score: boxes, circles, lines and arrows of various sorts. The annotations are attached to the musical events on the score, so if the score is resized, stretched, or otherwise transformed, the annotations will go along. These tools could greatly simplify the tedious task of creating analytic diagrams. [click] It will have an optional multi-page design, which will allow multiple screens to be created within a single lesson file. This may be most useful for teaching situations. Other tools will include a Rich-text editor and a graphics viewer. [click] Finally, annotations can be saved locally, where they are linked to the media file on the server.

We think that it is important that a digital music library should include music in symbolic form. Where would Google be if it couldn't access the *content* of web pages, but only their appearance? And amazon.com has just recently added full-text search capabilities to its search engine for books. The vision of an exhaustive library of encoded music may never be fully realized. Nevertheless, working seriously towards realizing that vision has the potential to open new realms of research in areas such as musical style, compositional practice, the creation and testing of musical theories, and the visualization of music. Although we are developing a robust framework for cataloging, organizing, and distributing music in its many digital forms, the need for a library of encoded music clearly transcends our own needs. We invite others to share their needs and ideas with us and to work together toward building a resource that we will wonder how we one day wondered how we lived without.

The Implications of Digital Music Libraries for Music Theory
Society for Music Theory 2003, Madison
Eric Isaacson and Brent Yorgason, Indiana University

References

Brinkman, Alexander R. "Johann Sebastian Bach's *Orgelbüchlein*: A Computer-Assisted Study of the Melodic Influence of the Cantus Firmus on the Contrapuntal Voices. (Ph.D. dissertation: Eastman School of Music, 1978).

_____. "The Melodic Process in Johann Sebastian Bach's *Orgelbüchlein*. *Music Theory Spectrum* 2 (1980): 46-73.

Brinkman, Alexander R. and Mesiti, Martha R. "Graphic Modeling of Musical Structure *Computers in Music Research* 3 (1991): 1-42.

Byrd, Donald. Candidate Music IR Test Collections. http://mypage.iu.edu/~donbyrd/MusicTestCollections.HTML Rev. 28 March 2003.

Byrd, Donald and Isaacson, Eric. "A Music Representation Requirement Specification for Academia." *Computer Music Journal* 27:4 (Winter 2003): 43-57 (Forthcoming)

Forte, Allen. "A Program for the Analytic Reading of Scores." *Journal of Music Theory* 10 (1966): 330-64.

Gjerdingen, Robert O. "Using Connectionist Models to Explore Complex Musical Patterns." *Computer Music Journal* 13 (1989): 67-75. Reprinted in *Music and Connectionism*, ed. by Peter M. Todd and D. Gareth Loy, 138-149. Cambridge, Mass.: MIT Press, 1991.

_____. "Meter as a Mode of Attending: A Network Simulation of Attentional Rhythmicity in Music." *Intègral* 3 (1989): 67-91.

_____, Learning syntactically significant temporal patterns of chords: A masking field embedded in an ART 3 architecture Author: Gjerdingen, Robert O. Year: 1992 Source: Neural networks, vol. 5, no.4 July-Aug 1992. p. 551-564

Humdrum Toolkit. http://dactyl.som.ohio-state.edu/Humdrum [site down as of 11/1/2003] http://humdrum.net/mirror/Humdrum [accessed 11/3/2003]

ISMIR: International Conference on Music Information Retrieval. http://www.ismir.net/

Maxwell, H. John. "An Artificial Intelligence Approach to Computer-Implemented Analysis of Harmony in Tonal Music." Ph.D. diss., Indiana University, 1984.

_____. "An Expert System for Harmonic Analysis of Tonal Music. In: Understanding Music with AI: Perspectives on Music Cognition. Ed. by Mira Balaban, Kemal Ebcioglu,, and Otto Laske, 335-353. Cambridge, Mass: MIT Press, 1992.

Nord, Timothy A. "Toward Theoretical Verification: Developing a Computer Model of Lerdahl and Jackendoff's Generative Theory of Tonal Music." Ph.D. diss., University of Wisconsin, 1992.

Variations2. http://variations2.indiana.edu/