What Motivates a Musical Query?

© 2000 Eleanor Selfridge-Field
Consulting Professor (Music)
Center for Computer Assisted Research in the Humanities
Stanford University
Stanford, CA 94305-3076
esf@ccrma.stanford.edu
http://ccrma-www.stanford.edu/~esf

There are three places we might look to establish a range of questions that might profitably be asked of musical-data searches. First, we might look at **who** might want to search musical data and why. Next, we might look at **what** has already been attempted, with a view to establishing what has worked (or not) and why. Third, we might look at **how** searches have been/are being carried out, with a view toward establishing the constraints that man-machine reciprocity impose on musical queries.

Who and What: Some Early Query Projects

Efforts to search musical data and select portions of the material that correspond to a user-defined model have been discussed since the 1960s. At first the discussion was theoretical, for there was no musical data to search. Apart from composition, however, the production of notated scores was the primary emphasis of computer applications in music in the 1960s and 1970s. Since networking as we know it today was not developed, each of the large projects developed its own method for representing musical data. Some leading projects were based on DARMS (the Digital Alternate Representation of Music; Columbia University, from 1965); SCORE (Stanford and Colgate Universities, from 1972); MUSTRAN (Indiana University, from the early 70s), and CERL (Illinois University, from 1973). [For details of these systems and many others, see Selfridge-Field (1997).] None of these systems was primarily focused on the support of data query and analysis.

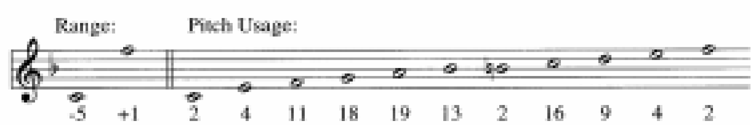
An early system that did have the primary goal of facilitating the query of musical databases originated at Princeton University in the mid-Sixties. Developed by Eric Regener, Michael Kassler, and others, it was based on something called IML/MIR [Intermediate Musical Language/Music Information Retrieval]. In contrast to the emphasis of early DARMS-based analysis on monophonic music and set-theoretic procedures, the Princeton work was exclusively concerned with polyphonic music (by Renaissance composers) and the minutiae of its notation, which are essential to assessing both proportional rhythmic values and contextual aspects of music theory.

The Princeton project initially encoded much of the music of Josquin des Prez, a Renaissance composer noted not only for the overall accomplishment and artistry of this music but also intricate allusions woven into the fabric of his works. Many works were encoded from multiple, divergent sources. What did people want to learn that they could not already find out? It was envisioned that the encoded sources (1) would support the reconstruction of the order in which manuscript sources were created; (2) would enable music theorists to establish the range and incidence of different kinds of harmonic procedures, such as the preparation of suspensions in cadences; (3) would improve methods for studying the complex hierarchies of rhythmic relationships which were enabled by mensural notation (a multi-level tree structure of sometimes twisted pairs of base-2 and base-3 groupings). Few of these objectives were met, but the work prompted the encoding of other repertories of Renaissance music and has undoubtedly contributed to some important research on modal and tonal theory (e.g., Powers 1981; Collins Judd 1992). [Ironically, these later studies have been conducted largely analogically, with computer involvement principally for the management of meta-data. Much of the encoded data was discarded after card-readers had become passè].

A project of greater achievement, in terms of preserved encoding and useful analysis, is one begun in 1982 at the Gesamthochschule fuer Musik at the University in Essen, Germany (see Dahlig, 1997, and its list of references). Building on a long-standing tradition in folk-music research, the project leader, Helmut Schaffrath, aimed to encode, store, display, and search monophonic repertories from diverse parts of Germany and German-speaking lands. The underlying materials encoded were variously recordings and scores. Because the music was monophonic, it was possible to encode a great deal of it fairly quickly, and the research quickly moved on to the development of a query apparatus. The group had encoded roughly 15,000 songs before Professor Schaffrath's untimely death in 1994.

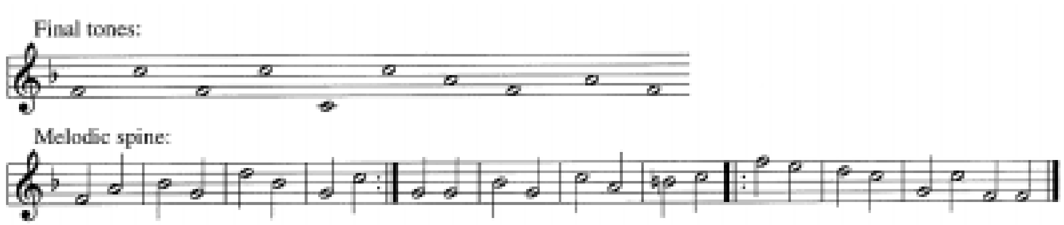
Schaffrath's data, which is distributed under license, has proved to be a legacy of considerable importance to those concerned with the development of query routines. In fact, contrasted with the Princeton work, use of his data and programming concepts have revealed an important fallacy in our thinking about musical-data-query. It seems reasonable to suppose that all queries pertinent to simple (i.e., monophonic) data can be derived *de facto* from software developed for the query of complex data. The case turns out to be otherwise because musical data can be simple or complex in an untold number of ways. One kind of complexity does not necessarily incorporate every conceivable kind of simplicity. The European cultivation of elaborate harmonic thinking has perhaps diminished respect for the modal and rhythmic complexities of other repertories.

In Western art music, for example, we like to categorize pieces according to their mode (major, minor) or key (A..G). One of Schaffath's queries assesses the range of the piece; an associated one lists the tones present, without bias towards categorization. In this instance (Figure 1) the incidence of each of the tones is also indicated.



In folk-song research, a particular pitch-list may be useful in identifying the region of the song's origin. Another departure from art-music "logic" is the query which retrieves only the final tones of each phrase.

In Figure 2a we see first the last tone of each phrase of the German anthem "Deutschland ueber Alles."



The value of such a schematic view to folk-song researchers is easily appreciated when this example is compared with the "melodic spine" shown in Figure 2b. This spine is a mechanical reduction at the half-note level of the entire song. Differences of rhythm (including presence or absence of repeated notes, rests, syncopations, and suspensions) can yield a description of the whole which is not entirely "true" to our cognitive impression. In folk-song research sorting by phrase-finals has been shown to cull more recognizable clusters of related songs than sorting by initial phrase content.

How? Some Query Considerations in Current Projects

The researcher intent on devising strategies for facilitating music query faces a formidable range of obstacles. A fundamental one is that music-query techniques are not simply straightforward extrapolations of text-query techniques. Another is that the term "musical data" is highly ambiguous: the information content for any given work can be selectively encoded in a host of ways. A third is that recent findings in human-factors research suggests that some of the features that matter most to listeners pertain to a particular performance of a musical work rather than to inherent features of the composition.

1. Music-Query vs. Text-Query
The diversity of goals which has always characterized musical query is compounded by an assumption that a musical query is essentially as **concise, specific, unidimensional, and unidirectional** as a text search. In addition, potential matches are tacitly assumed to have a length corresponding to that of the original query. These are potentially paralyzing expectations. The questions that can be answered by means as simple as those used in text searching are generally questions to which scholars and performers of music attach little value.

We have come to recognize these obstacles largely through our experience in two long-term research projects at the Center for Computer Assisted Research in the Humanities at Stanford University. (1) The *MuseData*® project is concerned with the development of a collection of fully encoded electronic scores. Like the Princeton work, it contains only polyphonic data and is designed for multiple kinds of applications (score editing and printing, query and analysis, etc.). (2) The *Themefinder*® project is concerned with data development and query software for a multifaceted melodic-search tool. Like the Essen work (and in fact incorporating data from the Essen project), it contains only monophonic music and is designed for a single application—thematic searching. *MuseData*® now contains more than 1,000 scores of classical music. These are preserved in a range of formats suited to different kinds of applications. *Themefinder*,® a Web-based tool, now contains more than 35,000 musical incipits, again in a range of formats. The URLs for both projects are given at the end of this article.

The quantity of data is much less interesting than the spread of repertories, the nature of the searches implemented, and the potential uses for retrieved data in new contexts. Briefly, *Themefinder*® incorporates incipits from several highly diverse repertories, including orchestral and chamber music (chiefly from the eighteenth and nineteenth centuries), polyphonic vocal music (chiefly from the sixteenth century), and folk music (chiefly from central Europe and from China). This breadth spans multiple notational, tonal, and mensural traditions. Therefore, although the brevity of incipits simplifies some needs, the range of repertories seems to complicate others.

Themefinder® currently searches at five levels of specificity/generality–by pitch string, intervallic string, pitch-class string, gross contour (relative direction only), and refined contour (relative direction, with step/skip discrimination). Items returned by the search can be pasted onto a clipboard and exported for further evaluation. They can also be heard as MIDI files. User feedback is welcomed, as are user submissions that meet the other criteria of inclusion.

Themefinder® has a growing supplement of textual data. The fields included identify information specific to individual pieces or repertories: e.g., composer, date of composition, and instrumentation and genre for classical instrumental music; social function, country, and region for folk-music. Links enable the extraction of all materials relevant to one long work (e.g., all the themes of a symphony) or all the works in a collection (e.g., the *Well-Tempered Clavier*). Text-fields enable preliminary results to be fed to other programs. In related work, for example, Aarden and Huron (2000) have fed "region" data to a geographical mapping program to show the distribution of musical features.

[The search algorithms for *Themefinder*® were originally devised by David Huron. The program implementation and Web interface, as well as data translation and the many tasks of verification, have been done largely by Craig Sapp, with contributions by Andreas Kornstaedt and a host of other Stanford students and visitors.]

We are led to believe that much of the most useful searching in musical repertories will be of the fuzzy variety. The degree of relationship between pieces already known to be similar is fundamental to many questions traditionally posed. It can aid the identification of tune-families in folk repertories; the degree of relationship between intended variations in classical instrumental music' or the covert agenda of parody and paraphrase in vocal music. Ambiguity and implication have played important roles in the detailed craft of musical composition over the past three hundred years. These subtleties beg for suitable ways of searching which are likely as heterogeneous as the repertories themselves.

2. Sound Data vs. notation Data

The results of fuzzy searching of musical material will be partially dependent on how the musical data have been defined. For one person "musical data" is necessarily derived from sound and is fundamentally temporally-ordered; to another it is necessarily dependent on notation, and is therefore spatially ordered. For search strategies, the difference is critical. Encoding schemes which privilege one domain tend to short-change the other, but multi-domain encoding within single files is enormously complex and difficult for programmers to decode accurately, efficiently, and completely. Music is full of contextual understandings and implicit meanings. The programming overhead on ferreting out either from a polyphonic score is considerable.

A simpler and yet very taxing problem is simply that the single symbol which in common Western notation which represents a "note" combines multiple parameters (pitch and duration; often additional features of articulation, dynamics, et al.), each of which may require contextual information to decode. The term "musical data" is almost unbearably vague.

Although it is generally true that meaning is determined by the order in which symbols are arranged, text is always searched and matched in one direction and on one dimension. Some appropriate matches in music may require selection from a large field of objects (with consequent weightings among them) or slanted views to locate composite patterns distributed across several parts instruments, or altered temporally or tonally from one movement to the next (or indeed from one work to the next).

For a textual explanation of such musical concepts, let us consider a cryptogram. In a cryptogram a cross-word puzzle is completed (1) by decoding verbal clues, (2) by transferring the letters associated with each number in each answer-line to the appropriate box in the puzzle, and (3) by confirming the result (a quotation extracted from a book) by decoding the title of the book. This is done by assembling the first letter of each answer word or phrase into a simple text-string. Many kinds of musical analysis in use today are more analogous, in their reductive tendencies to (3) than to (1) or (2). Whereas text analysis almost always depends on letters staying assembled into words and words into sentences, musical analysis allows for a lot of reductive selection. Let us construct a series of possible matches for the target *bag*. The results of a text search (ignoring case differences) would be these:

Principle: order is observed
Acceptable matches=bag Bag, BAG
Unacceptable: order is violated=gab, abg, bga

Principle: direction for comparison=L→R
Acceptable matches: processed L→R
Unacceptable: direction is violated =R→L; T→B; B→T; diagonals

Principle: literal content required for match=bag
Acceptable matches=*bag*gage, cab*bag*e
Unacceptable: literal content/semantic meaning is violated=*b*e*ag*le, *ba*n*g*, *b*ad *a*nd *g*ood

Briefly stated, certain conceptual principles used in understanding and describing music may be satisfied by the equivalent of some of the unacceptable matches shown here. The changed order may be acceptable in set-theoretic searches. The changed-direction matches may be acceptable in seeking retrograde statements in contrapuntally-conceived music. The dispersed matches (last example) mimic some important reductionist strategies used frequently in contemporary music theory. As we observed in Figure 2, reductive selection, an approach which is essential to many kinds of melodic searching (Selfridge-Field, 1998), can itself be practiced in many ways and is thus not an automatic solution to the need for data generalization.

3. Challenges from Human-Factors Research
Recent human-subject research on listening has revealed some surprising correlations that bear on search strategies. The following are only a few examples.

A. Among untrained popular-music enthusiasts, tempo may be a principal component of "style"-similarity judgments. Tempo is, of course, not an attribute of a musical work but rather of its performance and will be specific to each performance.

B. The level of vibrato affects judgments of "expression" in performance (Rapoport 1996). This is surprising because pedagogy preaches that dynamics and tempo change are among the principal means of expression. Vibrato is also a quality of performance, not composition.

C. Style and genre judgments of broadcast music can be made with 40% accuracy in .025 seconds by musically "untrained" teenage subjects, if the test repertories are familiar (Perrott and Gjerdingen 1999). This is surprising because samples are shorter than many single MIDI "events." Here human perceptual and cognitive abilities seem to defy the musical wisdom of ages. The general reliance on "events" to describe musical content in computer applications could fail if it can be shown that the threshold for "style" recognition is more granular, or more deviant, or more superficial than the essence of that range of musical "content" which events describe.

Conclusion
My purpose in contrasting the divergent goals of both earlier and current projects has been to demonstrate how broad the range of goals in music query may be. I have attempted to show how much query objectives may be colored and constrained by the repertories familiar to researchers.

The only common thread in music-query motivations, broadly defined, is--alas--human curiosity. Unless we lose that, designers of music-query software can expect to cater for an unending stream of "special" needs, as musical preferences continue to evolve and change.

References

Aarden, Bret, and David Huron (2000). "Geographical Localization of Musical Features: Mapping European Folksong," Computing in Musicology 12 (1999-2000), 115-128 [forthcoming].

Collins Judd, Cristle (1992). "Modal types and 'ut,' 're,' 'mi' tonalities: Tonal coherence in sacred vocal polyphony from about 1500," *Journal of the American Musicological Society* 45/3, 428-67.

Dahlig, Ewa (1997). "The Essen Associative Code: A Code for Folksong Analysis" in Selfridge-Field (1997), 343-361.

Perrott, David, and Robert O. Gjerdingen (1999). "Scanning the Dial: An exploration of factors in the identification of musical style," *Society for Music Perception and Cognition: Annual Meeting (1999)*, *Abstracts*, (Chicago: Northwestern University, typescript), p. 88.

Powers, Harry (1981). "Tonal Types and Modal Categories," *Journal of the American Musicological Society* 34/3, pp. 428-470.

Rapoport, Eliezer (1996). "Emotional Expression Code in Opera and Lied Singing," *Journal of New Music Research* 25, 109-49.

Selfridge-Field, Eleanor, ed. (1997). *Beyond MIDI: The Handbook of Musical Codes*. Cambridge: The MIT Press.

Selfridge-Field, Eleanor (1998). "Conceptual and Representational Issues in Melodic Comparison" in *Melodic Similarity: Concepts, Procedures, and Applications* (Computing in Musicology 11), ed. Walter B. Hewlett and Eleanor Selfridge-Field (Cambridge: MIT Press), 3-64.

URLs:
MuseData:® **www.ccarh.org/musedata/**
ThemeFinder:® **www.themefinder.org/**