After the search is over … the work begins

Tim Crawford (Goldsmiths College, University of London)

I am a musician who became interested in using computers in the mid-1980s in the hope that they might help me with the musicological work that was beginning to occupy more and more of my time. I am still waiting ...

Even though much research work has been done on the low-level feature-extraction and matching tasks of music information retrieval (MIR), relatively little effort goes into matching music at a higher 'semantic' level. Finding appropriate semantic levels for particular musical tasks is an open research field, one that will become increasingly important as low-level retrieval tools become more sophisticated and efficient, as they show every sign of doing, largely thanks to the ISMIR conference and its companion MIREX contest.

The semantics of music is a topic fraught with difficulty, at least, in comparison with text, which, after all, exists for the principal purpose of conveying meaning. The problem is not so much that musical objects cannot be grouped into hierarchical patterns that recur in possibly 'meaningful' ways – this seems to be a very strong feature of all musics – but rather that the hierarchical level at which such patterns are significant is not consistent. For example, we can usually agree that most (by no means all) music can be segmented into 'notes', which are defined by their 'pitch' (intimately related to the acoustic frequency of their sound). These might be regarded as 'atomic' units without individual meaning which combine to form meaningful groups (rather in the same way that characters in a language combine to form words, which are grouped into phrases and sentences), and indeed such groups certainly are perceptible. But things are very much complicated by the fact that music unfolds in time: the note-groupings can be found at a single point in time (a 'chord' can be thought of as a grouping of notes which all sound together at a point in time) or as a sequence of separate notes following one another in time, each with a certain duration (the 'tune' of a song is a large-scale grouping of this kind), or some intermediate situation (the 'harmony' of an certain passage of music can be thought of as a model of the process of selection of all the notes occurring within time occupied by the passage). Furthermore, pitch can be viewed at a variety of different hierarchical levels (octave, diatonic scale-degree, chromatic pitch) as indeed can musical time (time-offset, beat, measure). There is a complex interaction between these multiple hierarchies which makes the assignment of fixed semantic meaning to any such note-groupings generally unsound. It may be partly for this reason that it is usually only in very specific musical tasks (usually of a specialised technical nature) that music 'matching' can usefully be carried out on only a single one of these hierarchical levels.

Of the three note-grouping types that I outlined above the third, harmony, is of course essentially different from the others, as it comprises a *model* of the note-content that is 'possible' or 'acceptable' within a passage of music rather than the actual notes themselves. Such a model might be as simple as a histogram of notes collected periodically within a window moving through time, and the progression of such harmonic 'states' viewed as a Markov process. This is the basic idea of an MIR system developed by Jeremy Pickens and myself in the OMRAS project (www.omras.org), although we

Dagstuhl Seminar Proceedings 06171
Content-Based Retrieval
http://drops.dagstuhl.de/opus/volltexte/2006/747

used something more sophisticated than the simple note-histogram as the basis of our model. The point of this approach was to achieve approximate matching in a polyphonic context without incurring the kind of performance hit that generally results from the combinatorial explosion that note-for-note searching of polyphonic music can cause. Instead of matching the actual occurrences of notes present in musical database documents (or 'scores') against the actual occurrences of notes present in a query, we compare 'harmonic models' of scores and queries. Because the design of our model allowed for a perceptually-validated concept of 'harmonic distance' between states, it proved very robust in the task for which it had been designed – to allow matching between an encoding based on a written or printed score (actually in the form of a MIDI file) and an audio recording of a performance of the score; the present state of the art of audio transcription is still primitive, and prone to a good deal of error in terms of wrong and missing notes, although the wrong notes are usually perceptually fairly close to the correct ones. The models produced by the audio front end of the system, while imperfect, still generated enough notes to build a model which could be very successfully matched against the 'perfect' score model.

The robustness of our MIR system to a certain kind of 'noise' (wrong and missing notes) meant that it also had the property that it could recognise as similar two passages that may not contain the same sequences of notes although they share the same overall harmonic content. This means that, in general, it is rather good at recognising musical 'variations on a theme' of the kind that preserve the harmonic sequence of the theme; this was generally the case with sets of variations before the time of Beethoven.

Most of my presentation at this Dagstuhl seminar used a simple (and not very efficient) web-based version of the OMRAS system which is built on a database of around 5200 MIDI files, mostly of classical music, but also including a selection of pop, film and TV music, as well as Scott Joplin rags. It was intended to demonstrate how even a simple and imperfect MIR system can surprise and delight if operating at a suitable semantic level (in this case, that of harmonic similarity); in most cases I was especially interested in high-ranking matches that were strictly-speaking 'non-relevant' to the query (in the TREC sense). I then went on to show some of the motivating principles behind the 'Leitmotif' project that Michael Casey and I hope to begin (given suitable funding) in the near future (not discussed in this abstract).

After a brief introduction to the OMRAS web interface, the following searches were demonstrated:

1. **Scott Joplin Rags.** A search for the original MIDI file of an audio recording of a Yamaha MIDI-enabled piano performance using that file; the music was 'Gladiolus Rag' by Scott Joplin. In this case, the original score was ranked 6th; however, two other Joplin rags in the same key, 'Easy Winners' and 'Auto Mechanic' were ranked higher: 2nd and 4th, respectively! While this may seem unfortunate from a pure MIR-performance point of view, it does suggest that harmonic modelling may have something to tell us about compositional style (don't forget that the query was derived from an imperfectly-transcribed audio performance); three of the top six ranked matches are by the same composer.

2. **'Because' and the 'Moonlight' Sonata.** Jeremy Pickens had already mentioned in his own presentation the Beatles song 'Because' and its anecdotal connection with Beethoven's 'Moonlight' Sonata for piano. Using the MIDI version of 'Because' from the OMRAS database as the query returned itself ranked first (a useful sanity check!), and the Beethoven sonata at number 2; however, this was not the first and most famous movement ('quasi una fantasia'), but the rather more vigorous final movement, suggesting that John Lennon may have based the chord sequence of 'Because' on Yoko Ono's private performance not of that well-known first movement, but on the last movement of the same work. Adjacent 'non-relevant' matches were 'Le Cygne' from Saint-Saens 'Carnival of the Animals' ranked at number 3 and Chopin's Étude, Op. 25, no. 12 (which sounded in this MIDI performance rather closer to 'Because' than the others), ranked 4th.

3. **Gounod's 'Ave Maria'.** This is well-known to have derived from Gounod's instrumental 'Meditation' on the first prelude from J.S. Bach's 48 Preludes and Fugues. This added a melody (for violin or cello) over the notes of Bach's prelude, which is thereby relegated to the function of an accompaniment. There exist countless arrangements of this derivative work for all kinds of vocal and instrumental combinations, some from Gounod himself, some by his near-contemporaries, and some from nearer our own time, including a few electronic versions. Over time, Bach's prelude features less and less in terms of its note-detail, while Gounod's melody is usually fairly well preserved; what remains more or less constant, however, is the harmonic outline of the piece (except in some of the more outrageous renditions, such as two I have in MP3 format which claim to be 'hip-hop' versions). The ranking of different versions, and of Bach's prelude (or the prelude and fugue combined in a single file) in a search using a MIDI performance of Gounod's 'Ave Maria' showed a similar pattern; Chopin's well-known C major Étude, Op. 10, no. 1, which is known to be indebted to Bach's prelude, tends to be ranked highly as well.

4. **A Bach chorale not by Bach.** There is a well-known collection of 371 Lutheran chorale melodies or hymn-tunes set in four parts by J.S. Bach; this has been used as teaching material by music teachers since the early 19th century as a model of excellence in the art of harmonisation of simple tunes, until recently part of the formal training of any classical musician. Hidden amongst the 371 are a number which are actually not by Bach, and were included in error when the collection was first published in the 18th century. A well-known organist of the generation before Bach, Johann Pachelbel, exerted a strong influence on him; a few MIDI files of his works (apart from the omnipresent 'Canon') can be found on the web, including a set of variations for organ on the chorale, 'Alle menschen müssen sterben'. Using this work as a query, I found that one of Bach's chorales, with the same title, was consistently ranked fairly high, typically at no. 18 (out of 5200); on listening to the chorale it is immediately obvious that the 'Bach' setting is in fact the first movement – the 'theme' – of Pachelbel's set of variations, from which it differs only in minor details. (If the search is inverted, with the 'Bach' as query, the Pachelbel set is ranked still higher, at no. 7.) This is despite the fact that most of Pachelbel's variations depart considerably from the harmonic setting of the theme, all except the last, which is an elaborated version of the theme in a

similar setting. So the system seems to be modelling what might be thought of as the *predominant* harmonic content of the work; the more extreme aspects of Pachelbel's variations being effectively treated as 'noise' to the model.

5. **A Bach fantasia not by Bach.** In 1991, a rather mysterious work in J.S. Bach's list of compositions (included because he copied it out in his own handwriting around 1740), the Suite in A for violin and harpsichord, BWV 1025, was shown to derive from a suite for solo lute by the contemporary lutenist, Silvius Leopold Weiss (1687-1750); Bach had apparently arranged the lute work for harpsichord and added a new obligato part for violin. One movement from BWV 1025, the opening Fantasia, does not, however appear in Weiss's suite; but two slightly different versions of an anonymous prelude in the same key found in contemporary manuscripts loosely associated with Weiss share much of the same musical material. Although this concordance was not originally discovered by an MIR search, the process of discovery can be duplicated by using the prelude and the Fantasia in turn as search-queries to the OMRAS system; in each case, the two pieces are ranked 1st and 2nd even though they are not identical in terms of note-content and quite different in length.