“What was the question?”: Music Analysis and the Computer

Alan Marsden
Lancaster Institute for the Contemporary Arts, Lancaster University

This is a post-print (author’s final draft) of a chapter published in Tim Crawford & Lorna Gibson (eds) *Modern Methods for Musicology*, (Farnham, England: Ashgate, 2009), 137-147. Details of the definitive version are available at www.ashgate.com.

Abstract

Four decades ago, a gulf was identified between traditional and computer-based musical research. It seems to exist just as much today in the field of analysis, judging by the small number of studies using computers published in *Music Analysis*. On the other hand, there are now few impediments to analytical work with computers. Systems for encoding and processing musical information and substantial databases of encoded music are now readily available. However, most examples of research with computers come not from the field of music analysis, but from other areas of systematic musicology, and particularly from Music Information Retrieval. Not only is the methodology and terminology commonly different from those of traditional music analysis, but also the objectives. Using a computer imposes demands of systematic definition, but also opens up new opportunities, and it is inevitable that these should lead musical research with computers in new directions. Though not following the paradigms nor expressed in the language of music analysis, some of this work is nevertheless analytical in nature, and measures to bridge the gulf and bring analytical dividends from current computer-based musical research are proposed.

1. A ‘Gulf’?

‘In spite of the ever-increasing use and acceptance of modern data processing equipment for humanistic research, there would appear to be … a widening gulf between scholars who pursue the traditional methods of historical musicology and those who have adopted the computer as their chief research tool. This unfortunate and unnecessary division stems largely from a misunderstanding of the nature and limits of computer-aided scholarship by the former group; however, the situation is hardly assuaged by members of the other camp who are often unwilling to discuss their work in terms intelligible to the uninitiated.’ These are the opening sentences of an article on ‘Music Analysis and the Computer’ written nearly forty years ago (Erickson, 1968, pp. 240-241). I was surprised to find how apt they are to open my own contribution on the topic: little has changed, it would appear. We can now strike out the word ‘widening’, but there is little evidence that the gulf between ‘traditional’ analysts and those who use the computer ‘as their chief research tool’ is narrowing. The publication of *Empirical Musicology* (Clarke & Cook, 2004) and the establishment of the on-line journal *Empirical Musicology Review* in 2006 provide some evidence, but books and journals come and go. The journal *Computers in Music Research*, for example, seems to have ceased publication in 1999.

Certainly the pages of the journal *Music Analysis* do not provide evidence of the widespread use of computers in analysis. Out of 221 articles, I have found only six which report or refer to explicit use of computers, and they are spread evenly
throughout the journal’s issues. Two are concerned mostly with matters of theory (Baroni’s article on musical grammar (1983), and Temperley’s on the ‘line of fifths’ (2000)) but do give examples or refer to applications of software to particular pieces or a particular repertoire. Another two linked articles also principally concern theory, but in this case theory which is of a degree of complexity that a computer is required for its application: Pople’s exposition of his Tonalities project (2004), and Russ’s account of analytical examples applying the theory and software (2004). (In truth, traditional tonal-harmonic theory is similarly complex—Pople intended his theory very much as an extension of existing theory—but we learn to apply it without a computer only through years of practice. The essential role of the computer in Pople’s theory is to enable the assumptions such as remain tacit in our application ‘by hand’ of traditional tonal theory to be made explicit in the application of the software.) The author of the fifth article (Nattiez, 1982) does not use a computer himself but makes a comparison with another study of the same piece in which the use of a computer was essential (Tenney, 1980.) Tenney’s use of the computer is similar to Pople’s: as a means both for making discoveries and for testing and making assumptions explicit, though about segmentation rather than about harmony and tonality. It is only in Cook’s analysis of performances of a Bach prelude (1987) that the principal focus is on what can be discovered about the music rather than on the development of new theory, and the computer is an incidental tool, in this case for discovering the timings of notes in audio recordings.

Probably the computer has also been an incidental but unreported tool in other studies published in *Music Analysis*. The journal contains two brief comments by Forte on the use of computers in pitch-class-set analysis. In one he seems to look forward to an all-encompassing analytical tool: ‘one can envision … a powerful set-complex analyser with artificial intelligence aspects’ (1985, pp.54–56). Thirteen years later, the computer is mentioned not as an intelligent analysis tool, but as a kind of calculator: ‘you really need a computer program to generate matrices’ (Ayrey, 1998, p. 231). Certainly, computers must have been used in the preparation of a number of the contributions concerning pitch-class-set genera in that issue of *Music Analysis*, some quoting figures to four decimal places! Probably it is as a calculator and collator of pitch-class sets that computers are most widely used by music analysts (if we discount the widespread and equally not-worth-mentioning use of computers in preparing music examples and writing papers). Among software specifically intended for music analysis, programs for pitch-class-set analysis are the most common: a number are readily available, and there are even on-line tools to work out pitch-class set membership. Though unreported, the impact of this use of computer as quasi-calculator is not to be underestimated. It leads to greater productivity, and perhaps we should take this silent use of computers as evidence against the gulf Erickson saw forty years ago. On the other hand, this kind off use of computers is not worth mentioning in articles, and not worth discussing further here, because the actual *analytical* step in making pitch-class-set analyses is the segmentation—everything else is trivial calculation and collation—and it is precisely *not* in this step that computers are used.

### 2. Examples of Music Analysis by Computer

The latest volume of the series *Computing in Musicology* is dedicated to music analysis (Hewlett & Selfridge-Field, 2006), but even here actual analysis of pieces of
music is found only in a minority of the contributions: four out of thirteen. The remainder are principally about issues of representation and software design. (One of the four does not actually mention computers, but we must presume, since it is in this volume, that computers have been used.) What is interesting about these four, though, is that they provide examples of the three kinds of non-trivial contribution computers can make in analysis. One, like the uses found in Pople, Temperley and Tenney above, is to use the computer as a means for developing and testing a theory. The task the computer performs is one a human could do, but the computer can be relied upon to be accurate and, above all, impartial. An analyst approaching a piece of music brings with her or him a richness of experience of this and other pieces, which inevitably influences judgments. One way to make explicit the reasons for one analysis of a piece rather than another is to do so through a theory which does not rely on personal experience, and one way to lay aside personal experience in the application of a theory is to delegate it to a computer which has no personal experience. Mavromatis (2006) uses a Hidden Markov Model to model the structure of Greek church chants. (This is a formalism well established in computer science and so one about which there is a wealth of knowledge on which to draw.) An added bonus is that the computer can do some of the work in theory formation, in that a Hidden Markov Model can ‘learn’ from a set of examples. The resulting model is able to generate new chants in a similar style to the given examples, and from the internal structure of the model Mavromatis is able to draw conclusions about the stability of cadential formulas in this repertoire, in comparison to the variability of opening gestures. (Machine-learning research has become quite common in the field of Music Information Retrieval (see below), but often there is no focus on what has been learned, and so no direct music-theoretic or analytical benefit.)

A second kind of computer application also uses the computer to do something humans can already do, but not because the computer is more accurate and neutral, but because it is very much quicker. This kind of research typically analyses a body of data larger than a researcher could hope to deal with, and two examples are found in Music Analysis East and West. Veltman (2006) demonstrates that hierarchical metre can be found in sixteenth-century vocal polyphony by counting the occurrences of stressed syllables in notionally stronger and weaker positions in specific rhythmic patterns. Deguchi & Shirai (2006) count the occurrence of three- and four-note pitch sequences in the melismatic and non-melismatic sections of Japanese koto songs, concluding that the melismas use a more restricted range of patterns but are more variable. Neither of these studies uses particularly large datasets (seven motets in the first case and six songs in two versions each in the second case). It would have been possible to complete both pieces of research by hand, though extremely time-consuming. Computer-based studies do exist, however, which have used larger datasets than a single researcher could study in an entire lifetime. Meredith (2007) tested five classes of pitch-spelling algorithm on a corpus of 216 movements, testing several, and sometimes many, different versions of each algorithm. (A pitch-spelling algorithm converts pitches designated in a twelve-note chromatic scale to pitches with a specific letter-name plus accidental.) Pople’s software also falls into this category of enabling research which would otherwise be rendered impossible by the sheer quantity of data. In his case it is not the number of pieces which is at issue (Pople’s software could be applied to a corpus of many pieces, but the design of adjustable parameters is intended more for studies of single pieces); it is the number and complexity of interpretations within a single piece which would swamp the analyst
trying to follow the same analytical procedures by hand. To return to pitch-class-set analysis, if one wanted to list all the Fortean set classes found in all possible segmentations of a melody of a hundred notes, one would have to determine the classes for approximately seven hundred sets. In a polyphonic piece where sets can be found also by combining notes from more than one voice, the number increases rapidly. Thorough and systematic analysis of even single pieces thus often involves larger quantities of data than can reasonably be dealt with by hand. If one needs to consider not just all the possibilities at each point in a piece but also the different combinations of possibilities, the quantities can become larger even than a computer can handle. Imagine, for example, that each bar in a piece of music can be interpreted in one of two ways. The number of different combinations of interpretations for \( n \) bars is then \( 2^n \), and a piece of just 64 bars would have more than 18 million million million different interpretations. The most powerful computers in the world might just about deal with each of these within a century! Thus while a computer can make possible kinds of analysis which would otherwise be impossible, one must not think of it as a machine which makes possible any kind of analysis we can think of.

The third class of computer-based analyses also uses the computer to achieve something otherwise impossible: to extract data from sound at a level of detail too small or too precise to achieve by ear. Krishnaswamy (2006) examines pitch-time traces of Indian classical singing to argue that pitch in this music is based on twelve tones, and notes claimed by others (on the basis of assessments by ear) to be microtonal are actually distinguished not by their tuning but by their articulation or ornamentation. Cook’s measurements of timing referred to above, more precise than could be easily achieved by ear, and similar studies based on timing extracted from MIDI or other performance data also fall into this category. Computer tools for the analysis of audio are now extremely sophisticated, and increasingly easy to use with the development of software such as Sonic Visualiser (Queen Mary University, n.d.). This opens a large and very significant potential field of research in analysis of music-from-sound rather than music-from-score.

3. Tools for Analysis by Computer

In considering the lack of use of computers in music analysis, Erickson commented that ‘there are as yet no standard for the encoding of music, no generally available musical analysis programs … and no comprehensive theoretical system for computer-aided analysis’ (1968, p. 242). In our post-modern world, it is no surprise to find that we have many encodings for music, but no single standard, many music analysis programs, but few that are generally available, and—of course—many theoretical systems but none that are comprehensive.

Early efforts in music research with computers spent considerable amounts of time on designing a representation, and then on encoding music in that representation. A number of representations designed since then are described in Beyond MIDI (Selfridge-Field, 1997), but new ones are still arising, among them the increasingly significant MusicXML (Recordare, 2007). Representation is only an issue for those who either need specific information not already contained in existing schemes. Furthermore, large quantities of music are now freely available encoded in MIDI, MuseData or kern formats in the MuseData and KernScores databases (http://www.musedata.org and http://kern.humdrum.net, respectively), so there is a
reasonable chance that a researcher will be spared the work of encoding also. This, at least, has changed in the past forty years. Unfortunately, the promise of software able to derive encodings by scanning printed scores does not seem yet to have been realised. Researchers do need to be aware of what an encoding does and does not represent, and they should be concerned about accuracy also: the issues of interpretation of sources which apply to score-based studies do not disappear in the digital domain.

Some music analysis programs are now generally available. Pople’s Tonalities software has already been mentioned, and is available for download (Adlington, 2007). To describe what it achieves in a single sentence is not easy, but broadly speaking it allows one to see the consequences in analysis of considering a piece of music to use a particular vocabulary of chords and scales. Apart from Russ’s article, I am not aware of published cases of use of the software, however. Software by Sleator and Temperley (n.d.), which produces analyses of metre, grouping, key and harmony (though not always with a high degree of accuracy), may be downloaded as source code. Some published studies have used this software, usually as a quick means of determining harmony (e.g., Aarden & von Hippel, 2004). By far the most widely used analysis software, however, is Huron’s Humdrum (n.d.). Its beauty is its flexibility. It consists of many components each of which performs a small task, but which may be strung together to perform complex analytical tasks. The framework of representation is specifically designed to allow new representations to be incorporated, both to represent novel kinds of music and to add previously unrepresented detail to existing encodings. New tools can be written by those with the appropriate programming expertise and used along with existing components. As might be guessed from this description, however, it is not the sort of software that one can pick up and start using after just a few minutes’ perusal. Some indication of the flexibility is given by the fact that Veltman (2006) used Humdrum for his study of stress and metre, while Aarden & von Hippel (2004) used Humdrum to study doubling in chords.

4. New Analytical Directions

If we now have the benefit of examples such as those listed in section 2, and the tools listed in section 3, why does Erickson’s forty-year-old gulf still exist? Certainly humans are capable of extreme ignorance, arrogance and stupidity, but surely those cannot be the only causes of the lack of interest on one side and failure to communicate on the other, which Erickson lamented. One answer has already been adumbrated in the discussion of the capabilities of the computer: its ability to be neutral and impersonal, to deal with large quantities of music, and to reveal what cannot be heard by the ear. Yet the most common preoccupations of music analysis are to deal with what the ear can hear, to examine single pieces, and to expound and influence the personal listening experience. This does not mean that the computer is useless or irrelevant for analysis (the examples above demonstrate otherwise) but it is important to realise that an extra step is needed after the computer has done its work to make the connections back to the world of personal listening experience, to illustrate how conclusions drawn from a study of a many pieces influence our understanding of individual pieces, and to explain how the imperceptible details of tuning or timing do nevertheless have an impact on what we hear.
Often to take this extra step would be premature, though, because even after forty years computer-based analysis is still largely experimental and conclusions of a traditional analytical kind can only be tentative. Erickson warned in 1968 ‘there is a danger … that the System may become the end in itself’ (p. 260), and this remains the case today. (Indeed, I have fallen victim to it myself, when what was originally a simple consideration of how to represent information about temporal relations in music became an entire book (Marsden, 2000).) But this is only a danger from the perspective of traditional analysis. Scholars frequently pursue ends in themselves, and the application of computers in music has opened paths to many such ends. I commented above that instead of the comprehensive theoretical system for computer-aided analysis which Erickson hoped for, we now have a plethora of theories (see, for example, Marsden, 1993). The concepts of traditional music analysis are not entirely systematic and precise (their power is precisely in their allowance for expert knowledge and experience), so those who apply computers to analytical problems must redefine concepts for the impersonal, digital domain. What does it mean, for example, for a passage of music to be in a particular key (Marsden, 1999)?

Perhaps Music Analysis was the wrong journal in which to look for examples of analysis by computer. A similar survey of a journal where use of computers is the norm, Journal of New Music Research (formerly Interface), turns up, among 365 articles since volume 18 (1989) (a slightly shorter period than the 25 years of Music Analysis), eighteen articles which use a computer in the explicit analysis of a single piece of music or collection of pieces, and furthermore these occur with increasing frequency in more recent years. It must be said though, that five of these are by a single author (Juhász, 2000a, 2000b, 2002, 2004, 2006) who makes comparative analyses of collections of folksongs. Four others (De Matteis & Haus, 1996; Agon et al., 2004; Hoffmann, 2004; Keller & Ferneyhough, 2004) could be described as analyses by synthesis, where a piece is ‘reconstructed’ in a systematic manner using a computer tool. Two (Cambouropoulos & Widmer, 2000; Lartillot, 2005) apply systems to extract motives from scores. Issues of rhythm and metre connected with performance are examined in two others (Moelants, 2000; Fleischer & Noll, 2002). Two analyse performance, in one case from recordings (Rapoport, 2004) and in the other in a live set-up (Koppiez et al., 2003). Nettheim (1992) tests a theory of 1/f distribution of pitches by application to pieces by a number of classical composers. Guigue (1997) demonstrates software which extracts ‘sonic’ properties from configurations of notes and allows those configurations to be compared. Chew (2005) examines automatic segmentations in two pieces by Messiaen. As the occurrence in the titles of terms such as ‘adaptive redundancy filtering’ and ‘entropy-based learning system’ implies, most of these are not expressed in language ‘intelligible to the uninitiated’ (to quote Erickson); the readership of Journal of New Music Research is presumably not considered uninitiated. Furthermore, all of these to some degree concern exposition and development of method rather than focussing on analytical conclusions (not surprising in what is still a largely experimental field). Indeed, many other articles contain data drawn from individual pieces or segments of pieces, but which do not focus sufficiently on the analysis to be counted in this survey.

Besides its different methodology, all of this research has rather different objectives from traditional music analysis. While some familiar themes can be found in the articles referred to above—segmentation and motives, for example—even here the focus is different. The conclusion is not a single segmentation, but that if one
measures in this way this segmentation follows, whereas if one measures in that way that segmentation follows. Similarly it is not a single motivic analysis which is presented, but an array of possible motives of different strengths. Thus the analytical result of such studies is often not an interpretation of the piece in question but a mapping of the terrain of possible interpretations. While the analysis is undoubtedly therefore shallower in one sense, it is richer in another.

Different objectives even more markedly underlie the field of Music Information Retrieval (MIR), which has increased markedly since 2000, and is now probably the most productive field of musical research with computers. The driving forces here are commercial: companies would like to be able to process musical content in a manner to produce the most desirable services in a digital music marketplace. Nevertheless, interesting and imaginative research is being conducted with relevance wider than these narrow commercial interests. The MIREX competition associated with the ISMIR 2005 conference (Downie et al., 2005) included an evaluation of thirteen different systems for automatic genre classification. A database of audio recordings, labelled ‘ambient’, ‘blues’, ‘classical’, ‘electronic’, ‘ethnic’, ‘folk’, ‘jazz’, ‘new age’, ‘punk’ or ‘rock’ was used, and each system had to determine, from the sound alone, into which of these ten genres to classify each item. The best-performing system classified with an accuracy of 69–78% (depending on how it was measured). (Full results can be accessed from http://www.music-ir.org/evaluation/mirex-results/audio-genre/index.html.) Looking more closely at the detailed results for all tested systems, we find that ‘punk’ and ‘classical’ were most often correctly classified (90% of cases), but ‘punk’ more consistently so (i.e., some systems performed badly at identifying ‘classical’). At the other extreme, ‘new age’ music was correctly classified only in 25% of cases, and it was just as frequently mis-classified as ‘ethnic’. However, music which was labelled as ‘ethnic’ was only infrequently mis-classified as ‘new age’. Furthermore, there was considerable variation between the systems in how often ‘new age’ music was mis-classified as ‘ethnic’. The objectives in this research were not explicitly analytical, but analytical conclusions can be drawn. Specifically, ‘punk’ music appears to have the most distinctive sonic characteristics; ‘new age’ music has some characteristics which are similar to ‘ethnic’ music, causing some systems to mis-classify, but other characteristics which are distinctive.

5. A Gulf

The gulf seems wider than Erickson thought: it is not one of understanding only but one of objectives also. This is probably inevitable. Artificial intelligence has not enabled computers to behave just like humans, and I suspect that if it did we would lose the benefits which computers bring to research. If we want to use computers in music research, we cannot escape translating fuzzy musical concepts into precise computational terms: the language is bound to be different and so incomprehensible to some. Furthermore, we should not be surprised that the different capabilities of computers have influenced researchers to pursue different objectives. Like explorers who set off to find one thing but discover another, those using computers for research in music have inevitably forsaken the original questions of music analysis and pursued novel lines of enquiry.

Though the gulf is perhaps inevitable, we should not allow it to impede research. Three particular kinds of bridge are possible. Firstly, training programmes, especially
at postgraduate level, could increase the number of individuals with expertise in both
music analysis and computing. I suspect it is easier for a musician to learn the
necessary computing skills and concepts, because these are more circumscribed, than
it is for a computer scientist to learn the subtleties of music analysis. Secondly, interdisciplinarily research teams with both musicians and computer scientists could
achieve the same degree of productivity in music analysis as seen in the field of
computer music at centres such as IRCAM in Paris and CCRMA in California.
Finally, more could be done to effect the step I outlined above of drawing out music-
analytic conclusions from computer-based work. This requires both greater effort from
those who work with computers to relate their work to the preoccupations of
musicology and music analysis and greater vision on the part of music analysts to see
the new horizons being exposed by computer-based research.

References


