09051 Abstracts Collection
Knowledge representation for intelligent music processing
— Dagstuhl Seminar —

Eleanor Selfridge-Field[1], Frans Wiering[2] and Geraint A. Wiggins[3]

[1] Stanford University, USA
[2] Utrecht University, NL
frans.wiering@cs.uu.nl
[3] Goldsmiths College—London, GB
g.wiggins@gold.ac.uk

Abstract. From the twenty-fifth to the thirtieth of January, 2009, the Dagstuhl Seminar 09051 on "Knowledge representation for intelligent music processing" was held in Schloss Dagstuhl – Leibniz Centre for Informatics. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations and demos given during the seminar as well as plenary presentations, reports of workshop discussions, results and ideas are put together in this paper. The first section describes the seminar topics and goals in general, followed by plenary 'stimulus' papers, followed by reports and abstracts arranged by workshop followed finally by some concluding materials providing views of both the seminar itself and also forward to the longer-term goals of the discipline. Links to extended abstracts, full papers and supporting materials are provided, if available.

The organisers thank David Lewis for editing these proceedings.

Keywords. Music representation, music encoding, digital music edition, Music Information Retrieval, intelligent music processing, music informatics, data formats, data interchange, music collections, audio, MIDI, MEI, TEI, humdrum.

Cross-cultural study, folk song research, contour analysis, performance analysis, cognitive modelling, music analysis, diatonic set, pitch class set theory, musical scales, expressive strategies, artistic individuality, early music, partbooks, mensural notation rules, score assembling.

Synchronization, alignment, multimodality, segmentation, annotation, modelling, model evaluation, pattern discovery, partial similarity, harmonic similarity, melodic search, implicit learning, unsupervised learning, multiple viewpoints, machine learning, self-organising maps, multidimensional mapping.

Table of Contents

Part I

Introductory and Plenary Papers

Chapter 2

Introduction

1 Knowledge representation for intelligent music processing

The ubiquity and importance of music have made it an obvious candidate for applications of new technology throughout history, but most notably since the late 19th Century, when analogue electronics and then digital computers were brought to bear. There was initially an emphasis on the production of audible sound, but as computers became powerful, they were used in the generation of scores, and in recent years digital technology has approached the difficult problem of the understanding of music, both as what is heard and what is imagined.

This seminar aims to promote the computational study of music at levels of abstraction higher than the audio waveform. Doing so will enable automation of the kind of reasoning applied explicitly by music composers, analysts, researchers and performers as consciously-developed skills, and implicitly by informed listeners as high-level cognitive processes.

Many music encoding systems have been created since the 1960s, and large quantities symbolic musical data have been produced across the world, as the output of disparate projects, and represented for storage in ways which are not interoperable. Music knowledge representation research, as opposed to musical data encoding, emerged in the 1970s. Only after several decades of research, consensus on generally appropriate features for music representation was reached, and approaches—for example MEI, MusicXML, and MPEG7 Notation—have been developed which do model music more fully. Only recently, attempts have been made to represent music in ways which conform to the principles of Knowledge Representation, in that their specifications explicitly include inference systems. The inference aspect is fundamentally important: a computer encoding of data is meaningless without a method for interpreting it.

An important area of application is in digital critical editions of music. Whereas paper editions have the drawback of presenting a selective and static image of a composition, digital editions potentially provide a more complete representation of the source materials and allow different 'views' of these to be generated automatically. Suitable knowledge representations for these sources would allow inference of missing information that is considered essential for modern study and performance, such as accidental pitch changes in Renaissance music, voice leading in lute tablatures, realisation of implied chords in basso continuo accompaniment, and also suggest solutions for unclear, illegible, corrupted and lost passages. Finally they would allow the compositions to be processed by means of a wide range of music-analytical or music retrieval methods.

Keywords: music representation, intelligent music processing, music encoding, music informatics, digital music edition

Joint work of: Selfridge-Field, Eleanor; Wiggins, Geraint; Wiering, Frans

Chapter 3

Plenary 'stimulus' talks

2 Studying Music is Difficult and Important: Challenges of Music Knowledge Representation—or: Writing Poetry About Music is Like Dancing About Architecture

Donald Byrd (Indiana University—Bloomington, US)

In my view, music is one of the most complex, difficult to study, and important to study of all cultural phenomena—in fact, among the most complex, difficult, and important of any phenomena whatsoever. The following features of music in general make it so. Some of this applies particularly to computer processing of music, but most of it applies in any situation.

1. **Music is an art.** Therefore the composer/artist can use its elements any way they like—for example, to confound music-IR systems, as the amusing essay 'Composing to Subvert Content Retrieval Engines' points out (Collins 2006). Obviously, few if any composers/artists have that goal in mind, but a great many—at least in some cultures, including ours—try to use its elements in original and interesting ways, not straightforward and conventional ways. This (among other things) makes content-based retrieval a great deal harder with music than with expository prose, the type of text that text-retrieval systems usually deal with and that music retrieval is usually compared to. Similarly, there is a story that Marc Chagall said, in response to criticism of his drawing by an art critic, 'Of course I draw poorly. I *like* to draw poorly.' That is, in his art, Chagall had no intention of using the element of drawing the way it was ordinarily used. The story may be apocryphal, but the point is that a creative person can always find original ways to do whatever they do; that's virtually the definition of creativity.
   One of the implications of this observation is that we should not expect information retrieval of music to be much like retrieval, in the text domain, of prose; it's more like retrieval of poetry, where the denotations of words as given in dictionaries may be less important than their connotations or even their sounds. But this phenomenon applies to many problems of music informatics, and, indeed, to many problems of doing anything at all with music. For example, in Byrd (2009), I show and briefly discuss some surprising examples of 'rule violations' in music notation; some are purely graphical

curiosities, but many strike much deeper[4]. And I have commented elsewhere (Byrd 1994) that "It is tempting to assume that the rules of such an elaborate and successful system as CMN [Conventional Music Notation] must be self-consistent. A big problem with this idea is that so many of the 'rules' are, necessarily, very nebulous... But if you try to make every rule as precise as possible, what you get is certainly *not* self-consistent."

Obvious as it is, the fundamental difference between works of art and otherwise similar creative products that aren't intended as art is often forgotten. One result is the well-known quotation of uncertain origin (it has been attributed to people from Clara Schumann to Frank Zappa), "Writing about music is like dancing about architecture." This is certainly thought-provoking, but it's not a good analogy because writing about music is generally intended primarily to convey information and only secondarily (if at all) as art. For the same reason, the common variation, "Talking about music is like dancing about architecture", is an even worse analogy. Actually, *writing* [or *talking*] about music is like *writing* [or *talking*] about architecture. Of course, this formulation is not too exciting! Using arts throughout, we might put it like this instead: *Writing poetry about music is like dancing about architecture*. This is a more memorable statement, but really interesting only as a response to the misguided original.

2. **Music is fundamentally a performing art,** so that *performances*, *symbolic representations of general performances* (scores and performance parts), and *symbolic representations of specific performances* (transcriptions) all exist, and the relationships among them can be extremely subtle. (I say music is "fundamentally" a performing art to distinguish it from forms like poetry, which can be performed, i.e., read aloud by one person for the benefit of others, but rarely is.) To borrow terms from mathematics, the mapping is not one-to-one; it's often many-to-many. For one thing, what is the correct or most authoritative version of a musical work? For Western classical music, it's ordinarily a score, but there may be several that are equally authoritative, or it may be a performance—or a set of parts intended for performers. Bach's *B Minor Mass* is a well-known example of the problems (Frans Wiering, personal communication, January 2009; see also the Wikipedia article "Mass in B Minor"). And as you go back in time, the correct interpretation of any written notation gets more and more difficult[5]. For most traditions other than Western classical music, the most authoritative version is almost always a recorded performance, but there may be many performances with

[4] A particularly interesting one is an example from a Chopin *Nocturne* of what Julian Hook calls an "impossible rhythm", where one double-stemmed notehead is a triplet eighth in one voice but a normal eighth in another and both end at the same point (the barline); therefore they must begin at different times in the two voices! Hook (2008) lists dozens of examples in the works of Brahms, Chopin, Rachmaninov, etc.

[5] Along the same lines, Nina Fales observes (personal communication, November 2008) that "among ethnomusicologists, a commonly recognised reality of fieldwork is that the further in time and distance a particular item of information travels from its source, the less likely it is to communicate its original significance."

equally strong credentials. It's hard to see how one could decide what is the most authoritative version of the folk song *Greensleeves*. (Cf. Goodman 1976 or Talbot 2000.) Even for recent Western classical music, it's by no means unknown for the composer to record a performance that disagrees significantly from their own preexisting score.

The closest text analog of these first two characteristics together is the play, especially the play in verse.

3. **A tremendous amount of music,** especially Western music, **has complex synchronisation requirements.** This applies to most popular and radio and TV music as well as European art-tradition music. To my knowledge, it is not only unique to music among the performing arts and other physical activities of people (sports, etc.), but it may be more demanding than any other presentation of information intended for human consumption. The text equivalent is something like explicitly synchronised speeches in plays, which indeed occur. But there is simply no comparison: even the few examples of these (e.g., Churchill 1982) are far less demanding in terms of synchronisation than the vast majority of Western music of, say, the 14th century or later. (In some ways, a play in verse is more like a piece of music than one in prose, and plays in verse are an important category. But of course they almost never involve explicit synchronisation.) True, circus acrobats, jugglers working together, etc., do feats that require split-second timing, but at nowhere near the rate of musicians—and never, as far as I know, involving as many simultaneous coordinated events.
4. **Music involves "instruments", very often in groups.** This has several implications. (a) It makes arrangements and transcriptions of a given work possible. Of course, there are many different instruments, and therefore an enormous number of combinations. Enough of them are actually used that the Library of Congress Subject Headings include hundreds, if not thousands, of entries for different ensembles. (b) Versions of many works exist for players of different skill levels, mostly lower but often—for virtuosi who want to "show off"—higher than the original. (c) Music notation may represent the sounds to be produced or it may represent the actions the performers are to take to produce the sounds. Both are widely used. Conventional Western music notation largely represents sounds, though it has features (e,g., notation of artificial harmonics for string instruments) that represent actions. Tablature for guitar and similar instruments represents the players' actions ("put a finger in *this* position on *that* string").
5. **Music is often combined with text,** not only via singing but also in the cases of narration and background music, plus several usages in music-notation scores. As a result, the problems of handling music are in several ways a superset of the problems of handling text. For example, to use space efficiently, scores of orchestral works routinely show on a page only the instruments that actually play on that page; the instrument to play each staff is identified by a label at the left end of the staff ("flute 1", "fagotti 1 e 2", etc.). Therefore, to convert the score to symbolic form, an optical music

recognition program must first perform optical *character* recognition (OCR) on the label strings.

6. Finally, **music is extremely popular;** in many cultures including ours, it is among the most popular arts. If this were not the case, perhaps other phenomena would be as complex and challenging. But the great popularity of music overall means that really popular works are likely to exist in many recordings and many scores, in arrangements for several different ensembles of instruments, at multiple levels of technical demands on the performers. As a result, music wins the "most challenging" contest hands down. It also means that handling music's challenges is important in itself, even on purely economic grounds.

Implications for Computer Applications to Music

One of my own research interests is computer systems for music. For all the reasons above, it would be surprising if a computer system designed for music did not have features of great value to other domains. Of course this does not mean that we should expect every computer system for music actually to be useful in other areas. The trick is to keep general features of the system separate from music-specific features so the former can actually be used in nonmusic contexts, and few systems are designed that way.

Acknowledgement

The ideas herein about the quotation "writing [talking] about music is like dancing about architecture" originated in a 2001 e-mail conversation with Steve Larson. Thanks to him and to Douglas Hofstadter, analogy fanatic *par excellence*, who clarified some of my thinking about that quotation.

References

Byrd, Donald (1994). Music-Notation Software and Intelligence. Computer Music Journal 18, no. 1, pp. 17–20. [6]

Byrd, Donald (2009). Gallery of Interesting Music Notation. Available at http://www.informatics.indiana.edu/donbyrd/InterestingMusicNotation.html

Churchill, Caryl (1982). Top Girls. London: Methuen.

Collins, Nick (2006, Winter). Composing to Subvert Content Retrieval Engines. ICMA Array, pp. 37–41.

Goodman, Nelson (1976). Languages of Art: An Approach to a Theory of Symbols, 2nd ed. Indianapolis: Hackett Publishing.

Hook, Julian (2008). How to Perform Impossible Rhythms. Talk at Graduate Theory Association Symposium, Indiana University, Bloomington, Indiana, February 2008.

Talbot, Michael, ed. (2000). The Musical Work: Reality or Invention? Liverpool: Liverpool University Press.

[6] Available at:
www.informatics.indiana.edu/donbyrd/Papers/MusNotSoftware+Intelligence.pdf

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1987

3 Multimodal and Partial Music Synchronization

Meinard Müller (MPI für Informatik—Saarbrücken, DE)

Significant digitization efforts have resulted in large multimodal music collections, which comprise music-related documents of various types and formats including text, symbolic data, audio, image, and video. The challenge is to organize, understand, and search multimodal musical content in a robust, efficient and intelligent manner. Key issues concern the development of methods for analyzing, correlating, and annotating the available multimodal material, thus identifying and establishing semantic relationships across various music representations and formats. Here, one important task is referred to as music synchronization, which aims at identifying and linking semantically corresponding events present in different versions of the same underlying musical work. In this contribution, we give an overview of recent music synchronization techniques and discuss various challenges arising from the fact that different versions, even though similar from a semantic viewpoint, often reveal significant differences in aspects such as tempo, dynamics, and instrumentation. Even harder are the cases where the versions to by synchronized only reveal some partial similarity and differ, for example, with respect to their polyphonic structure or their musical form. Finally, we show how synchronization results can be used for various applications including content-based music retrieval, inter- and intra-document browsing, and performance analysis.

Keywords: Synchronization, alignment, multimodality, partial similarity

See also: Meinard Müller, Information Retrieval for Music and Motion, Springer, 2007

Part II

Workshops

Chapter 4

Workshop 1a: Cognitive Modelling, Pattern Discovery and Folk Song Research Moderator: Alan Marsden

4 Workshop Report

Daniel Müllensiefen (Goldsmiths College—London, GB)

The session included four talks two of which presented specific application of computer analysis tools to folk song research while the other two were concerned with music analysis questions and cognitive models. Zoltán Juhász presented an analytic approach whereby the similarity between folk song melodies can be mapped to reflect the relations between (musical) cultures. He uses pitch contour as his basic melodic representation and employs a neural network (i.e. a self-organising map (SOM)) to cluster the melodies from each musical culture into melody types. He then uses a second SOM to organise the melody types from several cultures and determine the overlap with regard to melodic structure. He interprets the resulting map that can be displayed graphically very nicely as map of 'deterministic connections between cultures'. A good portion of the discussion concerned the question whether different technical options in the modelling process (especially regarding the SOM) would have created different results.

Peter Grosche introduced a system that can aid in the analysis of non-professional audio recording of folk songs. The basic questions he aims to answer are how performances of the same song can differ between recordings and deviations from an idealised song template can be captured. In other words: How is a song brought to life? He chiefly considers pitch and tempo aspects and uses MIDI scores of the idealised songs to compare against. Central to all there comparisons is the synchronisation between audio recordings on one hand and audio and MIDI files on the other hand. This synchronisation process allows also for easy and meaningful annotations of an audio recording to highlight different aspects of the performance.

Martin Rohrmeier surveyed his framework for comparing experimental and computational approaches of musical training. He starts from the assumption that the well-studied psychological mechanism of implicit learning is of high relevance for music learning and the development of musical expertise. In order to make the training phase of computational models comparable to a typical experimental session in which humans participate, he puts various computational model (Markov models, grammatical parsers, neural networks, etc.) into same experimental framework. This framework includes a training phase and a testing phase. Questions that he poses and are also raised in the discussion include the conceptual inclusion of other memory effects such as interference and the abstraction level at which model and human answers are compared.

Christina Anagnostopoulou started from the questions how music analytic knowledge is represented and differs between a human analysis of a score compared to analytic listening and computational analyses. She summarised a study which included 12 music experts who analysed to (monophonic) pieces from score (parts from a cello suite by J.S. Bach and Berio's *Sequenza*). As results she observed the differences between the sub-groups consisting of historical musicologists (focussing largely on harmony and top level structures) and analysts with a more scientific background (rather bottom-up approach focussing on interval structures etc.). All experts paid a great deal of attention to rhythmic structures. From the results she is able to identify a few overall important analytic features including segmental descriptors, the coherence or distinctiveness of subsequent segments, rhythmic structure, *linked parameters* (i.e. the co-occurence of specific events in different musical dimensions).

The general discussion of this session evolved around a few important points:

- The sharing of music data was mentioned as one fundamental aim that would facilitate research as a whole in this area.
- Different types of models were discussed and the general question was raised whether computational models should behave like humans. As a consensus it was described as important that researchers define clearly whether they intend their models to behave like humans or not. This approach could also make modelling simpler since not parts of the parameter space would possibly have to be searched for certain human-like models.
- Often a distinction can be drawn between the representation of music that a model takes as input and how it represents it internally. As an advice it should be made clear on what level of representation the actual research is carried out and that any musical or analytic information that is not needed for the primary research goal is not thrown away.

Keywords: cognitive modelling, pattern discovery, folk song research

5 Analysis of different folk music cultures using self organising maps and principal component analysis

Zoltán Juhász (KFKI Research Institute - Budapest, HU)

A comparative study of different oral musical traditions requires a uniform and unbiased description of the national/regional folk music cultures. To do this, we developed a method mapping the contour of each melody to a point of a multidimensional metric space. In this model, the characteristics of a given musical culture are attributed to spatial characteristics of a point system which is built up by the melodies of a representative collection. The spatial characteristics are determined using principal component analysis.

Another tool of uniform description of different musical traditions is a system that determines the typical melody contours in a given set of melodies automatically, using a special form of artificial intelligences, the Kohonen's self organising map. In general, this method is able to classify melody collections from many points of view: interval and rhythm distributions can also be classified, and typical motives inside the melodies can also be identified.

Based on the above techniques, a comparative study of different musical cultures can be accomplished as follows:

– The analysis of the overlapping of the point systems corresponding to different musical cultures is able to determine common, as well as separate areas of the musical space. Thus, common and special musical principles can be distinguished using this analysis.
– We applied self organizing mapping for automatic classification of melody contours of 16 European and Asian folksong corpora. The cross-classification of the melodies using national/regional self organising maps detects those melody types that have variants also in the foreign culture, as well as special melody types that have no relationship "abroad".
– We characterized the strength of the contacts between musical cultures by a probability density function. We show that the relationships identify an "Eastern" and a "Western" sub-system that are associated due to the close relations between the Finnish and Irish-Scottish-English musical cultures to the Carpathian Basin. We also show that "Eastern" cultures define a very clear overlap of melody types, functioning as a common crystallization point of musical evolution.
– The results show an extended system of relations between different musical traditions, and propose the theory of certain musical parent languages. Using the above-mentioned techniques, the archetypes of these parent languages can also be reconstructed.

Keywords: SOM, contour analysis, multidimensional mapping, cross-cultural study

6 Computer-aided Comparative Musicology and the Music of a Mystic Islamic Sect (a case study for Zoltán Juhász presentation)

Janos Sipos (Hungarian Academy of Sciences—Budapest, HU)

I am an ethnomusicologist with comparative musicology at the centre of my interest and I have been carrying out research, mainly in Asia, for the past twenty years. I came to the seminar first of all to learn, observe and if it is possible to initialize common projects.

As I took my first degree in mathematics I am open to computer-aided methods and we have been working on analyzing, classifying and comparing (strophic and monodic) folk song materials. There are many unsolved problems here, let me now quote Béla Bartók: "I think that if we will have had sufficient folk music material and study at hand, the different folk musics of the world will be basically traceable back to a few ancient forms, types and ancient style-species."

In the last decades comparative musicology lost its hegemony, losing ground to (American) ethnomusicology, with its adoption of the same main question and sometimes the methods of the "social anthropologists": how do individual cultures function. However these days comparative musicology is regaining more and more strength, just by using computer aided tools.

In my presentation I show an application of Juhász's program system to the classification of the music of a mystic Islamic sect, we compare this musical world to the music of the neighboring people and we will see an example how the results achieved by different methods and areas of sciences help to achieve valuable conclusions.

7 Towards Automated Processing of Folk Song Recordings

Peter Grosche (MPI für Informatik—Saarbrücken, DE)

Folk music is closely related to the musical culture of a specific nation or region. Even though folk songs have been passed down mainly by oral tradition, most musicologists study the relation between folk songs on the basis of symbolic music descriptions, which are obtained by transcribing recorded tunes into a score-like representation. Due to the complexity of audio recordings, once having the transcriptions, the original recorded tunes are often no longer used in the actual folk song research even though they still may contain valuable information. In this paper, we present various techniques for making audio recordings more easily accessible for music researchers. In particular, we show how one can use synchronization techniques to automatically segment and annotate the recorded songs. The processed audio recordings can then be made accessible along with a symbolic transcript by means of suitable visualization, searching, and navigation interfaces to assist folk song researchers to conduct large scale investigations comprising the audio material.

Keywords: Folk songs, audio, segmentation, music synchronization, annotation, performance analysis

Joint work of: Müller, Meinard; Grosche, Peter; Wiering, Frans
Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1966

8 A framework for comparing different computational learning models with data from human experiments. And: how it becomes possible to learn "on the fly"

Martin Rohrmeier (University of Cambridge, GB)

As current work in progress, I am going to discuss a framework for evaluating a variety of computer models of learning in music with findings from (human) implicit learning experiments and show a couple of preliminary results.

There are several challenges for the evaluation of different machine learning models: one is to design an overarching framework that enables result comparison even though algorithms differ in their input and output formats, and their number and type of free parameters. A second challenge is to design the framework in a way that computational and human results can be related; therefore, computational results are gathered using comparable experimental conditions as in the human case.

As a second part, I will present a series of results that relate to a very interesting effect found with human subjects and computational models: subjects acquire structural features "on the fly" during a recognition or classification test, even if they had no or just insufficient exposure to the structures beforehand. Using the framework outlined above, computational modelling techniques indicate how this effect becomes possible and allow one to draw conclusions about the nature of the features that are acquired during such unsupervised online-learning processes.

Keywords: Machine learning, implicit learning, model evaluation, unsupervised learning

9 How do analysts think? Representations for semiotic music analysis

Christina Anagnostopoulou (Univ. of Athens, GR)

My talk will consist of two parts. In the first part I will present some work in progress, a short experiment with academic music analysts, who discuss two different musical pieces, by Bach and Berio. I will explain their descriptions and attempt a taxonomy of features they choose to describe the music, with an emphasis on segmental features.

In the second part I will show a computational analysis of a challenging piece of music by Xenakis, discuss the representation formalism chosen (multiple viewpoints), the musical knowledge, explain the problems we encountered, and show preliminary results in pattern discovery within a single musical piece.

Keywords: Music analysis, modelling, representations, multiple viewpoints, pattern discovery

Chapter 5

Workshop 1b: Representation, Encoding and Critical Editions
Moderator: Tim Crawford

10 Workshop Report

David Lewis (Goldsmiths College—London, GB)

This session was largely devoted to encoding formats and other mechanisms to assist in the interchange of musical information, with critical editions being, for the most part, a source of motivation, but not a topic of discussion. In focussing as we did on encoding formats, one runs the risk of rehearsing arguments and exploring issues that are as old as music computing itself, however the topic still has considerable importance, and those present were acutely aware of the need to progress without endlessly retreading old, well-worn paths.

Issues of the integrity of one's representation are vital to those who edit from source materials of any sort, but are most obvious for editors of music that uses notations badly served by traditional encoding formats. The first presenter, **Theodor Dumitrescu** [TD], is one such, and the motivation for creating a bespoke representation, *CMME XML*, centres not only on the ability to represent all the aspects of sources of mensural music that he considers significant, but also on the appearance of symbols that have remained in common currency in musical notation but whose meaning has changed. In the latter case, being able to encode the symbol in an interchange format may not be enough if there is no way of distinguishing the various uses of it over time and cultural and geographical space.

As TD pointed out, making as little interpretation as possible in the first, 'transcription' stage is a major goal for an editor. Christophe Rhodes suggested the importance of being able to specify the nature of interpretations, both implicit and explicit, so that later users can reason more usefully about the data. Tim Crawford made direct parallels with the practice in the primarily audio-based project OMRAS2 of representing processes as well as data (an idea also present in the work on CHARM by Geraint Wiggins who, alas, was not present).

Another issue that TD visited was the question of whether 'unusual' notational elements (that is, not central to CWMN) should be part of a core of an interchange format or of extensions to it. Perry Roland [PR] noted that design of the core of MEI is still in progress and most of what is potentially useful to encoders should go into the core, whilst Raffaele Viglianti [RV] discussed the mechanisms for extension that would subsequently apply as exemplified by TEI.

Where *CMME XML* is a specialised encoding format—designed to support a specific activity—and *MEI* aims to be universal and generalised— designed to support as many applications as possible, with as much information preserved—**Andreas Kornstädt**'s [AK] proposal, sketched out in his presentation, aims to model a generalised interchange without attempting the degree of completeness central to *MEI*. Although, as Joachim Veit would observe in his reflections, the precise form of AK's *Dagstuhl Core* developed over the course of the week, the principal concern that it aims to address remains the same: many potential 'clients' of music corpora do not require all the special information that appears in the authors' own internal representation of the data. More likely is that they will require quite common, simple musical information. A similar approach applied to corpus metadata is described by Daniel Müllensiefen for the corpora session below.

AK argued that a single interchange format 'hasn't worked and will not work' for political as well as technical reasons and taking the Dublin Core as a model, proposed a similar 'minimal' information set for music interchange. Compliance with the core would imply that this minimal set was implemented and that translator applications had access to them. It was generally observed and agreed that such a core would not represent any individual representation in its entirety, but could only form part of an interchange mechanism. Some concern was also raised (by PR) whether such an endeavour would interfere with the development of a single good standard (such as MEI).

Several people felt that a minimal form of exchange data would be beneficial, with Laurent Pugin [LP] noting that such a small, common subset might assist the uptake of MEI (c.f. TEI), whilst TD noted that many researchers were currently using MIDI-like data of absolute pitch (usually key number) and time (usually clock time) only and would be happy if all corpora at least provided that. TD also noted that there was a limit to the musics for which even this degree of basic information could be provided in an entirely uncontroversial way.

Unfortunately, no decisive discussion established firmly the applications that such interchange would support and, as such, the workshop discussion could not progress to what would form such a core and how this would be decided. AK's more pragmatic suggestion, to look at the list of features currently supported by existing encoding formats, and limiting notation to mainstream CWMN forms was later followed.

11 CMME and inter-format compatibility with non-CWN music

Theodor Dumitrescu (Utrecht University, NL)

The CMME Project (Computerized Mensural Music Editing, www.cmme.org) carries out ongoing development of a web-based software platform for publishing digital editions of repertories in "mensural notation" (used in 14th-16th-century Europe). Although mensural notation is the direct ancestor of Common Western

Notation, there are enough conceptual differences to make it not only fairly impractical to combine the two in a single representation grammar, but even potentially fundamentally misleading—due to conflation of notational elements whose functions and meanings changed radically over time while appearing largely the same. For this reason, CMME editions are based upon an XML encoding specific to mensural notation (CMME-XML), while custom viewing software allows users to visualize the scores in numerous forms of notational translation without changing the underlying representation.

A question which has arisen repeatedly over the years is whether the CMME musical data could be brought under the umbrella of a more general XML representation such as MEI or MusicXML, ensuring immediate compatibility of scores and software for widely different repertoires.

Since the basics of the CMME format are now fairly stable—encompassing not only representation of notational elements but also editorial interpretative data and variant reading encoding for philological analysis—the time is ripe for a reconsideration of how such an action of format-conversion could be envisioned and implemented. Can it be done without conceptual modification of the data already encoded with the system? If not, where are the incompatibilities to be found? How can we weigh the ultimate advantages and disadvantages of the conversion?

12 From Feet of Clay to the Dagstuhl Core

Andreas Kornstädt (Stanford University, US)

This is a kind of preview of what might come out of a discussion and initiative started during this Dagstuhl seminar.

It is evident that musicologists / music theorists are confronted with difficult choices when having to decide which music data formats to employ as a base for their research. Some formats are excellent for layout (such as Leland Smith's SCORE), for analysis (such as David Huron's Humdrum) or for standardised parsing (any XML-based format). For some there are lots of tools (Humdrum again) and for others there are lots of data available (CCARH's MuseData database stands out here). Finally, some corpora of very high quality (MuseData, SCORE from publishers) while others have vastly different quality levels (online MIDI archives). In many cases, researchers are overwhelmed by the number of choices to make or confused with by the number of details entailed by choosing one particular format over another so that they start over and come up with yet another "single-use" specialized format just for their particular study. This "throw away" culture greatly hampered research because quite often over half of the time of a research endeavor is spent settling the "format question" and / or inputting new data. Many a project isn't started at all because it appears unrealistic to encode a sufficiently large corpus of music (or to obtain it).

A central problem is the absence of reliable, high quality converters between data formats and descriptions of what they actually can convert from A to B - and more important: what they cannot. The aim of this initiative - taking cues from the "Dublin Core" initiative in the realm of library catalogues - is to provide a a core feature set of music data that is sufficiently complete for a wide range of Western music from 1750 to 1935 and to describe formats and converters between them with regard to the features. Formats not providing a means of to deal with the vast majority of these core features could be considered outside the "Dagstuhl Core" (formats for encoding monophonic music, insufficient analytic insufficient analytic encoding capabilities, etc.)

We plan to publish the results in a public Wiki so it can be used and extended by researchers.

Keywords: Representations, data formats, core, interchange

13 Drafting the Dagstuhl Core

Eleanor Selfridge-Field (Stanford University, US) [7]

Andreas Kornstädt proposed that as a means to facilitate the use of existing polyphonic corpora in common Western notation (1650-1935) that those interested form a consortium to adopt and promote a core set of features.

The Dagstuhl Core, as it was provisionally named, would serve as a frame of reference for parties interested in using encoded music, as well as creating corpora, formats, and converters. It would educate users about the possibilities and limitations of application-independent music data formats and the quality of programs that convert between them. It would provide a feature list for each format/converter containing the indications "Yes"/"No"/"by Extension" to indicate how the given feature can (not) be realised in a given format / by a given converter. The emphasis would be less on completeness than on essential features that any representation approach intended for use in scholarly work should be able to handle.

Over continuing discussion in ensuing days, a list of Core features was developed by a group including Ted Dumitrescu, Johannes Kepper, Andreas Kornstädt, Daniel Röwenstrunk, Perry Roland, and Eleanor Selfridge-Field. Additional input was received from Craig Sapp, who has translated extensively among four of the five data representation schemes (Humdrum Kern, MuseData, SCORE, and MusicXML) compared on the "core" feature list. (At the present time, no software to implement MEI is available.) The representation schemes were chosen because of the extensive repositories of music that already exist in them.

[7] This brief report summarises activities following on from the above workshop which culminated in the develoment of draft documents for a Dagstuhl Core

There was some latitude in viewpoints on the extent, elasticity, and application areas to be served by the Core but there was consensus on maintaining a focus on scholarly research and applications (printing, analysis, editing of virtual materials, and queriable data). Some ability to interchange data among applications was also frequently mentioned.

The list is a work in progress. A wiki (access currently restricted) has been set up at http://muwimedial.de/dagstuhl-core/. After a short phase of refinement, it will be made circulated for public comment. A provisional table of features is shown elsewhere in this proceedings.

Table of features: http://drops.dagstuhl.de/opus/volltexte/2009/1971/

Chapter 6

Workshop 2a: Synchronisation, alignment and performance analysis
Moderators: Sebastian Ewert, Meinard Müller

14 Case Study "Beatles Songs"—What can be Learned from Unreliable Music Alignments?

Sebastian Ewert (Universität Bonn, DE)

As a result of massive digitization efforts and the world wide web, there is an exploding amount of available digital data describing and representing music at various semantic levels and in diverse formats. For example, in the case of the Beatles songs, there are numerous recordings including an increasing number of cover songs and arrangements as well as MIDI data and other symbolic music representations. The general goal of music synchronization is to align the multiple information sources related to a given piece of music. This becomes a difficult problem when the various representations reveal significant differences in structure and polyphony, while exhibiting various types of artifacts. In this paper, we address the issue of how music synchronization techniques are useful for automatically revealing critical passages with significant difference between the two versions to be aligned. Using the corpus of the Beatles songs as test bed, we analyze the kind of differences occurring in audio and MIDI versions available for the songs.

Keywords: MIDI, audio, music synchronization, multimodal, music collections, Beatles songs

Joint work of: Ewert, Sebastian; Müller, Meinard; Müllensiefen, Daniel; Clausen, Michael; Wiggins, Geraint

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1964

15 Exploring Multiple Instances of Multimedia in Multiple Modes

Donald Byrd (Indiana Univ.—Bloomington, US)

It is widely accepted that *multimodal* access to music is important; music is so complex that no one way of looking at it is adequate. The world also needs access to multiple *instances* simultaneously. People constantly want to compare related (or possibly related) music, often within a single work or movement of a work. Similar statements hold for many fields besides music: witness the ubiquitous "diff" programs (and word-processor commands) for text, and the recent appearance of such features for XML. To enable comparison of music in symbolic form, Ian Knopke and I have worked towards a "MusicDiff" (Knopke and Byrd 2007). But a system that explicitly supports multimodal access to multiple instances is very desirable, if it can be done without limiting the modes available.

I'm in the early stages of developing a system, the General Multimedia Workbench (GMW), intended to address this situation. The "M" in "GMW" originally stood for "music". But music is so challenging in so many ways (Byrd 2008) that an appropriate design can be adapted for a tremendous range of other uses, from learning an operatic role to creating multimedia shows to learning to juggle to studying graphic novels. I believe the GMW has an appropriate design.

To clarify the value of simultaneous multimodal access to multiple versions, I'll briefly compare several versions of a very conventional and relatively simple piece of Western tonal music, the U.S. national anthem, the Star Spangled Banner. The versions I'll consider are some or all of:

1. a standard version in encoded CWMN (NB: "CWMN" is my favorite acronym for conventional Western music notation; I hope others will adopt it!)
2. a recording of a standard version
3. a recording of Jimi Hendrix's Woodstock improvisation on it
4. a published transcription of Hendrix's version in both CWMN and guitar tablature
5. as time allows, encoded CWMN for some silly variations of my own, inspired by David Levitt's challenge many years ago to rewrite the Star Spangled Banner to reduce its melodic range to an octave
   The encoded CWMN forms are files for my Nightingale score-editing program, so text views of the encoding and MIDI forms will be available as well as display of the notation. (I would have been reluctant to use this particular piece of music before our recent presidential election...)

References

Byrd, Donald (2008). Why Studying Music is Both Difficult and Important. (see above, 2)

Knopke, Ian, and Byrd, Donald (2007). Towards MusicDiff: A Foundation for Improved Optical Music Recognition using Multiple Recognizers. In Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007), Vienna, Austria, pp. 123-124.

Chapter 7

Workshop 2b: Computational musical analysis, mathematical music theory Moderator: Ian Knopke

16 Simple Pitch Class Cycles and Western Tonal Scales

David Meredith (Aalborg University, DK)

The diatonic pitch class set (i.e., the pitch class set associated with the normal major scales and the church modes) has many special mathematical properties that have received a great deal of attention from music theorists. However, traditional tonal music theory recognizes not only the diatonic scales but also non-diatonic scales such as the harmonic minor scale, and the ascending melodic minor scale. Some music theorists (including Schenker) also recognize the importance of other scales such as the "harmonic major" scale which is a combination of a major lower tetrachord and a "harmonic" upper tetrachord. Nevertheless, there seems to be little work to date on exploring the mathematical properties of these non-diatonic tonal scales.

In this paper, I present what I believe to be a new result that suggests that the emergence of the common non-diatonic scales is a natural consequence of tonal music being based on the pitch intervals present in the diatonic set. The result is based on the idea of a "k-spectrum" (Clough and Douthett, 1991). The 1-spectrum of a scale is the set of pitch class intervals associated with scale steps. For example, the 1-spectrum of the diatonic set is 1,2 because every diatonic step is either 1 or 2 semitones. Similarly, the 2-spectrum of a diatonic set is 3,4 since a third (i.e., a leap of two scale steps) in a diatonic scale is always either a major or a minor third and thus consists of either 3 or 4 semitones. The set of all k-spectra for the diatonic set is called the "diatonic spectrum set" or simply the set of "diatonic spectra".

In this paper, I show that the pitch class sets associated with the smallest simple cycles whose associated interval sets are diatonic spectra are either diatonic sets or the pitch class sets associated with the common non-diatonic tonal scales (ascending melodic minor scale, harmonic minor scale, harmonic major scale and "whole-tone-plus-one" scale). This result suggests that these common non-diatonic tonal scales may be an inevitable consequence of tonal music being based on the diatonic set.

Keywords: Diatonic set, pitch class set theory, musical scales

17 Melodic Search in Large Collections of Medieval Music

Ian Knopke (Goldsmiths College—London, GB)

Recent studies of medieval music have identified the importance of small, ornamented, memorized fragments in the compositional process, somewhat contradicting the traditional view of how such music is constructed using counterpoint rules. While small searches have been undertaken with great success, to date no one has conducted such a search across a very large corpus. This presentation discusses a set of techniques for efficiently discovering motives and patterns across the entire collection of Palestrina masses. There will also be some discussion of the author's PerlHumdrum and PerlLilypond systems that form the basic technical tools of the system, and of pattern-matching techniques for music in general.

Ian Knopke is a music technician, theorist, performer, and composer from Canada. He received his Ph.D. from McGill University, researching search engines for sound and music on the Internet, before teaching and working as a researcher at Indiana University in Bloomington, Indiana (USA). Ian is currently a researcher in the Intelligent Sound and Music Systems (ISMS) group in the Center for Cognition, Computation and Culture at Goldsmiths University.

Keywords: Music, humdrum, melodic search

18 Towards quantitative characterization of artistic individuality in music performance

Bruno Gingras (Goldsmiths College—London, GB)

As part of a larger project studying expressive strategies in organ and harpsichord performance, I am developing quantitative methods for describing differences and similarities between the expressive strategies used by different performers. For instance, I seek to compare the degree to which different performers modify their patterns of articulation, velocity, and/or tempo when asked to emphasize a different voice in a polyphonic piece. Because its aim is to identify what is unique to each performer's style instead of focusing on general tendencies, this research is central to the characterization of artistic individuality in music performance. One obvious difficulty that arises when analyzing the contribution of different musical parameters to a performer's global expressive strategy is the fact that they are generally measured on different scales. I will discuss the statistical approaches I have developed in order to circumvent this problem as well as related issues. The presentation will lead to an open discussion about the relative merits of various approaches to the study of artistic individuality in music performance, while raising questions regarding the most appropriate way(s) of representing the performance data thus analyzed.

Keywords: Performance; expressive strategies; artistic individuality

Chapter 8

Workshop 3a: Electronic Editions Moderator: Joachim Veit

19 Workshop Report

Tim Crawford (Goldsmiths College—London, GB)

Joachim Veit [JV], as chairman of this session, opened proceedings with a 'Lamento' concerning his own experiences over a decade of work on the Carl Maria von Weber complete edition. The edition began 10 years ago with the intention to produce 50 volumes of Weber's music by 2026 but with no idea of considering electronic publication. The intention now is provide a very comprehensive 'apparatus' of related contextual material (much of it non-musical) as well as electronic music editions and facsimiles of sources. This puts a huge burden on the editors (and also on the publisher) without any guarantee that users will 'read' the edition any better than a conventional publication.

Johannes Kepper [JK] examined the nature of electronic scholarly music editions, posing seven questions which need to be resolved:

- What is the editor's role in a digital edition?
- What are the qualities of participative/collaborative concepts of an edition (cf. Web 2.0)?
- How do digital editions serve scholarly and/or practical needs?
- What is the difference between a (commented) archive and an explicit edition?
- How should we deal with the historical context of the edited works from a conceptual perspective)?
- How do/should encodings and facsimiles relate?
- What are the actual benefits of digital editions?

These questions provoked a lively discussion. Ted Dumitrescu [TD] felt that the editor's role should be much the same—in terms of critical editing—as with printed scholarly editions. Eleanor Selfridge-Field [ESF] stressed the distinction between a digital surrogate for a book and a 'born digital' edition; she also pointed out that the activities of publishers can be very counter-productive (for valid or imagined business reasons). Furthermore, she mentioned the increasing reluctance of students to seek 'best' rather than 'acceptable' (or perhaps 'accessible') editions.

JV mentioned the issue of facsimiles as principal reading material (for which demand seems to be increasing)—readers need guidance. Perry Roland [PR] (from his librarian's perspective) spoke about the important distinction between 'archives' and 'editions' (as raised by JK). ESF thought that musicology in general is rather poor about disseminating information about projects.

PR put forward the view that (in several ways) Digital Humanities [DH] presented a good example for musicology. To which Laurent Pugin [LP] and ESF (without disagreeing with PR's advocacy of the DH model) gave examples of the additional complexity of music encoding and representation, while JK and Raffaele Viglianti [RV] discussed the problem of navigating the vast amounts of data involved. JK thought that the editor should act as a 'guide' through this maze, while RV mentioned the possibility of this role being taken by informed performers (as, perhaps, in the Online Chopin Variorum Edition [OCVE]).

Tim Crawford [TC] raised the significant methodological distinction between editing works with multiple sources (as in the Weber edition) and those with a single, more-or-less corrupt, single source (as in medieval music, or that for lute). Basically—as TD pointed out—the editorial process and skills are the same as always, but presentation can be much richer in a digital edition (if encoding is done scrupulously, perhaps 'transcription' should be thought of largely as a presentation issue rather than an editorial one). However, ESF said that her experience suggested that digital editions mean more work for editors than conventional ones.

TC pointed out that some performers of historical repertories (e.g. lute players) preferred to make their own 'editions' or use facsimiles exclusively. This led to a discussion of collaborative editions—e.g. the OCVE—and the possibility of maintaining an 'audit trail' of contributions to such an edition. The role of the editor then could become more like the moderator of a Wiki. One might have (perhaps will, inevitably, have) different collaborative editions for different communities.

After coffee, **Laurent Pugin** [LP] introduced his work with Ichiro Fujinaga and others on the Marenzio edition, working from scanned microfilms of printed 16th- and 17th-century partbooks in which mensural notation is type-set. This uses LP's own Aruspix software for optical music recognition as well as collating the printed sources to discover detailed typesetting variants that may occur between different exemplars of the same book. These variants can contribute to stemmatic analyses of the transmission patterns in the early sources. Again, the basic role of the editor is the same, but the task for the next editor is made easier and the data is richer.

TC raised the problem of IPR in digital editions; ESF recounted various issues raised by her work at CCARH. This led to some discussion of licensing, analogies with the Open Source movement, etc.

In the final wide-ranging discussion: TD thought that there needs to be more corpus-based analysis, drawing the analogy with current text-based work in DH; TC suggested that we need to persuade our sceptical traditionalist colleagues gently, JK suggesting that we (and our automated tools) should offer suggestions rather than definitive editorial decisions; ESF spoke of the complications raised by multiple sources whose variants seem to arise from circumstances of performance—for example, different editions of Vivaldi's music in which the different bass-figurings suggest different harmonies.

20 MEI Study Group

Johannes Kepper (EDIROM—Detmold, DE)

During the Dagstuhl Seminar on "Knowledge Representation for Intelligent Music Processing", a small (but vitally discussing) group of scholars was concerned with different approaches to the encoding of music notation from a critical editor's perspective. Major points in these discussions were the question of data interchange between different encoding systems (where there was an overlap to the discussions about Andreas Kornstädt's idea of a "Dagstuhl Core" description of music notation file formats), the possibility of finding common structures in all kinds of music notation (from different ages) that might offer a basic level of interchangeability between supported formats (this topic was mainly addressed by Ted Dumitrescu, Laurent Pugin and Perry Roland), and the especially the discussion about MEI as a potential model for editorial / scholarly music notation encoding.

Our several meetings lead to a consent that MEI seems to be a good starting point for further activities in this last sector. Therefore, a small Study Group for the future development of MEI was founded by Laurent Pugin, Perry Roland, Daniel Röwenstrunk, Joachim Veit, Raffaele Viglianti and Johannes Kepper. For a better communication of the activities around MEI a mailing list[1], a wiki[2] and a new Website[3], currently redirecting to the old page at UVa) were set up. Although the mailing list and the wiki are not open to the public yet, everyone is cordially invited to join and contribute to this group (a personalized login is available on request, please contact the author of this summary).

Important goals of the Study Group are a transition from DTDs to a modern Schema language (probably ODD), providing a comprehensive documentation with best practice examples, speeding up the general development, and finally offering a first stable version in early 2010. Although this is definitely very ambitious, a file format explicitly modelled on editorial problems seems to be a prerequisite for "born digital" editions and is therefore absolutely essential for the future of music editing.

21 Assembling Early Music Partbooks

Laurent Pugin (Stanford University, US)

1 https://lists.uni-paderborn.de/mailman/listinfo/mei-l
2 http://www.muwimedial.de/mei
3 http://www.mei-c.org

Aruspix is a software application for music editors that facilitates the comparison of early music editions when compiling comprehensive critical modern editions. It includes a collation workspace that enables different editions and re-editions of a book to be collated automatically. It is based on the digital representation of the content of the books generated by an optical recognition process, another feature of Aruspix, which enables the parts to be collated even when the layout is not the same. As most of the repertoire of the time has been published in partbooks, the task of the editor also requires the parts to be assembled and aligned. Having this task performed automatically (or semi-automatically) in Aruspix would not only expedite the editorial process, but also greatly enhance the feedback provided to the editor when there are differences between the editions and when editorial decisions have to be made.

Assembling the parts, however, is not a trivial task as it means determining the actual duration of the notes of the mensural notation, which requires different levels of knowledge. The first level is the notational level where the rhythmic context must be taken into account in order to be able to apply the basic mensural rules (alteratio, imperfectio, coloration, proportions). The second level encompasses a wider historical context, which is often necessary for solving typical ambiguities of symbols. Finally, the third level, by far the most difficult one, would have to incorporate a wider analytical context that has to be taken into account when errors appear in the text and when it has to be corrected. Integrating the first level of knowledge by implementing a set of mensural rules seems to be fairly straightforward, but I would be interested in designing a knowledge representation that would eventually allow the other levels to be incorporated as well.

Keywords: Early music; partbooks; mensural notation rules; score assembling

22 Introduction: MEI Goals

Perry Roland (University of Virginia, US)

This presentation provides background information on the goals of the Music Encoding Initiative.

Keywords: Music notation representation XML

Joint work of: Roland, Perry

23 Redesign of MEI using Literate Schema Design

Perry Roland (University of Virginia, US)

The Music Encoding Initiative will no longer use DTD syntax to express the MEI XML application. Instead, plans are being made to use literate schema design, most likely using the ODD (One Document Does it all) system designed for creation of the Text Encoding Initiative (TEI) guidelines. Key aspects of this approach include support of multiple schema languages; facilities for interoperability with other vocabularies; and facilities for user customization and modularization.

Keywords: Music notation representation XML RelaxNG TEI ODD

Joint work of: Roland, Perry

Chapter 9

Workshop 3b: Harmony and Pattern Discovery Moderator: Martin Rohrmeier

24 Towards meaningful segmentation and analysis of tonal harmony

Bas De Haas (Utrecht University, NL)

Background: The computational analysis of musical harmony is receiving a lot of attention the last decades. Even more so, because Music Information Retrieval (MIR) could greatly benefit from automated analyses of harmony. Recently, we presented a similarity measure for tonal harmony (Haas et al., 2008). This similarity measure is based on Lerdahl's Tonal Pitch Space (TPS; Lerdahl, 2001). The central idea behind the measure is to compare the change of chordal distance to the tonic over time. This is done by calculating the TPS distance between each chord of the song and the key of the song.

Plotting the chordal distance against the time results in a step function. The difference between two chord sequences can then be defined as the minimal area between the two step functions f and g over all possible horizontal shifts t of f over g. If discrete time units are used, this minimum can be calculated in $O(nm)$ time where n and m are the number of chords in the two matched songs. Finally, the minimal area between two step functions is normalized by the length of the shortest step function. We named our distance the Tonal Pitch Step Distance (TPSD).

Open issues:

- What is a good way to segment a song? When analyzing the harmony of a piece, the horizontal as well as vertical grouping of musical elements is crucial for a valid analysis. Most segmentation algorithms known from the literature deal only with homophonic music. How could these algorithm be extended to polyphonic music?
- What constitutes "harmonically similar"? Typical examples of harmonic similarity are cover songs, similar harmonizations of the same chorale melody, live performances of studio recordings, blues songs or songs based on rhythm changes, but what are the main criteria that human listeners use to determine if a two songs are harmonically related?
- What is the best way to evaluate these kind of measures of harmonic similarity? Retrieval results can very well be evaluated quantitatively on large datasets. However, the amount of available harmonically similar pieces are often relatively small. What is then the best way to analyze retrieval results?

– What are the main challenges, (dis-)advantages and caveats of porting the above mentioned problems of harmonic analysis and segmentation from the symbolic to the audio domain? For example timbre is often mentioned as an important cue for segmentation, but is usually disregarded in symbolic analyses.

Keywords: Harmonic Similarity, Music Segmentation, Tonal Pitch Space, Music Information Retrieval

25 Cross-modal structural analysis

Olivier Lartillot (University of Jyväskylä, FI)

Computer automation of structural analysis of music remains an open topic that has been investigated in both symbolic and audio domains. Previous studies have highlighted the close articulation between the complementary paradigms of local segmentation and pattern repetition. Computational models in the audio domain have concentrated in particular on the analysis of self-similarity matrices.

I am trying to develop a refined pattern description of audio recordings offering a detailed and multi-level description of phrase repetitions, encompassing motivic texture, phrasal structures and global form. The combinatorial redundancy of structures, inescapably engendered within such paradigm, needs to be curbed without weakening the richness of the results.

For that purpose, the model of articulation between segmentation and pattern extraction we recently proposed (Lartillot, Ayari, 2008) is adapted to audio within the similarity matrix representation. The main principle motivating the methodology is based on a generalization of solutions initially developed in specific fields of expertise, but that prove to offer valuable answers to more general questions.

This study in progress should result in a new tool for visualization of phrase repetitions, motivic content and general form in musical recording from various genres. As such, the tool, due to be integrated in a future version of the MIRtoolbox environment, should supplement the assortment of computational tools for musicological analyses. The general methodology may also guide further developments of computational modeling of music cognition processes.

Reference:

Olivier Lartillot, Mondher Ayari, "Segmenting Arabic modal improvisation: Comparing listeners' responses with computer predictions", Conference on Interdisciplinary Musicology, Thessaloniki, 2008.

Keywords: Phrase extraction, structural analysis, motivic analysis, local segmentation, pattern repetition, audio analysis

Joint work of: Lartillot, Olivier

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1968

Chapter 10

Workshop 4a: Theory and Cognitive Science Moderator: Christina Anagnostopoulou

26 Workshop Report

Bruno Gingras (Goldsmiths College—London, GB)

The workshop featured a presentation by Daniel Müllensiefen (Goldsmiths College, London), on the mental representations of melodies (the so-called M4S project for Modeling, Music Memory, and perception of Music Similarity). The aim of the project is to develop a cognitive model explaining human musical behaviour using a set of melodic features. The melody features considered in this project include summary features, such as pitch, interval, and harmony; characteristic text constants, derived from an analysis of sequence of intervals (this analysis is based on ideas from computational linguistics); statistically weighted text constants, and melody term frequency weights, with local and global weights reflecting the importance of short melodic elements. The model can be used to predict court decisions on song plagiarism, or whether some melodies are easier to remember than others. Using linear and logistic models, tree models, and similarity algorithms, the model may also be used to predict which songs will enter the charts. Eventually, the model could be applied to study other aspects of human musical behaviour, such as amusia, karaoke, or memory for lyrics.

There are still a number of open questions about the project. For instance, should the model of feature co-occurrence be built before the cognitive behaviour is modeled? How can the model deal simultaneously with features and features frequencies? How can it combine sequence-based features and summary features? Another challenge is to derive densities from awkward frequency distributions.

The discussion about the project centered on the presentation of the melodic features (how are individual melodic terms such as pitch intervals and duration ratios derived?) and the cognitive relevance of the statistical features and melodic terms thus identified. It was also asked whether the model could be used to develop a unified theory of melody.

The open discussion for the workshop focused on the integration of cognitive science and traditional music theory to create a 'cognitively enhanced music theory'. It was argued, however, that the goals of traditional music theory, namely analysis and interpretation, are not necessarily compatible with the goals of cognitive science (developing a model of music perception and cognition). This led to a more philosophical discussion regarding the ontological basis of music theory: should it be normative or descriptive? Should it emulate the principles of scientific enquiry? More concrete topics were also touched, such as the development of repositories of data, and the application of recent developments in music cognition research to music education and music recommendation systems.

27 Mental representations of melodies

Daniel Müllensiefen (Goldsmiths College—London, GB)

My current research project M^{4}S deals with how melodies are represented in memory and whether we can predict memory performance and memory failure in psychological experiments regarding melodic memory. We are currently using a feature-based approach where features are algorithmically extracted from monophonic melodies. From these abstract features and from information about their commonness in a large corpus of pop music we try to predict correct and false memory performances. This fundamental research on melodic memory has some interesting potential applications from music analysis to music information retrieval.

The two open issues I would like to discuss concern the organisation of the algorithmic feature toolbox and the prediction model that can make use of melodic feature values, feature frequency with respect to a corpus, and possibly information related to the human subjects.

Keywords: Melody;memory;automatic analysis;corpus-based analysis;

Chapter 11

Workshop 4b: Data structures, APIs
Moderator: David Meredith

28 Abstraction in Music Representation: a brief introduction to AMuSE

David Lewis (Goldsmiths College—London, GB)

I provide a short description and demonstration of the Advanced Music Score Encoding (AMuSE) toolkit, explaining its underlying philosophy and abstract data types and illustrating some of its more unusual functionality.

29 Towards robust geometric music processing using generalized representation

Kjell Lemström (University of Helsinki, FI)

For processing symbolically encoded music, geometric modelling seems to be in many ways superior to string modelling, for instance, in its inherent capability of dealing with polyphony and ignoring extra intervening elements, such as musical ornamentations. However, current geometric methods lack efficient ways to deal with some relevant musical properties that a string matching framework deals with rather more naturally. Firstly, it is a rather fundamental need to have global time-scale invariance (without losing the capability to deal with local time fluctuations) and, secondly, as the pitch intervals are often not precise, generalized pitch interval classifications (e.g., the contour representation) are needed in many applications.

We show that it is possible with the geometric framework to deal with these properties at the representational level as it is done when using the string matching framework. Moreover, the changes in the representation are done without losing the aforementioned important properties of the geometric framework.

Chapter 12

Workshop 5a: Applications and systems; musicological method Moderator: Jörg Garbers

30 Workshop Report

Alan Marsden (Lancaster University, UK)

This report concerns a workshop held at the Dagstuhl seminar *Knowledge Representation for Intelligent Music Processing* on the afternoon of Thursday 29 January 2009. The moderator of the workshop was Jörg Garbers.

The objective of the workshop was to consider examples of software applications and systems which have or could be actually used in musical research or musical activities. A secondary topic for consideration was appropriate and effective methods of musicological research in the light of such technological developments. This might be musicological methodology which might be subject to change because of novel possibilities, or it might be computer-science methodology in the still new area of research on musical software technology.

The workshop began with two presentations of particular projects. The first, by Meinard Müller, concerned a forthcoming experiment using *SyncPlayer* technology, developed as part of a larger project on synchronisation described elsewhere in these proceedings. An easy-to- use piece of software allowing switching between synchronised performances will be at the heart of the project. The project aims to

1. provide a novel feedback mechanism for use in music education,
2. test the usability and effectiveness of the software interface, and
3. investigate whether Music-Information-Retrieval tools have an application in music education

In the experiment ten student pianists will make recordings of Beethoven's *Pathétique* sonata (in both MIDI and audio format). *AudioSwitcher* software will synchronise the performances of all ten, and allow synchronised switching of playback between any of them. The students will be given tasks to perform using the software (such as determining which of the ten performances was their own). Data will be collected on the usability and effectiveness of the interface for the performance of the tasks.

Subsequent discussion concerned the details of the data to be captured in such an experiment. For example, some considered that it would be desirable to also capture video of the performances, or some other form of motion capture.

Other research on musical performance has shown that information about a performance can sometimes be more easily extracted from motion data (whether from video or captured in some other way) than it can be from audio. There was also discussion of the need for comparative data. Should not such an experiment compare how easy or difficult the tasks are to perform using existing technology or by other means? Finally, comparisons were made with the software *Sonic Visualiser* (from Queen Mary University of London) which also has a tool for synchronising performances and switching between them in playback.

The second presentation was from Jörg Garbers, concerning software frameworks for music processing. This is covered elsewhere in these proceedings, so description here is not necessary. The point stressed for this workshop was that the integration between different pieces of software was made possible by the fact that all were scriptable, though scriptable in different ways.

General discussion then concentrated on the following issues.

What makes a good interface for musicologists?

The impact of an interface on a novice user (and most musicologists are novice users) is enormous. Visualisations and sets of results which are immediately comprehensible to the user are much more effective than ones which require explanation or some kind of study before they can be interpreted. Musicologists already use a number of kinds of graphic portrayals of information (music notation, diagrams, etc.); interfaces which make use of these are more likely to be effective.

Pictures and music examples have a huge impact, partly because (in the terms of the preceding paragraph) they are immediately comprehensible. It is generally worth the effort of incorporating such things into an interface for musicians. Maps are also a good method of giving an interface into or visualisation of a collection of items or data

How can we best stimulate involvement of musicologists?

Most research in the field would benefit from greater involvement from musicologists, but past experience of those involved in projects in this field has shown it to be difficult to get musicologists engaged. Demonstrations and 'low-hanging fruit' (easily gained benefits) are good ways of making clear to musicologists that there is value in this field. One of the successes of *Humdrum* is that it has a large body of demonstrations associated with it which musicologists can consult and which make clear the benefits of the software. Musicologists are not generally technology-averse; they already often use software. However, they might reasonably be expected to be averse to learning new pieces of monolithic software. Plugins for existing software or adaptable tool sets are more likely to be adopted.

How can we best achieve effective co-operation?

It would be better sometimes to get input from musicologists at the beginning of a project rather than soliciting responses to a prototype. On the other hand, musicologists are not generally well equipped to give input to technological projects at an early stage.

Probably musicologists and computer scientists both need to learn more about each others' disciplines, and about how each other thinks. Musicologists need to learn to *operationalise* musicology and come to understand concepts, such as representation, whose precise meaning in the context of a technological project is likely to be different from their meaning in a purely musicological project. Computer scientists need to learn more about musicology and its preoccupations. (There are plenty of musicological questions which appear susceptible to computational study.)

At the root of such questions are issues of human-computer interaction, which most of us take for granted. Probably we should learn more about the subtleties of this interaction.

Problems

Two problems which often hamper effective projects were identified, but nobody had any potential solutions to offer:

1. User-interface development is time-consuming. Researchers understandably want to devote their time instead to the development issues specific to the musical problems at hand.
2. Musicology and computer science operate by different paradigms for good (and bad) reasons. Joint research will always run up against these differences.

31 Software frameworks for systematic music processing

Jörg Garbers (Utrecht University, NL)

Software frameworks are the key to systematic music processing. Thinking in frameworks makes it easier to provide other developers, expert users and end users the functionality they want or deserve (which is not always the same). And music processing is hard enough, because of its many aspects.

I will present software frameworks from the MaMuTh (Mathematical Music Theory) project and from the WITCHCRAFT (What Is Topical in Cultural Heritage: Content-based Retrieval Among Folksong Tunes) project. I would like to discuss the level of detail at which we should conceptually design, share and recombine components: representations, measures, algorithms, user interfaces.

To ease technical realization I am interested in existing frameworks, best practices, their benefits and practical limitations.

Keywords: Software framework, generic components

Chapter 13

Workshop 5b: Corpora of music and results Moderator: Daniel Müllensiefen

32 Workshop Report

Daniel Müllensiefen (Goldsmiths College—London, GB)

The session discussed the various issues that are related to creating, maintaining, researching, and the making available of music corpora for research. While this session was slightly biased towards music corpora using a symbolic representation format given the session participants, the outcome of the session are relevant to all curators of research music corpora (including audio and score corpora). After a brief presentation by Frans Wiering about the corpus of Dutch Folk Songs which is currently under investigation by the Witchcraft project, the participants gathered together the aspects of music corpora that are important for making research with these large coherent collections of music pieces more effective.

As a result the participants agreed on a small number of pieces of information that, taken together, form a short description of a music corpus for research and that could inform interested researchers very quickly about the content and extent of a corpus, its availability and how to obtain it. This information led to the development of a simple and short standardised description scheme using XML mark-ups from which an HTML page can be produced very easily. The goal of this scheme is to give researchers a standardised template to publish the most important information about their corpus very quickly and easily.

The description scheme is called *DagMuCorR: Dagstuhl Description Schema of Music Corpora for Research* and comprises for files. With the help of these four files a researcher hosting and maintaining a music corpus should be able to produce a standardised HTML page that informs other researchers about the given corpus. The four files (which will soon be made available for download at http://www.doc.gold.ac.uk/isms/research/resources/dagmucorr/) are:

- **dagmucorr.dtd**: XML Document Type Definition that defines the basic elements of the description schema.
- **goldsmiths_corpora.xml**: An XML example file that contains descriptions of three music corpora maintained at Goldsmiths College. An extract from this is printed below as an example.
- **corpora_list2html.xsl**: An XSL style sheet that can be used to convert an XML file containing corpus data to an HTML page.
- **goldsmiths_corpora.html**: An example HTML page that results from applying the XSL style sheet to the content of the XML file.

There are three increasingly time intensive ways for a researcher to describe her/his own corpus within this scheme and to publish this information on an HTML page:

1. **Edit HTML page**: Replace all the information in *goldsmiths_corpora.html* with information on your own corpus and publish the HTML file. Takes about 20 minutes.
2. **Edit XML file**: To have your corpus description in a more flexible form (e.g. to derive a database scheme from it) replace all the information in *goldsmiths_corpora.xml* with information on your own corpus. Then render the XML file to HTML using the XSL style sheet. (One easy option is to place XML and XSL files in the same directory and open the XML file with a recent version of Mozilla Firefox. Then save the resulting HTML file as complete webpage.) Publish the HTML file. Takes about 30 minutes.
3. **Edit XML and XSL file**: If you want a distinctive layout of you corpus information page edit the XSL file as well and choice colours and layout options. Then edit XML file and apply XSL style sheet to XML file (as above) to generate the HTML file. Takes 40 minutes or more depending on the sophistication of your design.

The goal of having a description scheme and a common (and unique) name for it (*dagmucorr*) is that comparable information about research music corpora can be found very easily on the web.

The following example is a description of one of the corpora hosted at Goldsmiths and is taken from `Goldsmiths_corpora.xml`[8]

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl"  href="corpora_list2html.xsl
   " ?>
<!DOCTYPE corpora-list SYSTEM "dagmucorr.dtd">
<corpora-list xmlns:dsmc="http://www.doc.gold.ac.uk/isms/
   research/resources/dagmucorr" lang="en">
<title>
    <long-title>Goldsmiths Music Corpora</long-title>
  </title>
  <corpus>
    <title>
      <long-title>The Goldsmiths-Geerdes corpus of MIDI-
          transcriptions of popular music</long-title>
      <short-title>GoldPop corpus</short-title>
    </title>
    <version>
      <version-number>1.0</version-number>
      <last-update>August 2006</last-update>
    </version>
    <ownership>
      <free-description>ISMS Group, Computing Department,
          Goldsmiths College</free-description>
```

[8] In this extract, lines are broken for ease of reading.

```
  <institution>Goldsmiths College</institution>
</ownership>
<data>
  <free-description>14,063 pretty faithful MIDI
      transcriptions of western commercial pop songs from
       about 1950 to 2006</free-description>
</data>
<creation-period>
  <time-period>
    <start>1990</start>
    <end>2006</end>
  </time-period>
</creation-period>
<rights>Goldsmiths has the right to use the corpus and
    parts of it only for non-commercial research and on
    in-house College computers</rights>
<availability>Available for research only within the ISMS
     Group</availability>
<analysis-data>
  <free-description/>
  <analysis-period>
    <time-period>
      <start>2006</start>
      <end>on-going</end>
    </time-period>
  </analysis-period>
  <person>
    <name>
      <forenames>Daniel</forenames>
      <surname>Müllensiefen</surname>
      </name>
    <institution>Goldsmiths College</institution>
  </person>
  <person>
    <name>
      <forenames>Jamie</forenames>
      <surname>Forth</surname>
    </name>
    <institution>Goldsmiths College</institution>
  </person>
  <how>Frequency analysis of melodic formulae</how>
  <how>Analysis of repeating drum patterns (SIA)</how>
  <publication>@incollection{MullensiefenWigginsLewis08,
      Address = {Frankfurt},
      Author = {Müllensiefen, Daniel AND Wiggins, Geraint
           AND Lewis, David},
      Booktitle = {Systematic and Comparative Musicology:
           Concepts, Methods, Findings},
      Date-Added = {2008-08-19 15:17:36 +0100},
      Date-Modified = {2009-01-09 17:05:12 +0000},
```

```
        Editor = {Schneider, Albrecht},
        Pages = {133-155},
        Publisher = {Peter Lang},
        Series = {Hamburger Jahrbuch für Musikwissenschaft
            },
        Title = {High-level feature descriptors and corpus-
            based musicology: Techniques for modelling
            music cognition},
        Volume = {24},
        Year = {2008}}
      </publication>
      <publication>@incollection{FW2009,
         Address = {London, UK},
        Author = {Jamie Forth and Geraint A. Wiggins},
        Booktitle = {London Algorithmics 2009: Theory and
            Practice},
        Editor = {Joseph Chan and Jackie Daykin and M.
            Sohel Rahman},
        Publisher = {College Publications},
        Series = {Texts in Algorithmics},
        Title = {Approaches for identifying salient
            repetition in
                  multidimensional representations of
                      polyphonic music},
        Year = {forthcoming}}
      </publication>
    </analysis-data>
    <documentation>
      <docs>none currently available; see publications</docs>
    </documentation>
    <contact>
      <person>
        <name>
          <forenames>Daniel</forenames>
          <surname>Müllensiefen</surname>
        </name>
        <institution>Department of Computing, Goldsmiths
            College</institution>
        <email>d.mullensiefen'at'gold.ac.uk</email>
      </person>
    </contact>
  </corpus>
</corpora-list>
```

Chapter 14

Demos and 'Lightning Talks'

33 Unmeasured Music and the Computer

Tim Crawford (Goldsmiths College—London, GB)

This is a kind of 'preview' of some issues that arise when contemplating a small research project involving computer-aided investigation of the little-known, and less-understood, repertory of 'unmeasured' preludes for harpsichord from 17th-century France and for lute from 17th- and 18th-century Germany. The research will be carried out jointly with Daniel Tidhar (Queen Mary, University of London) with the kind help of Bruno Gingras (Goldsmiths, University of London). Some preliminary 'results' are destined to appear in a session (organised by TC) on 'Unmeasured Music' at the International Musicological Society Conference at Amsterdam in July 2009.

For this seminar, the point I'd like to make most strongly is that these pieces (around 100 survive for harpsichord, and at least three times as many for lute), though at first sight they are notated in the system typical for their instrument (standard 17th-century keyboard notation for the harpsichord, tablature in the case of the lute music), in fact only use a 'subset' of the system, and in a manner that can only be described as 'pragmatic', rather than following the 'rules' of that system. In both cases, performers who have taken the trouble to immerse themselves in the repertory will tell you that this compromised notation is in fact ideal to convey what is somehow essential for a performance of the music in a manner that suits their instrument better than any transcription which attempts to impose 'order' by adding barlines and regularising durational discrepancies.

I hope to discuss not only some of the problems of music representation that arise with preludes for both instruments, but also whether these 'problems' are best thought of as 'opportunities' to free ourselves as music-researchers from the well-known 'tyranny of the barline' (well-known to performers, that is). If we can in a sense think of all music as, in the first place, 'unmeasured', maybe we shall be in a better position to deal with some of the problems of reconciliation between scores and 'real' music as manifest in performance.

My example piece is perhaps not very characteristic of the genre as a whole, though I intend to use it to demonstrate a few typical features. The Fantasia in C major by the lutenist Silvius Leopold Weiss (1687-1750), who was an exact contemporary and acquaintance of J.S. Bach, is numbered 24* in the complete edition of Weiss's works I am editing for Bärenreiter, and is closely related to a shorter Prelude (21*) which is to be found a page or two earlier in the same manuscript.

Keywords: Unmeasured music, prelude, harpsichord, lute

34 Identifying and reasoning about repetition in symbolically-represented polyphonic music

Jamie Forth (Goldsmiths College—London, GB)

Repetition, either exact or based on some notion of similarity, plays a crucial role in many types of musical activity. I will briefly discuss recent work based on the SIATEC pattern discovery algorithm (Meredith et al 2002), in which we build upon previously published methods for identifying 'interesting' patterns from symbolically represented polyphonic music. The larger context of this research is computational creativity. We are initially concentrating our efforts on discovering salient musical repetition for the purpose of automatically generating musical knowledge over which artificial musical agents may operate. Currently we have applied our analysis techniques to Bach Two-part Inventions and MIDI encoded pop song rhythm tracks. The future of this work will concentrate on the representation of hierarchical relationships between instances of repetition. For the purpose of the seminar, I would be very interested to discuss ideas for evaluating and improving the performance of the algorithm, which perhaps might include comparing, on a large scale, automatically generated analyses with those of musicologists, or possibilities for using more intelligent and adaptive heuristics.

Keywords: Repetition, pattern discovery, symbolic music analysis, multilevel representation, musical creativity

See also: D. Meredith, K. Lemström and G. A. Wiggins (2002). 'Algorithms for discovering repeated patterns in multidimensional representations of polyphonic music'. Journal of New Music Research, 31(4): 321–345.

35 Towards Bridging the Gap between Sheet Music and Audio

Christian Fremerey (Universität Bonn, DE)

Sheet music and audio recordings represent and describe music on different semantic levels. Sheet music describes abstract high-level parameters such as notes, keys, measures, or repeats in a visual form.

Because of its explicitness and compactness, most musicologists discuss and analyze the meaning of music on the basis of sheet music. On the contrary, most people enjoy music by listening to audio recordings, which represent music in an acoustic form. In particular, the nuances and subtleties of musical performances, which are generally not written down in the score, make the music come alive. In this paper, we address the problem of bridging the gap between the sheet music domain and the audio domain. In particular, we discuss aspects on music representations, music synchronization, and optical music recognition, while indicating various strategies and open research problems.

Keywords: Audio, sheet music, symbolic score, optical music recognition, music synchronization

Joint work of: Fremerey, Christian; Müller, Meinard; Clausen, Michael

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1965

36 Owning another's music: edition, arrangement and musical influence

David Lewis (Goldsmiths College—London, GB)

What do we do to another's music when we write it edit it? What does the music do to us in the process?

Whilst i do not attempt to answer either of these questions comprehensively, I do discuss some of the issues involved and talk about where computers (and good music representations) may help. This talk is illustrated with examples of composers' keyboard transcriptions of the compositions of others.

37 An Inscriptive Metaphor for Music Encoding

Richard Lewis (Goldsmiths College—London, GB)

The practice I address is the encoding of musical information in computers. I consider such encoding as an instance of inscription—making marks on a medium. Drawing on the work of Bruno Latour (a leading figure in the field of science studies), I consider some important and interesting properties of inscriptions such as their mobility, immutability, scalability, and recombinability. I assess both the notated music score and digital storage as classes of inscriptive media and describe how they embody Latour's properties of inscriptions. I consider what role inscriptions may play in defining or demarcating a discipline (consider mathematics or chemistry as the manipulation of symbols and formulae or music analysis as the manipulation of notation). I also draw on the work of the philosopher Nelson Goodman whose theory of notation provides an interesting instantiation of Latour's ideas with explicit reference to applications in the arts.

My work-in-progress questions rely on drawing a distinction between encoding and representation. Representations (following Wiggins and Brachman and Levesque) are ways of encoding information coupled with ways of understanding it. Encodings (in the digital domain) are just the ways of turning information into bits and bytes. There seems to be a tendency to consider encoding as, not necessarily a trivial problem, but certainly as subordinate to representation.

On the contrary, encoding is actually quite a difficult and important issue. For example, in digital biblical hermeneutics, the encoding of Hebrew text is very difficult as its conception of characters isn't identical to that of Latinate text. Problems include not only things like right-to-left ordering, but also the encoding of vowels as alterations to other characters. Ordinary encodings used in computers (e.g. ASCII) are very much Latinate in nature (e.g. assuming a one-to-one mapping between characters and letters) and so encoding non-Latin characters is made quite complicated.

There may well be a whole layer of infrastructure that deals with encoding problems and which is related but different to solutions for representation problems. What we need is an abstract model for encoding (rather like Wiggins's abstract model for representing). Such a model could take Goodman's notions of marks, inscriptions and characters as a basis and build encoding semantics (not the same as representational semantics) on top of them. It could then be linked to representation systems by providing a means for establishing relationships between abstract representational concepts and concrete encoding elements. Such an infrastructure may go some way towards providing general purpose input methods for any notational phenomenon which, in turn, may provide the basis for digital editing tools and general purpose virtual research environments which rely on dealing interactively with notational data.

Keywords: Inscription, music encoding, editing, notation

38 Folk Song Alignment

Peter Van Kranenburg (Utrecht University, NL)

To compare two (folk song) melodies, it is insightful to notate the one below the other such that the corresponding parts are aligned. Therefore, alignments of two or more songs are heavily used by folk song researchers. It would be valuable for folk song research to be able to generate these alignments automatically. In Computer Science, sequence alignment algorithms were developed some decades ago. Because alignments have found an application field in molecular biology, where they are used to find corresponding patterns in protein or nucleotide sequences, it is in that discipline that many algorithms, improvements and optimizations were developed, which have the potential to be employed for musicological research as well. For each concrete alignment a score can be computed. The higher the score, the better the sequences can be aligned. In the computation of the score, domain knowledge should be incorporated. In the case of folk song alignment, musical scoring schemes should be developed. As much musical knowledge as possible should be incorporated in these scoring schemes, especially knowledge or hypotheses about melodic transformations in the process of oral transmission.

In order to use an alignment algorithm to align musical sequences, the following problems have to be solved:

- Representation: How to convert a melody into a sequence of symbols? What do these symbols mean?
- Substitution scores: What is the score of substituting a certain symbol with another?
- Gap penalties: What is the gap penalty function?

Keywords: Melody, Alignment, Folk Songs

39 Towards a framework for the analysis of performance issues

Dougie McGilvray (Glasgow Caledonian University, GB)

Extracting attributes of a performance is merely an act of observation. To analyse performance requires analysing performance attributes in the context of the score. For this reason, Performance Markup Language was developed to facilitate the analysis of performance issues. PML comprises multiple domains including logical score, logical performance, audio and gesture. The PML representation and the framework which surrounds it will be described. The currently limiting factor to the creation of a framework and tools which non-engineers would use is the lack of adequate interfaces to analyse information in the context of the score. Preliminary work to create scores annotated with analytical data and an interactive interface for analysing performance using the musical score will be described.

Keywords: Representation, Performance Analysis, Alignment, Visualisation

40 Ambiguity and Multiplicity in Music Representation

Alan Marsden (Lancaster University, GB)

Proper representation of music must often be ambiguous and/or multiple.

I give examples of this and propose some high-level strategies.

Keywords: Music Representation

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/1969

41 Good Knowledge Representation for Intelligent Music Processing

Alan Marsden (Lancaster University, GB)

I propose five tenets of good representation/modelling in this field: 1. Acknowledge what is left out. 2. Know about others' work. 3. Do not claim you stand on firm ground. 4. Do not trespass outside your domain. 5. Do not pretend you are not involved.

Keywords: Methodology

42 Statistical modelling of musical context

Christophe Rhodes (University of London, GB)

Statistical models have gained currency in aspects of research into musical data and musical representation, from issues in music perception, through analysis and synthesis, to aspects of musical performance. This paper takes this observation as its starting point, and argues for the construction of explicit statistical models to represent knowledge about the music under consideration, for two reasons: firstly, the use of context may allow the construction of models which perform better at given tasks; secondly, the existence of such a representation of knowledge allows the 'knowledge' itself to be tested quantitatively with reference to existing musical data. These ideas are explored with reference to tasks involving inference of unnotated information in particular repertoires.

Keywords: Bayesian inference; statistical modelling; quantitative musicology

Joint work of: Rhodes, Christophe

43 Musical Data Translation at the Center for Computer Assisted Research in the Humanities

Eleanor Selfridge-Field (Stanford University, US)

This short presentation describes organizational principles of data translations from the MuseData and KernScores full-score databases of musical repertories.

Keywords: Musical data translation; Humdrum; MuseData; musical scores and analyses online

Joint work of: Selfridge-Field, Eleanor; Sapp, Craig

Full Paper:
http://musedata.ccarh.org

Full Paper:
http://kern.ccarh.org

See also: Beyond MIDI: The Handbook of Musical Codes (The MIT Press, 1997).

44 The TEI Special Interest Group in Music

Raffaele Viglianti (King's College, CCH—London, GB)

The Text Encoding Initiative (TEI), as defined on its website www.tei-c.org, "is a consortium which collectively develops and maintains a standard for the representation of texts in digital form". The XML encoding model proposed by the group is extensively used by scholarly research in Text Encoding and its applications in Digital Philology are renown. For example projects developed by the Centre for Computing in the Humanities in collaboration with other departments use TEI as representation of a text with its edition (structural and semantic knowledge) and as a base for developing an interactive publication on the web (among them: Inscriptions of Aphrodisias, Henry III Fine Rolls).

A similar XML-centred approach to Digital Philology in music is present and growing, but definitely in need of a music notation format strong enough to be shared and that would meet scholarly needs. So far, among other XML representations for music, MEI has embraced Digital Philology the most.

The Special Interest Group (SIG) in Music[9] is a group of users of TEI that are interested in working towards a wider representation of "text" that would include music notation.

The SIG rejects the idea of expanding the TEI encoding with new elements designed for music representation and instead promotes a plug-and-play system employed by other SIGs too. Therefore the main purpose of the SIG is to provide recommendations and tools for extending TEI with music representation formats, and, when possible, vice versa.

The (small) group has started working recently and met only once (TEI Members Meeting November 2008), therefore it is now in a delicate phase of decision making and testing. Our first focus will be on MEI for the reasons mentioned above; though it is the SIG's aim to consider other relevant encoding formats later on and all the output produced will take in account equivalences with ISO standards and make them clear with documentation.

Major works in progress consist in comparing TEI and MEI modules and define when there is overlapping in the semantics and decide how or whether to merge the two representations.

The entire TEI is defined with ODD (On Document Does it all), a TEI-based format for writing human-readable descriptions of XML files. Through ODD TEI is structured in modules and classes. Therefore there isn't a fixed TEI Schema: a customized Schema in the desired format can be generated via a tool (Roma) that interprets ODD instructions.

[9] http://wiki.tei-c.org/index.php/SIG:Music

The effort that the MEI community is putting in redefining MEI with a literate schema design will make this task easier and would help getting more complete results.

The group hopes to learn from the success of TEI in Text Encoding and hopes to meet the needs of musicologists and text encoders that work with text where music notation is an important component often overlooked.

Keywords: TEI MEI ODD Philology Encoding Text Score Symbolic

45 Towards meaningful music (information) retrieval

Frans Wiering (Utrecht University, NL)

Great progress made in music information retrieval (MIR) research during the last decade or so. On closer inspection, this progress has been caused mainly by technological development: more often than not, only elementary knowledge of music has been incorporated in the design of MIR methods. There is an obvious lack of connection to human musical behaviour; consequently, only few research results have found their way into successful software products or Internet services.

The solution to this problem is to found MIR research first of all on music itself, viewed as a cognitive behaviour, and specifically, on the meaningfulness of music. The principal reason for people to compose, listen to, study, play and buy music is that it is meaningful to them. Yet such meaning is subjective, difficult to express, and hard to relate to measurable musical properties. To involve musical meaning in MIR may therefore not sound realistic at the moment or even be feasible at all. However, musical meaning is an excellent long-term goal from which concrete projects with more tangible aims can be derived.

One such project, currently under development with external partners Henkjan Honing (University of Amsterdam) and Louis Grijp (Meertens Institute) involves a collective online annotation system for musical audio. The annotation data will be used to determine the most important events in a piece of music, how listeners experience these and what measurable properties they have. For these properties similarity measures will be created, which in turn will be employed to improve access to recording archives.

Keywords: Music information retrieval, cognition, musical meaning, annotation

Part III

Concluding Remarks

46 “So why do we bother?”: The goals of computational musicology and modelling music cognition

David Meredith (Aalborg University, Denmark)

Towards the end of the 2009 Dagstuhl seminar on music representation, after several days of intense deliberation over the details of file formats, algorithms and representational schemes, I felt that the real reasons *why* we were doing what we were doing had been pushed far into the background. I therefore organised a workshop on 29 January with the aim of identifying the ultimate *goals*, *problems* and *questions* that motivated the work of the participants.

I suggested that we might be able to elicit these ultimate goals by first asking ourselves *what* we wanted to achieve and then, like a curious child, repeatedly asking ourselves *why* we wanted to achieve these goals until we discovered some high-level problems that were worth solving for their own sake.

The participants split up into groups of 5–8 people, and each group worked for around half an hour on accumulating a list of high-level, motivating goals and questions. It was hoped that this exercise could serve both as an inspiration for new projects and as a reminder to us all of why we do what we do.

The following is a list of all the goals, questions and problems that were suggested, with similar ones clustered into provisional categories. The ordering of the categories themselves is arbitrary.

Category 1: Improving our understanding of musical cultures or repertoires and the relationships between them

- Finding cultural references in order to understand one’s culture, understanding the cultural and musical transformation of melodies
- Creating a database of folksongs to study high-level relations between musical cultures
- Identifying quotations, allusions and references in polyphonic repertoires (in order to create musical meaning)
- Developing an “intertextuality machine” for finding textual relations to other pieces from a given piece (in order to create musical meaning)
- Integrating historical and contextual data with musical datasets in order to improve our understanding of cultural development and enrich the evidence base for musicological discourse
- Understanding folk-song transmission
- Studying scribal habits
- Increasing our knowledge of the whole field of musical culture and history, not just the main, “lighthouse”, historical figures

Category 2: Preserving and conserving knowledge

- Conservation and preservation

- Archiving results to facilitate access and growth of knowledge
- Providing musical texts for practical performance
- Creating digital, open, multi-accessible, multi-perspective editions to create new historical knowledge

Category 3: Understanding music cognition, perception and sociology

- Answering the question "What is man?" (Kant)
- Modelling human intelligence
- Understanding music cognition - more specifically, understanding musical memory and how melodies are encoded in the mind/brain
- Understanding the power of music to affect people politically and emotionally
- Building corpora of music that have specific cultural purposes (e.g. political music, advertising music, film music)

Category 4: Tools for music analysis and retrieving musical information

- Developing tools for counting the frequency of "stuff" (motives, harmonic progressions, etc.) in pieces of music in order to test musicological claims that particular types of "stuff" occur more or less frequently in the music of one composer or repertoire than another
- Building tools that enable others to manipulate, understand and enjoy music
- Browsing and searching large collections
- Developing a specific music-search function integrated with notation programs to familiarise musicologists with technology (thus creating a need for improved search technology)
- Developing a corpus-based categorization and classification tool for finding exemplars of particular types in a large body of music
- Developing a mood recognition and music selection tool that identifies the mood of a listener and plays appropriate music

Category 5: Understanding musical performance and composition

- Investigating artistic individuality in performance and composition
- Studying and analysing performance practice in order to benefit future performances

Category 6: Tools for performance and composition

- Building an automatic editor that would allow an amateur musician to go directly from a source to a performing edition
- Building a sound-to-notation renderer that allows manipulation of musical ideas that originate in sound

– Building a “music-minus-1” system that takes printed scores as input, allowing, for example, a soloist to generate a performance of a concerto without the solo part that he or she could then practise or even perform with (i.e., “score-based karaoke”)
– Developing an intelligent interactive composer’s sketchbook that allows for structural relationships to be visualized, which, in turn, could stimulate further creativity

Category 7: Music education

– Developing a system for marking harmony and counterpoint exercises that could be extended into a tutoring system

As can be seen, the goals identified ranged from the purely philosophical (e.g., “What is man?”, “What is the nature of human intelligence?”) to the highly practical (tools for practising, performing and learning); and from the purely scholarly (e.g., understanding folk song transmission) to the purely creative (e.g., an intelligent, interactive composer’s sketchbook).

Attendance at this seminar was conditional upon being invited by one of the organisers. So the range of research interests among the participants is perhaps more an extrapolation from those of the organizers than the agreed goals of any well-defined field. Nevertheless, despite the breadth of topics represented, a couple of unifying characteristics could be identified apart from the fact that all the participants use computers intensively to do research on music.

1. The participants were primarily motivated by musicological (including analytical and historical), cognitive or creative goals.
2. The participants were primarily engaging with music information on a symbolic level (as opposed to, for example, the audio level).

Despite the diversity of their interests, the participants therefore formed a surprisingly coherent group representing a “subfield” of music informatics that overlaps well-recognized areas such as music perception, historical musicology, music analysis, music information retrieval, composition and performance. The seminar thus provided a valuable opportunity for interaction between researchers with overlapping interests from different, well-recognized domains.

This suggests that, apart from providing inspiration for future projects, the list of goals given above might help to establish the identity of a research domain at the intersection of music cognition, musicology and music informatics.

47 (Some very personal) Critical Reflections on the Seminar “Knowledge Representation for intelligent Music Processing”

Joachim Veit (Detmold/Paderborn)

When Frans asked me to present you some very personal thoughts about my personal perception of the Dagstuhl Seminar I was very astonished, because besides my heavy interest in digital editions I am a straightforward narrow-minded classical editor (or editor of classical music) with an increasing number of pecularities, so I am a sort of outsider here in this room. And when I frivolously responded to Frans' idea I didn't know what difficult task he laid on my shoulders and why the devil he chose me. But in the meantime I suppose I got the point: Critical editors are those persons used to compare sources and evaluate them and give a Critical Commentary. But: these critical editors have real handwritten or printed sources—and I have none! So I can't really do the job I wanted and extract the good sounds from my sources while leaving the bad ones for the Critical Apparatus—and that, in traditional editions, means: make them vanish.

Thus I have to keep to my very transitory impressions during these exciting days and try to do a transitory electronic edition of those impressions. And Frans has the disadvantage that in normal digital editions the Critical Apparatus is part of the text and the user is confronted with both: the cleaned text with all its fine thirds and sixths and the damned critical stuff with the fourth and seventh—I promise to exclude any *diabolus in musica*.

So please allow me to present you a few remarks concerning the following topics:

```
<dagstuhlSeminar id="09051" n="1">
  <date from="2008-01-27" to="2008-01-30"/>
  <dagstuhlHeader>
    <fileDesc>1. Some Background to the Seminar</fileDesc>
    <sourceDesc>2. The Structure of the Seminar: Advantages and
         Problems of this Kind of Seminar</sourceDesc>
  </dagstuhlHeader>
  <text>
   <body type="Joachim"><!--fill in: 3. Some Remarks concerning the
        Content-->
   </body>
   <back>4. Consequences</back>
  </text>
</dagstuhlSeminar>
```

1. Some Background to the Seminar:

As far as musicologists and especially editors are concerned there have been only three conferences about those digital matters in Germany: Mainz 2007—Paderborn 2008—Dagstuhl 2009 (and some of you participated in all three):

- The Mainz Conference: *Digital Media and Music Edition* was a sort of getting into touch with one another for the first time (in this case it was also an interdisciplinary meeting of people from literature and musicology).

- The Paderborn Conference *Digital Edition between Experiment and Standardisation* concentrated on problems of encoding standards for music and again for special fields of text (especially letters) on the one side and tools for digital music editions on the other one.

At both conferences we had prepared lectures and nevertheless some room for panel and other discussions. Besides the lively discussions with about 80 or 90 people involved there was much unofficial talk and planning of cooperations in between—but you had to do that either in the breaks on the corridors or in the beer tavern.

One of the most important results of both conferences was exactly the kind of thing we do here: we fully realized the importance of direct and long running contacts between the projects or even the collaboration of projects and the necessity to the facilitate the interchange of results as well as a better infrastructure for mutual information.

And now we have this fantastic one-week *Dagstuhl-Seminar* with a smaller number of people (which makes working together easier). Besides our five-minutes presentations we came with no prepared papers but only with the intention to spend the whole week with colleagues discussing problems of Knowledge Representation in a beautiful surroundings and in a—despite the freezing temperatures outside—very comfortable atmosphere which might have to do with the philosophy of this house with its open rooms, people changing places at lunch and dinner, having a hike and even playing and singing together. And we had room for endless discussions on the time-line as well as on the locality-branch. And even if somebody wanted to retreat for a while (as I had to do) there is no problem—so I felt like having a one-week sabbatical with a lot of nice people around me, with good talks and good and very inexpensive wine and beer—but only with very short nights. Let me please emphasize this atmosphere of warmth with which people met here and the unusual possibility to have such an intense contact with many of our colleagues. That might be only a psychological aspect but I think this promotes and facilitates future collaborations enormously. You cannot measure that in Euros and Cents but without this atmosphere there is no inspiring research I think!

So I end the first point accompaning my sweet melody only with parallel thirds and sixths.

2. The structure of the Seminar

I admit that I was very sceptical in advance and before leaving my urgent work for the forthcoming volume of our edition I was not sure if it would not be better to keep to my daily work. Even the idea of nearly 40 of these *5-Minute Presentations* seemed strange to me at the beginning. But I have to admit that I changed my mind: Although my memory for persons (and especially their names) is quite horrible, I was very astonished about the broad field of interests associated with “intelligent music processing” and I am happy to have had this kind of survey which we miss so hardly because we often know only a small sector of the field of "digital research" (please allow me this perhaps misleading expression).

The process of finding overlaps between these projects or at least overlapping interests went pretty quickly for such a big and perhaps inhomogeneous group. Based on this search we had our *Workshops* which seem to have covered shared interests in all groups—perhaps with the exception of the workshop on computational musical analysis on Tuesday which David [Meredith] found too disparate. Changing topics of workshops combined with changing participants in these workshops may cause some delay in getting quickly to effective discussions. I think we didn't had that problem in our multipart workshop on editorial and encoding matters—but instead we run into problems I shall mention in part 3.

We all know about the problem of parallel lectures or workshops from many conferences. In our case this working in parallel groups didn't allow many of us to inform ourselves—perhaps for the first time—about things going on in neighbouring fields of digital research. e.g. I was very happy to be able to go to the workshop on data-structures yesterday morning which gave me a lot of stimuli for our discusions in the edition-group the day *before*. Certainly we had the short summaries of the workshops so that we at least had a weak impression of what had been discussed in the other workshops—but for myself I should have liked more detailed reports, perhaps complemented by other participants.

The *Stimulus Talks* have been a very good opportunity to fill in the gaps of knowledge and for a future seminar I should like to have at least one in the morning and one in the afternoon and perhaps alternating between presentations of problems within certain narrowly defined research areas and more general discussions of methodogical or other problems.

The time for the *Demo Session* which wasn't really at the centre of our endeavours has been too short for me: I didn't manage to see more than two and a half demos—which I really regret because in this case we had the possibility to consult the developers directly.

The *"Unforeseen Sessions"* (i.e. sessions planned as a splitting up of a previous workshop or as a continuing of former activities in the same or smaller circles) proved to be one of the greatest advantages within this form of seminar. We, for example, split up our editorial group in several sessions on encoding with three to 10 or 12 people. And I think that there have been more of these "unforseen" sessions and they contribute enormously to the success of this Seminar. Dagstuhl offers ideal support for such a kind of thing because there are so many rooms in any size you want.

So, concerning the structure of the seminar, I in the end liked this open form which allows you to spend as much time for solving urgent problems as you have in one week. But for future seminars with scholars form such a wide variety of research I should like to recommend you: Please make sure that you—to a greater amount—force people of different backgrounds to work at the solution of shared problems. The confrontation of the interest of developers of tools or people with a more mathematical background with those strange folk from the humanities or musicology might lead to really controversial discussions and a better understanding for each others positions—which is necessary in order to come to convincing solutions instead of producing high-tech toys or clumsy do-it-yourself constructions.

So to come to an end with my second point: Despite these suggestions for future seminars—it has been an exciting experience taking part in this form of seminar which was really new to me. And I learned from this: Do trust in self-regulating social processes if you have people who don't want to go home from such a seminar without finding compensation for loosing one week of regular work!

3. Some Remarks concerning the Content of the Seminar

Let me choose only two examples from our editorial and encoding workshops and just have a short look on the processes initiated during our discussions here:

The first example is Andreas' shifted idea of a *Dagstuhl-Core* encoding which he put on the table in his short introduction on Monday. Many of us believed that this Dublin-Core inspired concept might help us to find the common essentials of music notation and allow easy translations between codes. In the course of the discussions Andreas modified this model step-by-step and ended up in collaboration with some of our group in order to develop a new kind of documentation of necessary core features and their coverage by software tools. I hope he allows me to abridge his ideas in this way—but my point here is this *modifying process* which was only possible within such a short time because people really met and had very target-oriented activities. And the whole community will profit from the publication of the results on the wiki.

The second example is Perry's *MEI*: The idea of a modular encoding structure which allows us to encode documents in Common Western Music Notation and to add modules for Neumes and Mensural Notation or Tabulature is as much fascinating as the possibility to encode variants, versions and even genetic aspects of an edited work which I as an editor urgently need. Surely some other formats like *MuseData* or *Humdrum* allow you to include variants, but MEI will guarantee you a much deeper coverage of the details of manuscripts. So representation within MEI seems to be perfect but at the price that there are no tools at the moment for displaying or for further use of these data. Nearly endless and thorough discussions arose about the question whether we should have a core encoding in this case too, whether a representation without tools and software doesn't make sense and to what extent Mensural Notation or the manifold notational variety are challenging questions for the main principles of MEI's basic concepts. Again I am not interested in the details of these discussions here but in the *process* that was launched by the Dagstuhl surroundings: Several of our discussions really exhausted at the end—nobody seemed to have any suggestions to solve the problems. But at the same time nobody seemed to be content with this situation and everybody endeavoured to find a new starting-point for some compromise. I nearly had the impression that people thought that it would be a shame if this week ends without new prospects for these intentions "to encode written documents"—as Perry called it.

I think, visiting one of the other workshops would have helped us in this situation too: Geraint's distinction between *format* (as syntax) and *representation* (as meaning) or his request: "Do not represent music *per se* but information *about* music" might have helped us as well as Alan's warning: "To include everything would be to copy the universe!" or his question: "What is the *task* of your representation"? And in the data-structure session I was impressed how far encoding of simply pitch and time might lead us in very intelligent analytical processing procedures. The relatively simple encoding is absolutely efficient for them.

So should we take the back seat instead of bloating up our encodings? And should we always remember that editorial encoding is a representation of written documents and nothing else?

It is really great to see how many new activities these discussions about encoding created—I think somebody will report about that later. Thus, concerning the results of our Seminar, Eleanor, Frans and Geraint must be very pleased if such things as here happened in many of the other workshops and "unforseeen" sessions too. —That leads to my last and very short point:

4. Consequences

My transitory digital edition of the Dagstuhl meeting was programmed to early in the morning of this meeting's last day to draw any conclusions from such a preliminary and bug-full version. But I suppose that this Seminar opened the disk drive for many new CD's labeled: "cooperations" and "new projects", "new goals", "further Dagstuhl Seminars with 40 or with only 10 or 20 people". I hope that we shall have a still closer cooperation and more communication about our results—even perhaps a new platform for digital research.

I'll end here and even if it is not my task here to thank officially—I at least want to express my warmest thanks to the organisers for this inspiring meeting which will hopefully contribute to a broader perception of digital activities in German Musicology too and especially will motivate young scholars to climb on our slowly moving bandwagon in order to promote the future development of *Knowledge Representation for intelligent Music Processing.*

Index