

MUSIC INFORMATION RETRIEVAL, MEMORY AND CULTURE: SOME PHILOSOPHICAL REMARKS

Cynthia M. Grund

Institute of Philosophy, Education and the Study of Religions
Philosophy
University of Southern Denmark
Odense, Denmark
cmgrund@filos.sdu.dk

ABSTRACT

The burgeoning field of Music Information Retrieval (MIR) raises issues which are of interest within traditional areas of discussion in philosophy of music and of philosophy of culture in general. The purpose of this paper is twofold: the first goal is to highlight and briefly discuss a selection of these issues, while the second is to make a case for increased mutual awareness of each other on the parts of MIR and of humanistic research. Many traditional debates within the latter receive infusions of new perspectives from MIR, while research within MIR could be fruitfully pointed in directions suggested by questions of interest within traditional research in the humanities, e.g. the relationship of individual memory to cultural memory, issues regarding cross-cultural understanding and the importance of authenticity in artistic contexts.

Keywords: Philosophy of music, culture, memory, ethics, authenticity

1 INTRODUCTION

Earlier this semester (spring 2005) I included a section on music information retrieval in an invited lecture on music and meaning which was part of a semester-long series of lectures on Semiotics and Meaning at the Institute of Language and Communication at the University of Southern Denmark (SDU) at Odense, sponsored by SemioNet, The Network for Studies in Semiotics and Meaning (SDU), Odense. I am an Associate Professor of Philosophy at SDU in the Institute of Philosophy, Education and the Study of Religions. I provide this admittedly anecdotal preamble to underscore the solidly humanities-based context. As part of the illustrative material in the PowerPoint presentation, I used a segment of streaming video from BBC World's Click Online, August 1, 2002 (Eddo, 2002) (thus already nearly three years old) describing the tune-tagging mobile telephone service called *Shazam* that was then being launched in

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2005 Queen Mary, University of London

the UK. My spring 2005 audience – as has been the case with virtually every humanities-oriented audience I have encountered since I began addressing MIR-related issues in talks some three years ago – had never heard of MIR. I supplemented the video segment with an overview of the various kinds of work which had been pursued in MIR since I was introduced to the field at two cross-disciplinary conferences at SDU in 2002 sponsored by Netværk for Tværvideenskabelige Studier af Musik og Betydning/The Network for Cross-Disciplinary Studies of Music and Meaning (NTMSB).¹

During the question-and-answer session after the conclusion of the three-hour lecture, the question from the floor which sparked most discussion related to the BBC-clip on *Shazam*. The issues raised nearly all had to do with the consequences this and related technology (i.e. MIR) would have for our perception and use of human memory with regard to music and sonic phenomena, and the broader implications for human culture in general. These issues – and related ones – are the issues to which I will now turn.

2 THE ISSUES

The specific question which launched the discussion dealt with what the sort of music mining represented by *Shazam* would mean for us as individuals:

(*). Does the ability to access music information in this fashion somehow impoverish the quality of our connections with music and the connections music facilitates between an individual and her fellow human beings by – additionally – eroding our abilities to concentrate, get information from our environment into our memories and contextualize it, so *we* can remember it?

As will become evident in what follows, my considered position is that this is not the case, that MIR in its various aspects will only result in – on balance – an improvement with regard to music-remembering, both on an individual and on a cultural level. This question, moreover, provides an elegant way to introduce and comment on the following relations between MIR, on the one hand, and several issues of relevance to philosophy and philosophy of human culture on the other:

¹ I am particularly grateful to Esben Skovenborg for drawing my attention to *Shazam*. My thanks also to Marc Leman and Declan Thomas Murphy for fine introductory insights into MIR back in 2002-2003.

- The advent of intramedial external interactivity (defined in what follows) with regard to musical materials occurs with music mining and related areas of MIR.
- In the long run, MIR can contribute to the discussion of the role and philosophical significance of memory with regard to questions of memory and of metaphysics in the debate about the ontology of musical works. MIR may well also be able to contribute to the growing field of research which regards a better understanding of memory as being a key subtext for understanding human culture.
- Continued striving within MIR research to find out how people categorize music and thus will be led to search for it yields insights as to how various cultures hear their own music and provides a portal through which non-members of a culture also can be party to these insights.
- MIR can introduce new aspects into discussions of authenticity as a factor to be taken into account in philosophy of performance and reception.
- New ethical issues arise with regard to MIR and some traditional ones are put into new and challenging perspectives.

There are a multitude of other issues which also could be brought to the fore, but in the interest of focus and brevity, these will suffice for the time being.

2.1 Intramedial external interactivity

2.1.1 *The Distinction Between Internal Interactivity and External Interactivity*

Querying by humming, singing, playing and the like has, of course, gone on throughout the history of humankind. In order to determine the name of the tune one had running through ones head, it was necessary to be lucky enough to encounter another human being who happened to recognize the tune which was reproduced and, in addition, who knew its name or could identify it some way. For lack of a better term, I have chosen to call this *internal interaction*, signalling the fact that a tune running through the consciousness of one human agent was then reproduced by this agent in order to run it by another human agent, who, after listening to it, then could identify it. Key steps in this process involve the intermediation of conscious, intentional human behavior.

Contrast this now with what is going on in something like the *Shazam* example. Here, “you can pull out your mobile phone, which you already carry with you everywhere, and you can dial four digits, wait fifteen seconds, and immediately found [*sic*] out the name of the artist, the album and the song” (Chris Barton in Eddo, 2002). In the querying situation at hand no person has to remember the melody in order to reproduce it for the purposes of querying, and no person has to recognize it in order to identify it. Here, *external interaction* is an appropriate term, in that in the interaction involved in the querying – all except the action of initiating the

query itself – takes place outside of any human consciousness.²

There are, of course, examples of mixed internal/external interaction: standard query by humming in music mining involves the ability of a human agent to reproduce to some extent or other a tune, which then is identified by an algorithm.

This distinction becomes most interesting when paired with the following one:

2.1.2 *The Distinction Between Intermedial Reference and Intramedial Reference*

When referring to music, particularly in order to access it in order to perform, *intermediality* is what has characterized the referential relationships between played, sounding music and score, chord chart, or whatever. Granted, there has been limited *intramedial* reference on an internal mnemonic level when, for example a conductor of an orchestra hums the portion she wants to hear again in rehearsal or in the sort of pre-MIR query-by-humming traditionally practiced among human agents.

2.1.3 *The Upshot*

What I think is interesting to think about in the music-mining case as it is exemplified by *Shazam* is that here we are crossing over into the realm of *intramedial external interactivity*. Music is used to find information about music, without the need for any of the music to be in any way represented or internally “heard” in anyone’s consciousness.

On the face of it, this would seem to feed the fears which might be seen to be lurking behind the question labelled (*) in the foregoing. Some afterthought will reveal, however, that this potential for finding and identifying music when we don’t have access to other human beings who can help us will only contribute to enhanced possibilities for us as human agents to get hold of accurate information. We will thus be better enabled to find out what we want to know in an increasingly complex musical/sonic environment in which we are exposed to a heretofore unheard plurality of music and musical styles, so that we can resolve our own queries

² An anonymous reviewer has remarked: “ ‘External’ is typically used to refer to person-to-person interactions,” referring to work by Hülsmann as well as Moore and Kearsley. S/he writes: “. . . the author seems to use terminology in a manner inconsistent with the literature, since she uses the term “external interaction” to refer to a ‘learner-content’ or ‘learner-resource media’ relationship.” To avoid any further misunderstandings among readers regarding my use of the term “external interaction,” I would like to point out that it is my own invention in this context, and – in this example – it refers to the interaction taking place between the mobile phone as the mediating presenter of sonic information to the database of digitalized music which is then subjected to an algorithm-based search for matches. At no time in this scenario does any human agent need to reproduce, remember or in any other way relate consciously to the music, save to be spurred by it to hold up the mobile phone and to subsequently check the information yielded by the phone.

I would like to take this opportunity to thank all four anonymous peer-reviewers for suggestions regarding conceptual clarification and copy editing.

and find our own personal “tags.” Both the quantity and quality of what *we* will be able to remember musically is enhanced.

2.2 Memory, metaphysics and culture

2.2.1 *Memory and Metaphysics*

In his very recent *Philosophy of Music: An Introduction*, a book solidly based in traditional, humanistic treatments of the topic, and which can be recommended both to those who are familiar with the issues dealt with in the philosophy of music and those who are new to the field, R.A. Sharpe writes:

You cannot reduce a work either to the class of its performances or to the possibility of a performance. The work is not the same as a performance. Let’s begin with a few truisms. A work can exist unperformed, as long as the music has been written out and preserved in a library or a study. It might, as well, be remembered accurately by somebody even if no notated copy exists. While it is remembered, it exists. If the last copy is destroyed it continues to exist as long as the last person to remember it can do so. After that the work is lost. (Whether it is right to say that it no longer exists is something we shall consider shortly.) (Sharpe, 2004, p. 59)

The “something we shall consider shortly” is, of course whether or not a musical work might have some sort of eternal existence as Platonic entities. It would be quite astonishing if MIR – or anything else, for that matter, could end up actually providing us with tools for resolving metaphysical quandaries of this latter sort.

To the extent, however, that MIR can enlarge, refine or otherwise give us new twists on what is involved in remembering a piece of music, the foregoing quote about the relationship of existence of a piece of music in the sense of it being remembered reveals at least one point of relevant contact between MIR and philosophizing about music. Although this concern from the humanities side of the fence, by virtue of its very metaphysical nature, may not be the sort of thing that causes the vast majority of researchers within MIR to lose sleep at night, there are issues dealing with memory and music which have intrigued and perplexed those who philosophize and theorize about music which are well-suited to examination by means of the tools afforded within MIR. For example, Sharpe contrasts the processes of supplying missing information in a spoken conversation with completion of a musical phrase:

But with music it is the syntactic features alone that enable us to supply what is missing, for there is no semantic element. Different notes and chords do not approximate to nouns, adjectives, prepositions or pronouns. If we attempt a glossary of music, we will take either whole phrases and whole harmonic sequences or notes in relation to the tonic and give them a character of the sort I have described above. Thus we may think

of rising major sequence as optimistic or a chromatic descending passage as sad. But the syntactic element that enables music to be followed is crucial to the way we value it. Music that is merely a concatenation of agreeable sounds does not have the same appeal. The very memorability of music connects with this capacity to be followed (Sharpe, 2004, p. 94).

Since the ability of a human agent to remember at least some sonic, non-textual aspects of a piece of music is crucial in many search contexts – exceptions are, of course, those involving intramedial external interactivity, on the one hand, and those involving search in terms of title and the like, on the other – this is also an area of relevance to MIR. Since MIR is replete with tools which can compare huge assortments of sonic data with which reactions on the part of test subjects have been associated, it should only be a matter of time, creativity and experiment design before MIR can provide us with substantial insights into how we remember musical material and how memory of musical material is related to other sorts of memory.

2.2.2 *Memory and Culture*

MIR-based research into features which make a piece of music more or less memorable for an individual should, in turn, be able to shed light on the multifaceted use of music in a wide variety of cultural contexts. There is a growing amount of research now being done on the ways in which many elements of human culture are increasingly being understood as rooted in pre-literate contexts as part of our constant struggle to remember a past which is always receding and eluding us, both on an individual and on a collective, cultural level.³ The tools which MIR brings to the table with regard to isolating and analyzing musical features with regard to the matter of memorability in large corpora of sonic material within and across cultures can be utilized to provide concrete, empirically-supported insights into a large and otherwise unwieldy area of inquiry.

2.3 Cross-cultural insights

The study of how various groups and cultures categorize the music in which they are interested (or not!) becomes a cardinal point in the development of more sophisticated tools for MIR. An important side effect of this research is increasing refinement of our insights into how different groups approach their music.⁴

This is, of course, interesting in the traditional philosophical debate with regard to “pure” music, where the work of interpretation is not short-circuited by the presence of a text. After a fine discussion of how we talk about music, R.A. Sharpe – once again – writes:

³ For fascinating insights into the way in which the need to remember can provide an explanation for behavior for which it is otherwise very difficult to give an account, see Whitehouse, 2000.

⁴ Debopam, Pappu, and Prabhakar, 2004, is a fine example of this.

In attempting to show how expressive descriptions become attached to music I have made much of the fact that what we may call “concert music for instruments only” is comparatively recent and comparatively restricted to Western and Indian cultures. Music has been and still largely is a mixed-media affair. . . . For most of its history, music either set tests or was an accompaniment for dance and it may well be that purely instrumental music would not have the effect it has if we were not brought up in a culture where our first experience of music is likely to be through singing and being sung to and moving in time to music. Perhaps this is where we should look for the primitive basis for our expressive descriptions of music (Sharpe, 2004, p. 107).

The enormous databases of digitalized music, much of which consists of music with lyrics, are indeed, “mixed-media affairs.” The work currently being done by Stephan Baumann, Tim Polhe, Vembu Shankar, Beth Logan, Dan Yang and Wonsook Lee⁵ is of great interest to any philosopher interested in how traditional concerns within music-and-meaning studies can increasingly be placed within empirically-based contexts as MIR continues to develop.

2.4 Authenticity

Questions of authenticity are often lurking beneath the surface in matters of musical performance, be it with regard to interpretation of score, instrumentation or any number of related issues. As will be remarked in the next section, MIR raises some new, interesting and not always unproblematic issues in this regard. There are, however, respects in which MIR and related technologies, can give us access to realities which, by their very nature, only can be virtual. I will only name one example here, but it is very illustrative: The CAHRISMA Project. CAHRISMA is the acronym for Conservation of the Acoustical *Heritage* by the *Revival* and *Identification* of the *Sinan's Mosques Acoustics*. A brief description from the project homepage reads: “The CAHRISMA project is an European Commission Fifth Framework INCO – MED Programme. This programme is directed towards cooperation between EU-countries and the Mediterranean countries. . . . The main focus within the project is to innovate [*sic*] the concept of Hybrid Architectural Heritage being a new way of identification that covers *acoustic* as well as *visual* features, the idea being that for spaces having acoustic importance, the architectural heritage concept, considered in conservation and restoration projects should be upgraded to cover acoustical as well as visual aspects.”

The results are stunning (and can be heard on the website). Thanks to the simulated auralizations, the listener is given the opportunity to hear, to *experience* music from diverse historical periods as it sounded in the

⁵ See Baumann, Polhe, Shankar (2004), Logan (2004), Shankar & Baumann, (2004) and Yang & Lee (2004).

venues in which it was performed as they were constructed and situated at that time, where the effects of later renovation and the distractions of irrelevant ambient noise such as that of modern traffic are removed.

2.5 Some Ethical Issues

MIR, however, can also raise potentially troubling issues with regard to authenticity. Once again, the framing for these remarks is provided by Sharpe:

We owe it to the composer to play it [the music] in the way he conceived it and it normally sounds better that way. Admittedly, unlike moral considerations elsewhere, these vary with the stature of the music. . . . Great works were conceived to be played in a certain way. Undoubtedly, within these constraints there will be more than one way of doing it. Bach might not have worried as to whether his keyboard music was played on the harpsichord, clavichord or early piano. He might have had preferences, but any of these instruments would do. But I suspect that ornaments would have mattered as intensely to him as they did to Rameau, and we know Rameau cared, given his instructions on the matter (Sharpe, 2004, pp. 82-83).

One can imagine a point in the development of MIR and related computer modeling techniques, such as beat extraction which allows for the conducting of audio files⁶ at which rank amateurs can “reconduct” performances by world renowned orchestras which were originally under the baton of a master. This raises some interesting issues of authenticity and ethics which, though perhaps anticipated by the foregoing quote and similar considerations within ongoing debates about performance practice within traditional musicology, introduces a sufficient number of fresh considerations to warrant renewed discussion.

3 . . . AND WHERE DO WE GO FROM HERE?

I conclude with some remarks which I found myself making repeatedly in various unpublished fora before and during ISMIR2004 in Barcelona; this gives them a chance to go “on record,” and the remarks in section 2 are somewhat more amplified than anything that I had an opportunity to formulate in these fora.

It is important to stress, that although I am sort of a token philosopher/humanist with the context of ISMIR, I am not advocating any kind of “MIR for poets” sessions or topics. It is my experience that when humanists – who often feel unnecessarily intimidated by an IT-heavy field of research such as MIR – are finally lured into lectures and talks on MIR, they are able to understand what is going on and are often stunned by the implications they realize that MIR has for, say, aesthetics and epistemol-

⁶ Again thanks to Declan Thomas Murphy for insights into this area and his work within it.

ogy. Some of these have been outlined in the preceding section. Of course it demands some preparation on the part of a speaker to present highly technical topics to mixed audiences, but I have seen on numerous occasions in Denmark that there can be very productive cross-fertilization when the “hard” and “soft” sciences meet.⁷

While humanists may have to overcome initial feelings of intimidation when confronted by MIR, on the other hand, I have heard from some very technically-proficient MIR and computer music researchers who have given talks to mixed audiences in which humanists are well represented, that they have found the experience to be both fruitful and rewarding, in that they have felt somewhat isolated and alone in the technical environments in which they commonly work and in which their own speculations of, say, an aesthetic nature, are derided and not taken seriously.

To sum up: The goal of this paper has been to indicate how at least some of the advances in MIR may be contextualized within the framework of a selection of questions within philosophy of music and philosophy of culture. It is my hope that not much more time will pass before we see increased awareness of – and interest in – the tremendous multidisciplinary implications of MIR and its potential for providing a potent interface between science and the humanities.

REFERENCES

- Baumann, Stephan: Tim Polhe, Vembu Shankar. “Towards a Socio-Cultural Compatibility of MIR Systemes”, (2004) Proceedings of the 5th International Conference on Music Information Retrieval, Universitat Pompeu Fabra, Barcelona, Spain, October 10-14, 2004, 460-465.
- The CAHRISMA Project
<http://www.dat.dtu.dk/cahrisma.htm>
- Eddo, Mark. (2002) "Shazam!!", BBC Prime Click Online Interview August 1, 2002 with, among others, Chris Barton and Vijay Solanki of *Shazam*, http://bbcworld.com/content/clickonline_archive_31_2002.asp?pageid=666&co_pageid=4
- Hülsmann, T. (2000) *The Costs of Open Learning: A Handbook*. Oldenburg, Bibliotheks-und Informationssystem der Carl Ossietsky Universität Oldenburg.
- JMM: The Journal of Music and Meaning, www.musicandmeaning.net
- Logan, Beth. (2004) “Music Recommendation from Song Sets.” Proceedings of the 5th International Conference on Music Information Retrieval, Universitat Pompeu Fabra, Barcelona, Spain, October 10-14, 2004, 425-428.
- Moore, M. and Kearsley, G. (1996) *Distance Education: A Systems View*. Wadsworth Publishing Company, Belmont, CA.
- Murphy, D. T., H. Andersen, and K. Jensen. (2003) “Conducting audio files via computer vision.” Proceedings of the Gesture Workshop, Genova. Springer Verlag, (Available at <http://www.diku.dk/~declan/pub/papers.html>.)
- NTSMB: Netværk for Tværvideenskabelige Studier af Musik og Betydning/Network for Cross-Disciplinary Studies of Music and Meaning, www.ntmsb.dk, Institute of Philosophy, Education, and the Study of Religions, University of Southern Denmark, Odense Denmark. See specifically programs for *Nature, Culture and Musical Meaning* and *Music, Logic and Technology* both held at SDU in 2002.
- Roy, Debopam, Nagaraju Pappu, T.V. Prabhakar. (2004) “Parichaykrama – An Exploratory Interface of Indian Classical Music Using Experiential Framework”, Final Proceedings of Computer Music Modeling and Retrieval, International Symposium. CMMR 2004, Esbjerg, Denmark, May 2004, Revised Papers, 359-370.
- SemioNet, The Network for Studies in Semiotics and Meaning, www.semionet.sdu.dk, Institute of Language and Communication, University of Southern Denmark (SDU).
- Sharpe, R.A. (2004) *Philosophy of Music: An Introduction*. Acumen, Bucks.
- Vembu, Shankar, Stephan Baumann (2004) “A Self-Organizing Map Based Knowledge Discovery for Music Recommendation Systems”, Final Proceedings of Computer Music Modeling and Retrieval, International Symposium. CMMR 2004, Esbjerg, Denmark, May 2004, Revised Papers, 119-129.
- Whitehouse, Harvey. (2000) *Arguments and Icons: Divergent Modes of Religiosity*. Oxford University Press, Oxford.
- Yang, Dan; WonSook Lee. (2004) "Disambiguating Music Emotion Using Software Agents", Proceedings of the 5th International Conference on Music Information Retrieval, Universitat Pompeu Fabra, Barcelona, Spain, October 10-14, 2004, 52-27.

⁷ A forum for this sort of dialog is JMM: The Journal of Music and Meaning, www.musicandmeaning.net