

The Musical Metacreation Weekend: Challenges Arising from the Live Presentation of Musically Metacreative Systems

Oliver Bown
Design Lab,
University of Sydney, Australia

Arne Eigenfeldt
School for the Contemporary Arts,
Simon Fraser University, Canada

Aengus Martin
Faculty of Engineering,
University of Sydney, Australia

Benjamin Carey
Faculty of Arts and Social Sciences,
University of Technology Sydney, Australia

Philippe Pasquier
School of Interactive Arts and Technology,
Simon Fraser University, Canada

Abstract

This paper discusses the approach to curation of the MuMe Weekend (held in Sydney, Australia, 2013) and the experience the authors gained from hosting the event. We identify open challenges arising from this applied demonstration of musically metacreative systems. Identified challenges are not technical, but rather methodological, concerning pragmatic aspects of presenting and collaboratively innovating musically metacreative work. The paper presents our approach to curating the event, the categories of performances offered – and responses to these – issues to do with the selection and presentation of the work, and issues to do with the evaluation by audience and performers.

Introduction

The Musical Metacreation Weekend (MuMeWe2013) was held in Sydney, Australia, on June 15th and 16th 2013, hosted by two faculties of the University of Sydney: the Design Lab at the Faculty of Architecture, Design and Planning, and the Conservatorium of Music. The event was timed to coincide with the 19th International Symposium on Electronic Art (ISEA), of which it was an associated event, and also occurred alongside two conferences: the International Conference on Computational Creativity (ICCC), a AAAI conference now in its 4th year, and the ACM Creativity and Cognition Conference, now in its 20th year. MuMeWe2013 can also be contextualized as part of the MuMe workshop series, the first of which, MuMe2012, occurred in October 2012 at Stanford University as part of the AAAI International conference on AI in Interactive Digital Entertainment (AIIDE 2012).

The MuMeWe2013 program, which lists all work exhibited at the MuMe Weekend, along with miscellaneous documentation, can be found online¹.

This paper discusses the approach to curation of the MuMe Weekend and the experience the authors gained from the event. The main purpose of the paper is to identify open challenges arising from this applied demonstration of musically metacreative systems. Identified challenges here are not technical, but rather methodological, concerning pragmatic aspects of presenting and collaboratively innovating

musically metacreative work. The paper presents our approach to curating the event, the categories of performances offered – and responses to these – issues to do with the selection and presentation of the work, and issues to do with the evaluation by audience and performers. We list significant open questions as they arise in the discussion.

Organisation and context

The MuMe Weekend was proposed as a creative-output counterpart to the MuMe Workshop series, which was initiated in 2012 as part of the AIIDE conference (Pasquier, Eigenfeldt, and Bown, 2012). The 2013 MuMe Weekend focused on disseminating the creative musical outputs associated with the kind of research presented at the workshop, both to the small community of MuMe researchers, and to a wider audience consisting of attendees of the ICCC conference, the ISEA symposium and the general public. The event offered the opportunity for researchers to meet and present and discuss systems, and for MuMe work to engage a wider audience. Being a formative field (its academic history is long but irregular, straddling different banners and communities), this inaugural event was set up with the expectation of inviting future innovation in such gatherings through experimentation and reflection around issues such as communication, clarity, relevance and artistic engagement. A parallel paper proposing a taxonomy of musical metacreation was written in response to the weekend's presentation of metacreative systems (Eigenfeldt et al., 2013), which stands in contrast to this paper's focus on organisational aspects of research and presentation.

The MuMe event was free and publicly advertised both independently and via the ISEA program, for which it was an associated event. The emphasis was on informal presentation in nonacademic settings.

Approach to curation

The MuMe event attempted to strike an appropriate balance between academic rigour and creative quality. Materials were gathered through an internationally disseminated call for participation. A curatorial committee was set up to provide some level of peer review of the work, but only as input to a centralised curatorial process. Reviews and reviewers were not revealed to the contributors, but reviewers

Copyright © 2013, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹<http://www.metacreation.net/mumewe2013>

were given full access to the submission materials including names and biographies of the contributing artists.

A fully peer-reviewed evaluation was not attempted, allowing for the organisers to maintain direction over the range of aesthetic and conceptual elements presented in the series: whilst the rigour and fairness of the peer-review process is desirable, many festivals and music events do not follow this model, partly because the peer-review of creative works is challenging, partly because the organisers view curation as a creative act for which they respectfully assume a level of authority and finally because many such events do not require submitted work to be in the form of novel research outputs. An interesting discussion on these issues in the curation of a computer music concert series is given by Guedes and Rebelo (2010).

We feel that such an event will benefit from moving towards a more rigorously peer-reviewed process as an implicit understanding of the musical landscape emerges, including reflection on how reviewers are asked to review works. In this case the selection of a program/curatorial committee should accurately reflect contributions to the field. This was only partially attempted in this first instance. However, we also feel that some curatorial control should be maintained, including invited works, with the aim of maintaining diversity, outsider opportunities and general audience appeal.

At the same time as being comparable to a number of digital art and electronic music events, the specific MuMe focus requires balancing contrasting criteria:

- the musical quality: irrespective of the quality of the process, algorithmic complexity or conceptual foundation of the work, is the work suitable to be presented to a public audience?
- the relevance to the MuMe research agenda: in the call for participation, the relevance requirement was defined loosely as addressing, through a software implementation, issues of software autonomy. A definition of autonomy was not offered and respondents were invited to make their own case for how autonomy was addressed.

As this paper and our previous discussions have elaborated (Bown and Martin, 2012), it is challenging to provide specific and precise criteria for what counts as a MuMe-relevant work. In particular, although all involved agreed that autonomy was key, the range of approaches to and concepts of autonomy was broad enough to require case-by-case consideration of the domain of autonomy being considered. This was an ongoing talking point throughout the MuMe Weekend, and it was clear that multiple notions of autonomy in musical contexts are needed, in a similar way to those discussed by Castelfranchi (2000). This is also discussed by ?.

Open challenge #1: *identify a set of clearly defined classes of MuMe-relevance, and their relationships.*

The question of relevance is guided not only by the curator's preference but also by what they see as the purpose of the event. For the 2013 MuMe Weekend the approach was taken that tools, technologies and methods that are capable of having a direct impact on MuMe research were sufficiently relevant, even if they did not clearly and directly

embody autonomy. This included, for example, "live coding", on the basis that flexible and dynamic tools for creating and reconfiguring musical algorithms on-the-fly may provide fertile environments for researching and implementing autonomous behaviour. This relevance was evidenced elegantly in Olofsson's installation piece, *low-life*, where a software agent written in the SuperCollider language mangled its own code to produce new musical effects, and in a visually engaging way. It should be reiterated, however, that in some of these cases software autonomy was not directly evident, despite the verdict of relevance to the theme of autonomy.

Through discussion the authors decided that this more inclusive approach to relevance was valuable, at least in light of *open challenge #1*, in that (i) it raised the musical quality, (ii) it stimulated discussion on the topic of relevance, and also more generally provided a broader pool of ideas to stimulate new approaches, and (iii) it favoured the avoidance of false negatives over that of false positives in the selection process. A work may be only minutely metacreative, but nevertheless constitute a meaningful and groundbreaking piece of MuMe research. Whilst this inclusive approach was deemed overall beneficial, (i) it should not be pursued at the cost of undermining the core focus on metacreation, and (ii) there remains a clear commitment to raising the quality and quantity of MuMe relevant submissions in future years, to which responses to *open challenge #1* will contribute. Whilst this goal demands further refinement of what constitutes good work on the field, and may promote some degree of factional specialisation, we support maintaining an inclusive breadth.

Audience and Contributor Engagement

The use of printed program notes is standard in the presentation of contemporary and electroacoustic music, both in the context of public concert series, and at academic conferences (such as NIME, ICMC, SMC etc.). In such text, technical detail falls far short of that provided in a paper, in part due to length, but also due to the potentially conflicting norms associated with such notes: typical discourse is that of an artist communicating their creative intent. This was identified as an area of dissatisfaction, but without any clear suggestions for development.

A panel discussion was held at the beginning of the weekend in which the system designers present discussed their ideas and responded to questions about the musical and technological intentions behind their works. The panel discussion proved interesting in mediating between an 'artist talk' form of discourse and technical and theoretical forms of discussion. Two areas of discussion to emerge from the session were (i) contributors comparing notes on individual working methodologies, finding some commonality in an iterative approach to development that combined artistic and engineering-based techniques, and (ii) the framing of individual works and practices in terms of a bigger picture of possible future directions, from developments in music-consuming industries such as games and advertising to new forms of metacreative art in the mainstream. In both cases this helped contextualise existing practice in terms of better

known creative methods for both the audience and contributors. Structured panel sessions could also be used to go into greater technical detail via interrogation, to collaboratively develop new methods, to consolidate use of terms and, after events, to critically examine works, and in these cases would benefit from narrowing to an expert audience.

In addition to the call for performance and installation works, we also included a call for a tentative alternative submission format. We requested video documentation that could be used for a ‘listening room’ environment, with the visual material acting as supporting technical documentation to the recorded musical work (i.e., *not* artistic in nature). Whilst the audio submitted should consist of complete musical works, the call asked for each work to be combined with video material that provided a technical description of the works, possibly tracking the development of the music with pointers to what the audience should listen out for (an entertaining example of this is given by Al Biles in his TEDx talk²).

In fact no submissions were made in this category. Since the documentation of work is time-consuming and often demanding, this was understandable. Nevertheless, we propose that MuMe needs innovation on the presentation of work that communicates both artistic and technical aspects of the work, and the proposed format may still be of value.

Genres and Submission Categories

The MuMe Weekend focused on three broad musical categories: improvisation, electronic dance music and electroacoustic composition, represented in three different concerts. An additional electroacoustic concert was scheduled to incorporate related work. Three installation works were also hosted. The video documentation of works was also included as a submission category, but this did not attract responses.

Improvisation

The “Improvising Algorithms” concert followed a format that has become common amongst researchers interested in musical autonomy: a free-improvised duet between a human instrumentalist and a software system, typically interacting only via the audio produced and heard by each participant (except in some cases where MIDI is treated as a substitute for audio). The format has been most clearly established in the works of Lewis Lewis (2000), but other researchers in the field such as Rowe (1992) and Biles (2001) are also known for using the improvised duet format in less freely improvised (more structured) forms. The Live Algorithms for Music Network in the UK (Blackwell and Young, 2005) developed this as a showcase format amongst gatherings of researchers.

The “Improvising Algorithms” concert was hosted at the Old Darlington School in Sydney, a 19th century school building owned by the Sydney Conservatorium of Music. The building is frequently used as a venue for new and improvised music by various Sydney groups, typically with an informal atmosphere.

²<http://www.youtube.com/watch?v=rFBhwQUZGxg>

Electronic dance music

Electronic dance music was presented in an “Algorave” at the ‘artist-run space’ *107 Projects*. The term was coined in 2010 by McLean and Collins (pers. comm.) and the description on the Algorave Facebook page (<http://www.facebook.com/algoravers>) - “sounds wholly or predominantly characterised by the emission of a succession of repetitive conditionals” - captures the general focus on electronic dance music that is algorithmically generated (neither specifically MuMe focused, nor specifically limited to live coding, with which the format is closely associated). Past Algorave events have had a strong focus on live coding performance, typically with live code projected on screens providing a visual component to performances. The MuMe Algorave followed this new concert format by engaging live coding performance, but also involved machine generated compositions, both pre-rendered or generated in realtime, for which there is no widely recognised precedent (the absence of any performer being conspicuous).

Electroacoustic composition

The “Studies in Autonomy” concert followed traditional electroacoustic concert lines, such as the performance series found at the International Computer Music Conference, with a mixture of fixed media pieces, live interactive pieces involving human musicians, and generative works. This entailed a less genre-specific set of works (indeed use of the term “electroacoustic” belies the breadth of pieces) than either the “Improvising Algorithms” or “Algorave” concerts, instead allowing more of a focus on the conceptual variation between works.

Related work

Finally, an additional concert was put on to accommodate works that were not deemed directly relevant to the MuMe research agenda. Quality works were accepted that were interesting examples of the use of computational processes, and that may be thought-provoking to an audience engaged with the theme of autonomy. Perhaps despite the curators’ decisions, these works may have appeared to the audience as better engaging artistically with notions of autonomy than other works, providing a possible avenue of feedback regarding selection and evaluation, discussed elsewhere.

Reflections

Arguably, the major classification between types of MuMe work lies between performative agency and memetic agency – terms used by Bown, Eldridge, and McCormack (2009) to refer to agency that occurs within a concert performance and during an offline composition process respectively. It may be counterproductive to view genre distinctions, such as between dance music and electroacoustic music, as fundamental. Nevertheless, besides the common sense of keeping these genres distinct from an audience-experience point of view, these forms, and other genres that could be added here, also differ in various senses with respect to the social process of their production, for example in the use of remix or multi-artist sub-genres found in dance music.

The differing conventions and expectations of these areas suggest a certain amount of domain-specificity within the broader MuMe agenda. MuMe research could be subdivided into differing domain-specific sets of objectives, perhaps in some cases carving out original genres (this is discussed further below). At the same time, the crystallisation of domain or genre-specific interests may be detrimental if exclusive of outlying work. In all cases, such work arguably needs to be presented in ways beyond the traditional combination of (a) academic paper and (b) musical performance or recording.

Open challenge #2: (a) *determine pertinent MuMe genres and establish best practice for their presentation, (b) establish suitable genre-agnostic presentation formats.*

What issues were there in the selection and presentation of the work?

A challenge in the curation of works was attempting to understand in detail from the submission materials exactly how each proposed system worked, in order to make judgements about the relevance to MuMe research. Applicants were asked to provide sufficient technical information about the working of the system for the curators to be clear about what the system did, albeit within a tight 500 word limit. Some submissions did not achieve this clarity, but it was unclear whether this was a misinterpretation of the submission requirements or more conscious avoidance of an explanation of the system details. This was tricky in cases where the limited information and demonstration material provided did imply conceptually and technically advanced systems. Another possible danger in the review process is that mere mention of common MuMe-relevant concepts and technologies such as statistical modelling, or the implementation of goal formation, may be considered a suitable-enough trigger for acceptance of the relevance requirement. As with other peer-review tasks, it is possible to obtain a surface understanding of the composite elements within a system, without gaining a deep understanding of the role of these elements, or to correctly position the work in terms of its innovation (e.g., compare a truly novel innovation to a less innovative recombination of elements).

Yet the requirements of reproducibility and technical clarity typical in scientific peer-review are not well established in the submission of creative works. System descriptions are of limited utility without supportive results, but as discussed above, there are no clear standards for what constitutes a persuasive result. Saying that 8 out of 10 audience members believed the music to be human composed is insufficient due to the “context problem” discussed below, and providing technical results, for example showing accurate gesture recognition or the size of a search space, can only be indirectly applied to the ultimate goals of musical metacreation.

Where musical examples were provided, as required, curators could call on their own aesthetic experience to judge the quality of outputs. A common problem, related to the preceding issue, is understanding the distinction between human and machine input into a composition. There are various ways in which human abilities can be disguised as software abilities. Serendipitous events captured in recordings

can be picked to give the impression of greater musical intelligence. Equally, the capabilities of a piece of software may be obscured: a system output may be relatively uneventful, lending weight to the response that the system is uninteresting, despite many human compositions utilising minimalism, repetition and duration to great effect. In undramatic works it becomes harder to reasonably ascribe ability to a system. This biases judgement, just as we would struggle to praise a machine that produced something akin to Yves Klein’s minimal paintings rather than something akin to Van Gough’s painterly dynamism. This can be resolved either through clarity of description as discussed above, or through the more concrete evidence of a vast range of examples of system output, or even an interactive interface to explore the system behaviour. Note that a great amount of human involvement in the aesthetic quality of the output need not weaken the claims of the system as this depends on what claims are being made. Such improved communication may enable MuMe creators to take the baby steps required of scientific research whilst producing artistically strong work.

Open challenge #3: *Devise requirements for the communication and review of creative MuMe results that enable clarity, the accurate attribution of agency in metacreative processes, and the cumulative, community-driven development of new works and systems.*

Did the event exhibit the state of the art in musical metacreation?

Well-known systems from the literature include those of Cope (1992), Lewis (2000), Rowe (1992), Biles (2001), Young (2008) and Pachet (2004). The techniques used by these developers were found in other pieces presented at the MuMe Weekend, although arguably not to the same standard. Lewis’ *Voyager* system is particularly well-known as an expert system based on rules and strategies derived by Lewis from his musical expertise. The system is particularly effective, producing a rich variety of style and content and is often compared to the painting system AARON, developed by Harold Cohen (McCorduck, 1990). In both cases, the system is typically perceived as an extension of the creator and a sort of embodiment of their (real or desired) artistic capabilities. In both cases the systems may not employ the most advanced AI techniques: neither learn, adapt or are self-evaluating in any deep way. Nevertheless through conceptually relatively simple techniques they excel in establishing a sense of creative autonomy.

Meanwhile, cutting-edge techniques in AI, computational creativity, music information retrieval and related fields may not have made it into MuMe works, as the leading technical literature take time to filter into systems suitable for performance. There is not always an obvious path for adapting these advanced techniques to be used in MuMe systems. Although a number of MuMe performance projects have arguably innovated novel algorithms not found elsewhere in the literature, the question remains as to whether building and deploying a live performance or composition system will ever stand at the frontline of AI algorithms research: much of the groundwork in testing algorithms for

their suitability has been typically achieved in lab environments rather than open-ended creative environments, and no clear method for systematic evaluation exists that would unambiguously confirm the value or applicability of algorithms in this domain.

Open challenge #4: *Provide pathways for integration of leading MIR, AI and professional tools into MuMe outputs, such as standards, services and libraries, or by making open MuMe systems available for use by non-MuMe researchers.*

What were the issues in audience and performer evaluation of systems?

Themes of software autonomy, musical generativity and artificial intelligence in human-computer performance are very new for most concert audiences, even those habituated to computer music and new music performance practice. In addition, the artistic goals and technical innovations of MuMe systems are often intimately linked, often requiring of an audience member a sophisticated appreciation of the relationship between the two. During the presentation of a work of musical metacreation, audience members may also be actively trying to define the varying degrees of agency attributable to system designers, performers and software in a work. This is a peculiarity of the reception of metacreative systems that is not often found in other performance contexts. Given such issues, it may be argued that MuMe-related works benefit from a certain degree of ‘de-mystification’, allowing audience members to engage with the concepts that have given rise to both software and the resultant musical works.

Many of the issues related to curatorship also apply to the evaluation and characterisation of MuMe systems that is required to clearly appreciate the state of the art. While many systems incorporate well-known algorithms from artificial intelligence or machine learning, they are also the result of artistic micro-decisions made during their development. As established, these systems are simultaneously engineering artefacts and artistic creations. The tension between the two disciplines is evident, as Lewis writes, “I don’t feel the need to ‘scientifically’ prove the validity of any process I use to get my music to sound the way I want it to sound. I feel utterly free to work in an intuitive way while programming computers to create improvisations” (Lewis, 1999). Furthermore, arguably, in many cases the intuitive and creative choices to which Lewis refers have an impact on the system’s output that is comparable to the choice of AI or machine learning algorithms.

That many MuMe systems are as much artistic creations as engineering feats may be the reason that few creators of such systems distribute their work for others to use or study; it is not common practise for artists to disclose their methods completely. Similarly, published papers rarely contain sufficient detail to reproduce the systems in their entirety (in many cases to achieve this with clarity would be a feat of technical writing). That published research should be reproducible is an ideal that has received considerable attention in the engineering signal processing community in recent years (Vandewalle, Kovacevic, and Vetterli, 2009)

and throughout science and engineering, while in the field of computer music, reproducibility is part of the agenda of the ‘Sound Software’ initiative funded by the UK Engineering and Physical Sciences Research Council³. However, reproducibility has not yet been widely adopted in the MuMe community and only such evaluation as appears in the papers themselves can be used to compare different systems. As alluded to above, evaluation is a difficult issue and many authors resort to anecdotal characterisation of their creations.

An open question, therefore, is how to achieve greater transparency in published work, possibly by re-focusing artists’ efforts on engineering or scientific outcomes, despite Lewis’ influential call to freedom. Greater transparency would have benefits to the MuMe community by enabling those science and engineering-oriented members to contribute by, for instance, providing theoretical frameworks by which to characterise MuMe systems. This in turn may drive further artistic and technical innovation. Clearly, however, this does not need to be stated in separate terms from *open challenge #2*.

Numerous discussion of the evaluation by audiences and performers of MuMe system can be found in the literature (e.g., Pearce and Wiggins, 2001; Collins, 2006; Hsu and Sosnick, 2009; Ariza, 2009; Eigenfeldt, Burnett, and Pasquier, 2012; Blackwell, Bown, and Young, 2012; Banerji, 2012). During the MuMe Weekend, no formal evaluation was undertaken, but the weekend provided an informal forum for attendees to discuss their impressions of different systems. Four issues concerning the presentation of systems arose as recurring themes in discussion:

The context problem

A single presentation of a piece is severely limited in its ability to convey sufficient information about the system. Even multiple plays, accompanied by a detailed description of the workings of the system often still leaves open questions about what the system achieves. Anecdotal reports of surprising results would even indicate that system designers themselves often don’t have a complete grasp of how their systems work, which has been interpreted both as damning the designer and praising the system.

In an improvised duet performance this is confounded further by the role of the performer. The performer may affect the perception of the system in different ways. They may skilfully exaggerate the appearance of two-way interaction, shedding a positive light on the system (this was evident at the 2005 LAM gathering, as observed by Michael Young, pers. comm.). Equally, they may detract from the system. They may play with too much expectation, posing phrases with the expectation of a response, or expecting things like rhythmic entrainment or the anticipation of a developed trajectory such as a crescendo. The claim of human-like abilities is here an obstacle to achieving the best interaction. Or they may play with too little hope of successful interaction, for example by consistently soloing over the system, treating it like a tape, assuming its behaviour is unresponsive to what they do. Between these extremes we expect to find perform-

³<http://soundsoftware.ac.uk/>

ers striking an appropriate balance given the capabilities of the system (which may vary widely).

The question of how a performer should perform with a system is therefore raised (see Banerji (2012) for an original take on this from an ethnographic perspective), but only as a peculiar subset of what performers routinely think about: how to perform, how to interact with other humans. The same two distinct goals may interfere here as in the creation of systems: to make the best music in the context of a concert, and to best bring out the workings of the system. These goals may have different influences depending on the context. At a significant musical event, the musical quality may take precedent. At an academic demo the system working may lead. It is necessary to recognise the role of context and the effect it has on presentation.

An obvious approach to this is to aspire for the performer to play “naturally”, meaning that the role of the system is not a cause for having to perform, interact and behave in novel and unusual ways that would not be found in human interaction. There is a wide range of responses from performers with respect to this notion. One performer’s response to the first author was that they would always think they were playing with the system creator, not the system, even if the creator is not present or active. In saying this, the performer – who was only moderately experienced playing with improvised electronics – proposed an appropriate naturalisation mechanism. This process also involved numerous rehearsals in which the system was tweaked, reinforcing the notion of system as complex instrument rather than as participant. In other cases, improvisers have been content with an approximate knowledge of the system’s working. Others are just happy to play and allow discovery to take place as part of the performance.

The above examples are likely affected by a multitude of factors specific to each individual performance context, including the performer’s personal style, the amount of time the musician has had in rehearsals with the system, and also the amount of knowledge the performer has of the system’s responsive and generative capabilities. This final point may have a direct bearing upon certain types of interactive scenarios outlined above. For example, if an improviser was told simply that the system listens and responds to their performance, they would be likely to come to personal conclusions about any perceived cause/effect dynamic experienced within an interaction. On the other hand, if a musician is given more specific information about exactly what is being listened to (pitch relationships, timbral characteristics), or details about how the system uses this knowledge in its decision-making, we might expect that the musician’s performative interaction with the system would be coloured by this knowledge. Placing the musician in dialogue with a machine changes this familiar performative scenario, forcing the musician to strike a balance between natural performative expectations and the specific musical context presented by the system.

It has been proposed that audience evaluation should be a central part of MuMe events (e.g., Eigenfeldt, Burnett, and Pasquier, 2012). Whilst all authors agree with this, careful consideration needs to be given to the form and purpose of

the evaluation. With a strong computer science element in our work, it is tempting to look to related fields such as music information retrieval (MIR), where a series of MIR challenges are established with quantified results that can be incrementally improved upon. Such challenges provide a ‘winch’ allowing the field to incrementally draw itself forward. With audience evaluation this may not be achievable. Firstly, asking audiences to engage in a Turing Test (Turing, 1950) evaluation of the music, essentially asking them whether they believe the music to be human made, presents only a plateau around which incremental progress cannot be made. Arguably, it is already trivial to make computer generated music that is capable of sounding like human generated music, even though a number of MuMe research challenges remaining unresolved (Ariza, 2009). Secondly, asking audiences to provide purely aesthetic evaluation is partially tangential to, and thus potentially disruptive to, the objectives of MuMe research, reducing in the worst case to a popularity contest that fails to take into account the system design. While MuMe systems may aspire to be aesthetically appealing to general audiences, optimising such a response is neither necessary nor sufficient to qualify as a successful MuMe system.

Open challenge #5: *Devise assessment formats that would have an equivalent effect to the Turing Test, but that are well suited to the context of music.*

A suggestion by Pasquier is to run events involving audience evaluation, such as ‘battles’ between agents. An example would be a series of improvisation bouts in a performance context mixing multiple agent-pair combinations so that audiences can see the systems acting in a real and open-ended interaction environment and derive judgements about agents’ capabilities. Since as far as musical genres go this is may rather an unreal situation, such a suggestion may be considered in conjunction with the following *open challenge #6*. That said, examples exist of sparring interactions between musicians. For example, in a club series called *Fightpod*, DJs competed in a winner-stays-on process judged by their audience.

An alternative approach is to subjugate evaluation in favour of more open-ended critical discussion, typical of musical criticism. In terms of stimulating discussion between participants, the MuMe Weekend could be claimed to be successful, even when responses to certain pieces were negative. The long-term hope for a regular series of MuMe events is that, as with the workshop series, the discussion evolves along with the maturation of works and the conceptual language used to describe the works. Evaluation may then be a secondary concern. As repeated throughout this paper, the discussion requires clear communication of the working of systems. It could also go beyond this towards forms of presentation that better integrate artistic and scientific discourses.

The modification of genres through the introduction of MuMe elements

Another approach is to design behaviours into systems that offer affordances to an improviser. The “continuation” approach of Pachet (2004) is a good example. Here the system

aims to model the style of the performer and play continuations of their soloing. This requires a turn-taking approach to performance. Although this kind of mimicry is the kind of thing a performer would do, and can be seen as a sign of musical intelligence, the system does this systematically as its central feature, a fact which the interacting performer can latch onto for creative effect. Pachet describes the process as mirroring. A number of other simple strategies are discussed below.

However, such innovations present an interesting scenario: that of potentially changing the genre that is being imitated. In other words, they suggest that rather than human-machine duet improvisation successfully imitating traditional human improvisation, it will, over time, establish itself as a ‘style’ in its own right, with its own conventions, expectations and rules of engagement. Similarly, issues of performance representation in generative dance music are another active site of exploration. This need not be a problem. Instead, such a move provides room to define modes of interaction and listening that are more suited to the ‘in between’ space, discussed below. Parallels to the uncertainty principle in physics or issues of participant observation in anthropology may contextualise this idea.

Open challenge #6: *Innovate new musical forms/genres/contexts that provide new meaningful avenues of expression for MuMe works, and enhance MuMe outputs.* This may be able to subsume “open challenge #5” as well as “#2(a)”.

Long-term structure

Long-term structure is a challenge for musical metacreation for two reasons: it is poorly defined and therefore even more difficult to judge; and it is difficult to integrate both long and short-term structure into a single system, with most developers focusing their efforts primarily on short-term structure, and secondly on the relationship between consecutive elements (i.e., working up towards increasingly large time-scales).

Long term structure in music has been well studied but little has been proposed in terms of good strategies for generativity beyond learning the structure of typical styles. In the MuMe domain Eigenfeldt and Pasquier (2013) recently presented a combined top-down and bottom-up approach with different computational methods being used to contribute to these different components. Martin et al. (2012) present an AI system for working with high-level blocks to achieve control of long-term structure. Since long term structure is loosely defined, a naïve but successful strategy is simply to change parameter values for the system over time. For example, in one work, Young (2008) periodically retrains a neural net that controls his playback system. As discussed, many systems, such as the influential work of Pachet (2004), allow long-term structure to be either led by the instrumental musician, or to develop as a result of an implementation detail of the system (in the case of learning systems, an example may be because their database accumulates content during the course of the piece).

Open challenge #7: *Develop a clearer theoretical understanding of how to discuss and evaluate musical long-term*

structure, its associated decision-making, and its relationship to short-term structure, in MuMe contexts.

Tricks, bugs and serendipity

We have discussed a number of strategies for live algorithm development based on simple analyses of essential elements in human improvisation. Simpler still are tricks that give a sense of meaningful musical interaction despite being fundamentally elementary. An example is to use onset detection to cause computer generated events to happen in perfect temporal alignment with human performer events. In human performance this alignment happens because players are playing in time, that is, are rhythmically entrained to each other. It can also be achieved using visual cues. Computers can respond effectively instantaneously, giving the illusion of entrainment or some other form of anticipation (perceived as such even if listeners are aware of the low-level process). Likewise, systems can instantaneously match the pitch of a note, and can thus perform an instantaneous harmonization of that note. Similarly, serendipity has been reported in many instances as playing an important role in the perception of creativity, or as a creativity-enhancing process in its own right. Unintended behaviours in the system, caused either by bugs or other unknowns, can also be effective in being surprising to both developer and audience.

Minimal strategies exist for achieving meaningful musical interaction that are clearly over-simplifications of human behaviour, but serve MuMe goals. Such minimal strategies may be completely unrelated to human behaviour, such as the triggering discussed above. We may describe these as part of an ‘in between’ space of creative MuMe possibilities, where the behaviours appear to lie in between the mechanical and the human-like, and are purposefully simplistic in pursuit of the goal of meaningful machine interaction.

Open challenge #8: *Discover and document ‘in between’ strategies for musical metacreation, where computers exhibit MuMe-relevant behaviour without necessarily doing so in human-like ways.*

Conclusion

We are confident that the MuMe Weekend was successful in combining the goals of (i) creating an engaging musical experience for a wide audience, (ii) stimulating constructive debate about notions of autonomy and the relevance of various approaches to software autonomy and (iii) bringing together excellent work in the MuMe field. However, we feel that the informal and inclusive approach to this event can be tightened somewhat through innovation in the various non-technical methodologies surrounding the emergence of musical metacreation as an active research area. Our open challenges set targets and define a set of focal points for consolidating MuMe musical events into engaging, constructive research gatherings. These cover communication (challenges #1,3), presentation (challenges #2,6), evaluation (challenges #1,5,7) and development (challenges #2,4,7,8). Although we do not propose solutions here, we feel that the MuMe Weekend was useful in stimulating the discussion that has helped clarify that these are important challenges, and this is in itself a constructive outcome.

Two recurring observations throughout the paper, however, are that (i) work must be communicated with clarity and detail, possibly requiring multiple modes of presentation (e.g., writing, musical examples and software) combining rigour in both artistic and scientific domains, (ii) musical metacreation may invoke an ‘uncertainty principle’ in which genres are distorted and redefined as domains for this kind of research-based creative practice; arguably this distortion should be embraced with new performance formats.

Acknowledgements

We thank all of the contributors, committee members and volunteers who contributed to the 2013 MuMe Weekend. A full list of credits can be found at <http://metacreation.net/mumewe2013/>. The collaboration and event was supported by an International Program Development Fund grant from the University of Sydney.

References

- Ariza, C. 2009. The interrogator as critic: The turing test and the evaluation of generative music systems. *Computer Music Journal* 33(2):48–70.
- Banerji, R. 2012. Maxine’s turing test - a player-program as co-ethnographer of socio-aesthetic interaction in improvised music. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment (AIIDE’12) Conference*.
- Biles, J. A. 2001. Autonomous genjam: Eliminating the fitness bottleneck by eliminating fitness. Available from <http://www.it.rit.edu/~jab/GenJam.html>.
- Blackwell, T., and Young, M. 2005. Live algorithms. <http://www.timblackwell.com/>.
- Blackwell, T.; Bown, O.; and Young, M. 2012. Live algorithms. In McCormack, J., and d’Inverno, M., eds., *Computers and Creativity*. Springer-Verlag.
- Bown, O., and Martin, A. 2012. Autonomy in music-generating systems. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment (AIIDE’12) Conference*.
- Bown, O.; Eldridge, A.; and McCormack, J. 2009. Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound* 14(02):188–196.
- Castelfranchi, C. 2000. Founding agent’s ‘autonomy’ on dependence theory. In Horn, W., ed., *Proceedings of the 14th European Conference on Artificial Intelligence*, 353–357. IOS Press.
- Collins, N. 2006. *Towards Autonomous Agents for Live Computer Music: Realtime Machine Listening and Interactive Music Systems*. Ph.D. Dissertation, Centre for Science and Music, Faculty of Music, University of Cambridge.
- Cope, D. 1992. Computer modelling of musical intelligence in EMI. *Computer Music Journal* 16(2):69–83.
- Eigenfeldt, A., and Pasquier, P. 2013. Considering vertical and horizontal context in corpus-based generative electronic dance music. In *Proceedings of the Fourth International Conference on Computational Creativity*.
- Eigenfeldt, A.; Bown, O.; Pasquier, P.; and Martin, A. 2013. Towards a taxonomy of musical metacreation: Reflections on the first musical metacreative weekend. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment (AIIDE’13) Conference*.
- Eigenfeldt, A.; Burnett, A.; and Pasquier, P. 2012. Evaluating musical metacreation in a live performance context. In *Proceedings of the Third International Conference on Computational Creativity*, 140–144.
- Guedes, C., and Rebelo, P. 2010. Reflections on music programming for conferences: The case of smc 2009. *Computer Music Journal* 34(3):10–19.
- Hsu, B., and Sosnick, M. 2009. Evaluating interactive music systems: An hci approach. In *Proceedings of New Interfaces for Musical Expression (NIME)*.
- Lewis, G. E. 1999. Interacting with latter-day musical automata. *Contemporary Music Review* 18(3):99–112.
- Lewis, G. E. 2000. Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal* 10:33–39.
- Martin, A.; Jin, C.; Carey, B.; and Bown, O. 2012. Creative experiments using a system for learning high-level performance structure in ableton live. In *Proceedings of the Sound and Music Computing Conference*.
- McCorduck, P. 1990. *AARON’s Code: Meta-Art, Artificial Intelligence, and the Work of Harold Cohen*. W. H. Freeman and Co.
- Pachet, F. 2004. Beyond the cybernetic jam fantasy: The continuator. *IEEE Computer Graphics and Applications* 24(1):31–35.
- Pasquier, P.; Eigenfeldt, A.; and Bown, O., eds. 2012. *Proceedings of the 2012 AIIDE Workshop on Musical Metacreation*. Palo Alto, CA: AAAI Press.
- Pearce, M. T., and Wiggins, G. A. 2001. Towards a framework for the evaluation of machine compositions. In *Proceedings of the AISB’01 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences*, 22–32. Brighton, UK: SSAISB.
- Rowe, R. J. 1992. Machine composing and listening with Cypher. *Computer Music Journal* 16(1):43–63.
- Turing, A. 1950. Computing machinery and intelligence. *Mind* 59(236):433–460.
- Vandewalle, P.; Kovacevic, J.; and Vetterli, M. 2009. Reproducible research in signal processing. *Signal Processing Magazine, IEEE* 26(3):37–47.
- Young, M. 2008. NN music: Improvising with a ‘living’ computer’. In et al, R. K.-M., ed., *Computer Music Modelling and Retrieval: Sense of Sounds*. Springer Verlag.