Musical Metacreation: Past, Present, and Future

Paul M. Bodily and Dan Ventura
Computer Science Department
Brigham Young University
Provo, UT 84602 USA
paulmbodily@cs.byu.edu,ventura@cs.byu.edu

Abstract

In order for Musical Metacreation (MuMe) to continue to develop as a community it becomes increasingly important to have a clear definition of what MuMe is. Although some definitions have been promoted, there remains some question about how well they really reflect and drive the field of MuMe research. We present the results of surveying the 80 MuMe papers that have been presented at the first five International Workshops on Musical Metacreation. Using this information we examine the trajectory of MuMe, identifying some of its current strengths and weaknesses. In the interest of identifying goals that distinguish and unify our community, we propose a formal, concise definition of MuMe and suggest ways to build the momentum of the field.

Introduction

This year marks the 6th International Workshop on Musical Metacreation (MuMe). The five previous workshops spanning the years 2012 to 2017[1] have yielded 80 publications in the form of technical papers, demos, position papers, studies, and performance reviews. This represents a strong history for the workshop, and in order for MuMe to continue to progress as a field, it seems critical that it develop and maintain a strong, unique, and focused identity.

We have surveyed the 80 publications to identify what MuMe claims as its identity and how well we live up to this in practice, collecting the following metrics [2]:

- *Common topics* within MuMe literature;
- *Frequently cited sources* in MuMe papers;
- The number of internal and external *citations* of MuMe;
- The *return rate* of authors at MuMe; and
- *Evaluation strategies* used in MuMe research
- *Musical genres* represented in MuMe; and
- *Computational approaches/technologies* used in MuMe.

The purpose in collecting this data is to identify what, in practice, defines MuMe research, what is our collective focus, and in what ways do we need to improve to reach our stated objectives. We start by examining several of these stated objectives—most of which have been made only informally. Based on these observations, we propose a formal definition of goals of MuMe research. We use the data collected to identify patterns within the community that may be preventing us from effectively reaching these goals and accordingly make suggestions for improvement.

What is MuMe?

Key to the success of any field is the definition of scope: what *is* and *is not* included within a particular domain of study. This does not necessarily require the scope to be narrow, but it should be reasonably well-defined in order to promote cohesion and collaborative growth. Several informal definitions of the scope of MuMe have been presented over the past years, and we examine those here.

In the preface to the *First International Workshop on MuMe*, the organizers wrote that MuMe "aims to bring together artists, practitioners and researchers interested in developing software and systems that autonomously (or interactively) recognize, learn, represent, complete, accompany, compose, or interpret music" (Pasquier, Eigenfeldt, and Bown 2012). Several aspects of this definition have been repeatedly emphasized within the MuMe community (e.g., see (Eigenfeldt et al. 2013)), specifically

1. the bringing together of artists, practitioners, and researchers alike, and
2. the focus on autonomy or interactivity.

Beyond these two aspects, however, the definition of what MuMe "people" are attempting to accomplish *through* this autonomy and interactivity has been somewhat less consistently defined. Within the same preface, for example, is extensive discussion about *computational creativity* and *metacreation*, the latter being defined as "the idea of endowing machines with creative behavior" (Pasquier, Eigenfeldt, and Bown 2012). This emphasis on designing systems that are autonomously creative has been repeated several times throughout the MuMe literature:

- "Musical metacreation, borrowing from the more general definition of computational creativity, is the idea of endowing machines with creative behavior" (Eigenfeldt et al. 2013).

[1] A tutorial was held in place of a workshop in 2015

[2] Raw data is available as an online appendix at *anonymized*

- "Musical Metacreation is a nascent field emerging from the larger umbrella field of computational creativity, itself an area that has mostly materialized in the last two decades. MuMe focuses on endowing machines with the ability to achieve creative musical tasks" (Bown et al. 2016).
- "Musical metacreation (MuMe), also known as musical computational creativity, is a subfield of computational creativity that focuses on endowing machines with the ability to achieve creative musical tasks" (Pasquier et al. 2016).

These and other definitions also commonly include lists of what constitute "creative musical tasks". Eigenfeldt et al., for example, have asserted that MuMe can be divided into a number of "canonical problems" (2013).

As a field, MuMe cannot accomplish its goal of "endowing machines with creative behavior" without making some minimal attempt at defining "creative behavior". Indeed, this has been a significant point of emphasis in the broader field of computational creativity (Wiggins 2006; Colton 2008; Colton and Wiggins 2012; Ventura 2016). The difficulty in defining "creative behavior" has precedent in the challenge confronted by the artificial intelligence community when attempting to define "intelligent behavior". Years of debate yielded a definition that has become widely accepted if for no other reason than a better one could not be found: "The performance of tasks, which, if performed by a human, would be deemed to require intelligence" (Wiggins 2006). This ultimately led Colton and Wiggins to formalize a similar definition for *computational creativity* (CC): "The philosophy, science and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative" 2012. This is suggestive of a definition of MuMe as

> *The philosophy, science and engineering of computational systems which, by taking on particular responsibilities, exhibit* musical *behaviours that unbiased observers would deem to be creative.*

This definition highlights two aspects of MuMe:

1. MuMe includes philosophy, science, and engineering; and
2. MuMe systems are and are not creative as a function of how they are perceived by unbiased observers.

So far we have done nothing more than summarize and clarify the scope and objectives of the MuMe field. In the following we examine, using aggregate data collected from the 80 MuMe manuscripts, the extent to which the workshop represents the scope of MuMe and the degree to which the workshop is contributing to its stated objectives.

Survey Results

Here we summarize the results of surveying the manuscripts for a variety of criteria and interpret the results to make some suggestions as to what they indicate about how we can further strengthen our community. Results are aggregated over all five years of the *International Workshop on MuMe*.

Table 1: External Non-Auto-Citations by Year

	'12	'13	'14	'16	'17	All
Papers	16	21	10	20	13	80
Papers with citations	11	16	6	3	0	36
Total citations	59	90	20	4	0	173
Avg citations/paper	3.7	4.3	2.0	0.2	0	0.46
Max citations/paper	18	26	22	3	0	26

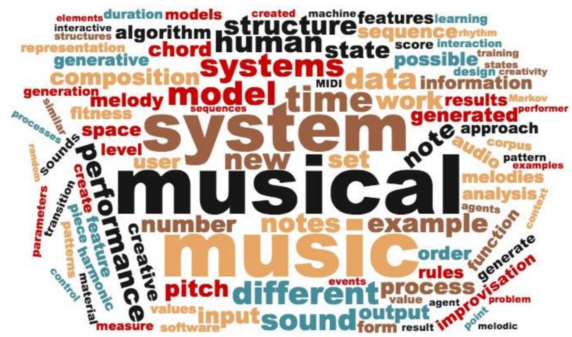

Figure 1: Word cloud of 100 most common words from MuMe papers from 2012-2017 where size represents frequency.

Common topics To identify common topics discussed in MuMe, we performed a word-frequency analysis across all 80 manuscripts (see Figure 1). The most frequent terms (and their counts) are *musical* (1361), *music* (1318), and *system* (1125). Other frequent terms included: *time* (598), *model* (545), *data* (491), *sound* (488), *performance* (443), and *example* (429).

Frequently cited sources To identify the common theoretical bases among MuMe researchers, we collected aggregate bibliographic information across the manuscripts. Sources that are cited by more than 15 MuMe papers (barring self-citations) across multiple years include:

- John Biles' work on genetic algorithms for generating jazz solos (GenJam) (e.g., (Biles 1994; 2001));
- Francois Pachet's work on interactive musical systems (e.g., (Pachet 2003; 2004));
- Shlomo Dubnov, Gerard Assayag, and Arshia Cont's work on Variable Markov Oracles (e.g., (Dubnov, Assayag, and Cont 2007; 2011)); and
- David Cope's work on Experiments in Musical Intelligence (EMI) (e.g., (Cope and Mayer 1996))

External Citations To examine the impact that MuMe research is having outside of the community, we examined the number of citations[3] (barring self-citations) from publications in venues outside of MuMe (see Table 1).

To identify communities with overlapping interests, we compiled a list of conferences and journals that cite the work being done in the MuMe community. Besides published reviews of the MuMe workshop, venues which cite MuMe research the most are (in order) the *International Conference on Computational Creativity* (ICCC); *Computers in Entertainment* (CIE); the *Computer Music Journal* (CMJ); the

[3] We use citation counts as reported by Google scholar

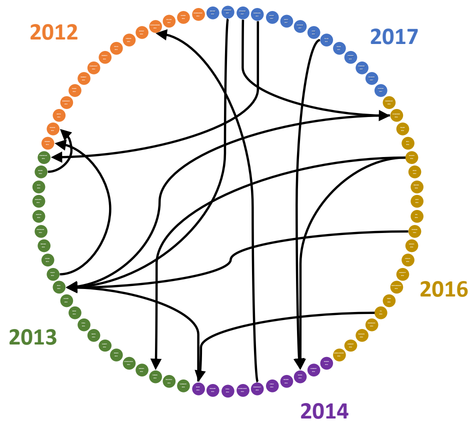

Figure 2: *Internal Citations.* Each node represents a MuMe publication color-coded according the year it was published. An arrow from A to B (in a previous year) means A cited B.

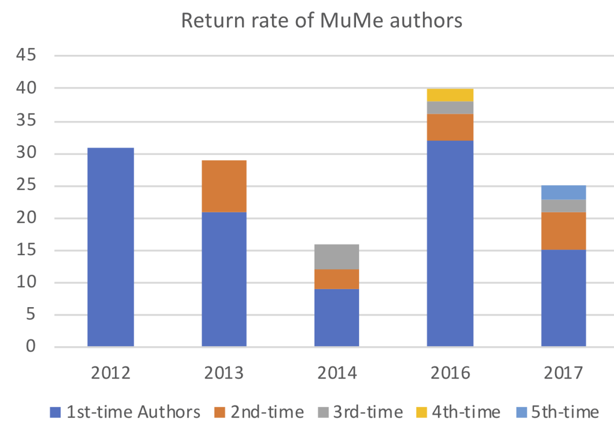

Figure 3: A break-down of the number of authors each year categorized by the number of of previous MuMe workshops at which the author has published.

Sound and Music Computing Conference (SMC); and the *International Computer Music Conference* (ICMC).

Internal Citations Of significant interest is how often MuMe researchers cite each other. For each MuMe workshop we tallied the number of citations (barring self-citations) of papers from previous workshops (see Figure 2). Over the five workshops, we find a total of 13 instances where MuMe researchers cite other MuMe researchers.

Return rate The identity and focus of a growing community is affected in large part by the number of members who are regularly participating in the community events. As a superficial gauge of this involvement, we looked at the number of workshops at which authors (lead or co-author) have presented work. Of 111 authors, we find that 88 (79.2%) have published once; 13 (11.7%) have published twice; 8 (7.2%) have published three or more times (see Figure 3).

Because it is possible that some authors return to MuMe even without having papers accepted for publication, author return rate (as measured by re-publication rate) represents an upper bound on the churn rate within the community.

Strategies for Evaluating Creativity Evaluation is a fundamental component of the scientific method, is necessary for any academic setting, and is critical to the progress of any kind of community goal. In creative domains it is important to distinguish between *self-evaluation* (which has been defined as an important characteristic of creative systems (Wiggins 2006; Jennings 2010) and *external validation*. We looked primarily at external validation strategies, although we estimate that roughly one third of MuMe systems had arguably some form of self-evaluation (genetic fitness functions; formal metrics such as information rate, likelihood, entropy, and accuracy; or input from a panel of critics).

Some publications demonstrated systems with live audiences. Many of these had solicited user feedback and had included quotes or anecdotes about the system's ease of use or creativity. Other systems included informal self-assessments of the system's performance by the authors themselves. A few analyzed the output of their systems using measures such as entropy or predictive accuracy on training data.

A few examples exist of more formal evaluations. Murray and Ventura asked music musicians to perform manual clusterings of generated music to evaluate how well songs evoked intended musical styles (2012). Yu et al. created a system for extracting composition rules and compared these rules to known rules for Bach chorales (2016). Teixeira and Pinto issued a questionnaire asking audience members questions such as "Do you consider these artifacts as music?" and " Would you associate these artifacts to the Rock genre?" (2017). Although several papers cited plans of including formal evaluations in future work, these were the only systems we found that included formal evaluation.

Music genres Does MuMe focus more on some music genres than others? We looked at the genres of music for which systems were implemented. We found that of systems presented in the last two years, 8 (27.6%) were implemented for electro acoustic or EDM music; 5 (17.2%) for jazz; 4 (13.8%) for experimental or feedback music; 3 (10.3%) for classical; and 3 (10.3%) for pop/rock. One system was designed for video game music and 5 did not specify a genre.

Computational approaches and technologies We were also interested to see what sort of breadth of computational approaches and technologies have been used in MuMe. General findings of systems presented in the last two years showed a significant variety in computational approaches including: restricted Boltzmann machines, evolutionary algorithms, grammar-based systems, neural networks (NN), recurrent NNs, long-short term models (LSTM), hierarchical modeling, hidden Markov models, *n*-gram models, rule-based models, random forests, and variable Markov models.

Similar variety existed with regards to technologies used for implementing systems: openGL, OSC, JSON, Magenta, GenJam, MIRToolbox, Max/MSP, Extempore, Python MIDIUtil, ChucK, Pure Data, and SuperCollider. Music formats included mp3, WAV, MIDI, and music XML. Languages included C, C++, Java, Python, R, Swift, and MatLab.

Strengths and Opportunities

These survey data demonstrate areas in which the MuMe community is doing well and also some important ways in which the community can grow to achieve its goals.

Strengths

Overall the MuMe workshops have been very successful at bringing together artists, practitioners, and researchers alike. This is evidenced in the MuMe literature by the many references to concerts, exhibits, and demonstrations in which authors have exposed their work. The most prominent evidence of the balance between research and artistry is perhaps the MuMe concerts (held in 2013, 2014, 2015, and 2017 in conjunction with the MuMe workshops and tutorials) in which autonomous and/or interactive MuMe systems perform their works autonomously or interact with their metacreators.

This also evidences how MuMe has been successful at maintaining a narrow focus on autonomous and interactive systems. This is evidenced in topics that are frequently mentioned in MuMe papers (e.g., *performance* (443), *interaction* (180), *agents* (174), *agent* (164), and *interactive* (161)). Of the frequently cited sources in MuMe publications, most focus on these aspects of MuMe.

MuMe has to some degree limited churn, with 40% of 2017 MuMe authors having previously attended and presented at a MuMe workshop. The leadership and guidance of those returning to the workshop each year (which likely includes others *not* included in that statistic) is critical to maintaining a focused discussion about MuMe and its goals.

Despite the young age of the workshop, publications at MuMe enjoy a reasonable rate of external citation. This includes citations from notable venues including *Advances in Neural Information Processing Systems* (NIPS) (h5-index:101); *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP) (67); *IEEE MultiMedia* (24); *ICCC* (19); and *Artificial Intelligence and Interactive Digital Entertainment Conference* (AIIDE) (19).

Opportunities to Grow

Although MuMe has at times been claimed to focus on "musical computational creativity", there are several indications in the data that suggest a lack of focus on (system) creativity. This discussion is critical to satisfying the claim that we are "endowing machines with creative behavior" (Pasquier, Eigenfeldt, and Bown 2012). The list of frequently cited papers focuses heavily on concepts related to engineering of musical systems; however, noticeably absent are papers relating to creativity. Several of the most seminal works on the philosophy of computational creativity (including (Wiggins 2006), (Jennings 2010), and (Colton and Wiggins 2012)) are never cited in MuMe papers ((Ritchie 2007) is cited once). Many of the buzzwords in the philosophical discussion of creativity (several of which derive from Boden (1977), Wiggins (2006), or Ritchie (2007)) are infrequently referenced or entirely missing in MuMe papers including *novelty* (46 individual references), *surprise* (25), *intentions* (23), *intention* (13), and anything related to the terms *inspiring* (as in *inspiring set*) (0), *typicality* (0), and *self-awareness* (0). Although a few MuMe authors focus on defining and demonstrating *creativity* in their systems (e.g., (Surges and Dubnov 2013; Lynch 2014; Tatar and Pasquier 2017)), they are by and large exceptions to a body of work that is otherwise typified by a lack of discussion about fundamental aspects of creativity.

The lack of discussion about what defines creativity in musical systems is possibly correlated with two additional vulnerabilities exposed by the data. First among these is the low internal citation rate. The fact that MuMe authors infrequently cite one another should be cause for some alarm. After all, one purpose of a community is to draw from and build on one another's ideas (incidentally this is also frequently cited as an important aspect of creativity in general). Are we talking past each other? Do we not share common goals? Do our systems draw on common concepts? Or are we simply failing to communicate/demonstrate the commonality? Is this difficulty in engaging in productive dialogue caused by churn within the community? Or is it *causing* churn? We suggest that this problem may resolve itself as we as a community strive to focus on underlying philosophical and academic principles about what *is* a creative musical system and what *concepts* are *shared* across MuMe systems that we can deliberate and develop (see also (Bodily and Ventura 2017)).

A third opportunity for growth is a focus on formal evaluation—demonstrating "musical behaviors that *unbiased observers* would deem to be creative." Without this external validation, there is no contestable argument for or against creativity. Casual assessment by the authors of a system or a few individual users is insufficient to assert that a system is creative; there is no statistical rigor in such evaluations and therefore no significant conclusions can be drawn. Furthermore it makes it difficult for the community to evaluate and discuss the merits of a given approach to musical metacreation because we do not actually know how well it works. This lack of formal evaluation likely contributes to a failure to internally cite other MuMe research. Progress in any academic discipline depends on formulating hypotheses, designing experiments to test formulated hypotheses and then measuring the statistical significance of the hypothesis. What should be measured? We suggest that engaging in the broader discussion (e.g., see (Ackerman et al. 2017)) of *what* defines creativity and *how* creativity should be assessed will strengthen our ability to meet needs in the community.

Diversity Though some diversity exists among the musical genres addressed by MuMe systems, there is still room for improvement. The benefits of such diversity are obvious: different genres expose different aspects of musical creativity (e.g., consider the difference in the creative tasks in monophonic jazz improvisation, choral arrangement, and ragtime piano composition). Such diversity will be both the cause and effect of increasingly focused discussions about foundational aspects of musical metacreation.

MuMe, like other areas of computational science, also suffers from a significant lack of gender diversity. Unlike other areas of computational science, however, MuMe has the distinct advantage of having overlap with the arts and humanities. We suggest the community consider ways to promote diversity including proactively recruiting keynote speakers and submissions from target demographics.

Discussion

We suggest that renewed focus on principles of CC in MuMe will help to promote a common language, develop a productive dialogue in the community, and encourage improvements to external evaluation in MuMe systems. What then are practical ways to increase this focus? As a starting point for the discussion, here are a few ideas: choose wording in MuMe calls, topics, and paper types to emphasize CC; encourage paper reviewers to look for elements of CC; dedicate special sessions of MuMe to CC topics; invite guest speakers/special panels at the workshop to focus on CC; develop education tools for teaching MuMe to students; and consider ways to accentuate CC in MuMe concerts.

Computational creativity—the question of what makes a computer creative—is perhaps the most defining feature that distinguishes MuMe from other music conferences. We have presented a definition of MuMe that highlights this characteristic and have conducted a review of MuMe literature showing that in addition to our strengths, we have much to gain from improving how we explore computational creativity in action in MuMe.

References

Ackerman, M.; Goel, A.; Johnson, C. G.; Jordanous, A.; León, C.; Pérez y Pérez, R.; Toivonen, H.; and Ventura, D. 2017. Teaching computational creativity. In *Proceedings of the Eighth International Conference on Computational Creativity*.

Biles, J. A. 1994. GenJam: A genetic algorithm for generating jazz solos. In *Proceedings of the International Computer Music Conference*, 131–137.

Biles, J. A. 2001. Autonomous GenJam: eliminating the fitness bottleneck by eliminating fitness. In *Proceedings of the Genetic and Evolutionary Computation Conference Workshop Program*.

Boden, M. 1977. *Artificial Intelligence and Natural Man*. Harvester Press.

Bodily, P. M., and Ventura, D. 2017. HBPL: a framework for debating, developing, and reusing foundational models of musical metacreativity. In *Proceedings of the Fifth International Workshop on Musical Metacreation*.

Bown, O.; Eigenfeldt, A.; Pasquier, P.; and Dubnov, S. 2016. Special issue on musical metacreation, part II. *Computers in Entertainment (CIE)* 14(3):1.

Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In *Proceedings of the European Conference on Artificial Intelligence*, 21–26.

Colton, S. 2008. Creativity versus the perception of creativity in computational systems. In *AAAI Spring Symposium: Creative Intelligent Systems*.

Cope, D., and Mayer, M. J. 1996. *Experiments in Musical Intelligence*. AR Editions Madison, WI.

Dubnov, S.; Assayag, G.; and Cont, A. 2007. Audio oracle: A new algorithm for fast learning of audio structures. In *Proceedings of International Computer Music Conference*.

Dubnov, S.; Assayag, G.; and Cont, A. 2011. Audio oracle analysis of musical information rate. In *Proceedings of the IEEE International Conference on Semantic Computing*, 567–571.

Eigenfeldt, A.; Bown, O.; Pasquier, P.; and Martin, A. 2013. Towards a taxonomy of musical metacreation: Reflections on the first musical metacreation weekend. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment Conference*.

Jennings, K. E. 2010. Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines* 20(4):489–501.

Lynch, M. F. 2014. Motivation, microdrives and microgoals in mockingbird. In *Proceedings of the Tenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Murray, S., and Ventura, D. 2012. Algorithmically flexible style composition through multi-objective fitness functions. In *Proceedings of the First International Workshop on Musical Metacreation*.

Pachet, F. 2003. The continuator: Musical interaction with style. *Journal of New Music Research* 32(3):333–341.

Pachet, F. 2004. Beyond the cybernetic jam fantasy: The continuator. *IEEE Computer Graphics and Applications* 24(1):31–35.

Pasquier, P.; Eigenfeldt, A.; Bown, O.; and Dubnov, S. 2016. An introduction to musical metacreation. *Computers in Entertainment (CIE)* 14(2):2.

Pasquier, P.; Eigenfeldt, A.; and Bown, O. 2012. Preface. In *Proceedings of the First International Workshop on Musical Metacreation*.

Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines* 17(1):67–99.

Surges, G., and Dubnov, S. 2013. Feature selection and composition using pyoracle. In *Proceedings of the Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*, 19.

Tatar, K., and Pasquier, P. 2017. MASOM: A musical agent architecture based on self-organizing maps, affective computing, and variable Markov models. In *Proceedings of the Fifth International Workshop on Musical Metacreation*.

Teixeira, J., and Pinto, H. S. 2017. Cross-domain analogy: From image to music. *Proceedings of the 5th International Workshop on Musical Metacreation*.

Ventura, D. 2016. Mere generation: Essential barometer or dated concept? In *Proceedings of the 7th International Conference on Computational Creativity*, 17–24.

Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.

Yu, H.; Varshney, L. R.; Garnett, G. E.; and Kumar, R. 2016. MUS-ROVER: A self-learning system for musical compositional rules. In *Proceedings of the 4th International Workshop on Musical Metacreation*.