Musical Metacreation: Papers from the 2013 AIIDE Workshop (WS-13-22)

Prediction and Proactivity in Real-Time Interactive Music Systems

Andrew R. Brown and Toby Gifford
Queensland Conservatorium Griffith University, Brisbane, Australia
{andrew.r.brown} {t.gifford}@griffith.edu.au

Abstract

We advocate for the use of predictive techniques in interactive computer music systems. We suggest that the inclusion of prediction can assist in the design of proactive rather than reactive computational performance partners. We summarize the significant role prediction plays in human musical decisions, and the only modest use of prediction in interactive music systems to date. After describing how we are working toward employing predictive processes in our own metacreation software we reflect on future extensions to these approaches.

Background

This position paper discusses proactive uses of prediction in real-time interactive music systems. Proactivity, in contrast to reactivity, involves committing to musical actions on the basis of predictions of future musical events—generally predictions of the future actions of other agents in an ensemble, be they human or computer. We underscore the importance of prediction for perception, cognition and aesthetics generally, and suggest prediction could be more widely utilised in computational music systems.

One benefit of prediction is that it allows for timely action during synchronous interaction. By not having to wait for complete information to be available, actions can be anticipated and synchronised. More generally we suggest that such behaviour in computer systems could help convey an impression of musical agency, rather than having the human always 'in the lead'. We hope that predictive features may move systems toward behaviour that seems more like a musical partnership.

There is evidence of the cognitive significance of prediction and proactivity from several quarters, including theories of perception (Bruner 1967), aesthetics (Meyer 1973) and intelligence (Hawkins 2004). In the following sections we briefly outline a few pertinent examples of such.

Prediction in Perception

The importance of prediction in perception is discussed by Bruner, who takes pains to highlight the "anticipatory and exploratory nature of much human categorization. In the case of most categorising, we attempt to find those defining signs that are as *sure* as possible as *early* as possible to give identity to an event" (1967:13).

Bruner also argues that anticipation is an important aspect of creativity, and states "an act that produces effective surprise ... I shall take as the hallmark of a creative enterprise" (Bruner 1979:18). He suggests the experiences of art involve "a flow of rich and surprising fantasy, a tangled reticle of associations that gives fleeting glimpses of past occasions, of disappointments and triumphs, of pleasure and unpleasure ... at this level, thinking is more symphonic than logical, one theme suggesting the next by a rule of letting parts stand for wholes" (1979:70-72).

Prediction in Aesthetics

Meyer (1973) was interested in musical patterning and how, through repetition, these patterns lead to expectations. It was his contention that degrees of conformity and deviation from expected patterns give rise to musical meaning. Exposure to music contributed to the formation of expectations and, in turn, to a sense of aesthetic appreciation.

A number of studies in psychology and neuroscience corroborate Meyer's insights. Rohrmeier and Koelsch's (2012) recent survey concluded "predictive information processing is fundamental to music in three ways … Prediction and expectancy incorporate the essence of the dynamics of musical temporality. Further they make the experience of local or large-scale goal-directed processes in music possible … Predictive processing constitutes a major process involved in musical interaction and synchronisation … Finally, processes of expectancy and prediction are understood to be linked with specific emotional and aesthetic musical effects" (2012:164).

Prediction in Intelligence

Hawkins suggests prediction "is the *primary function* of the neocortex, and the foundation of intelligence" and that "even behaviour is best understood as a by-product of prediction" (2004:89). In his theory, perception and cognition involve constant prediction, with sensory information categorised into invariant structures, acting as schemata for extrapolation.

Music is a telling case, and one Hawkins regularly refers to, because it affords natural description in terms of invariant representations such as intervallic motion, rhythmic ratios and spectral signatures. However, these require specific starting values (a note's pitch, duration and tone color) and local context (such as current key, metre, and room acoustics) to be operationalized.

A prediction can also be characterised as an "analogy to the past" (Hawkins and Blakeslee 2004:205) because it involves comparison and extrapolation to previous experiences. The case for analogy as being at the centre of cognition is argued by Douglas Hofstadter and Emmanuel Sander (2013) with particular emphasis on language usage to illustrate their points. Their description of the way analogy functions to trigger ideas bears resemblance to discussions above of perception, aesthetics, and intelligence. They suggest that "at every moment of our lives, our concepts are selectively triggered by analogies that our brain makes without letup, in an effort to make sense of the new and unknown in terms of the old and known" (2013:3). These analogies then condition the operation of ideas.

Bar (2007) gives evidence in support of this notion from cognitive neuroscience: "Rather than passively 'waiting' to be activated by sensations, it is proposed that the human brain is continuously busy generating predictions that approximate the relevant future … information is extracted rapidly from the input to derive analogies linking that input with representations in memory. The linked stored representations then activate the associations that are relevant in the specific context, which provides focused predictions. These predictions facilitate perception and cognition by pre-sensitizing relevant representations." (2007:280).

Given this widespread support for the role of prediction in cognition we hypothesise that operationalizing and implementing predictive processes may be of benefit to computational creativity generally, and musical metacreation particularly in the case of interactive music systems.

Prediction in Computational Creativity

Human cognition aside, what evidence is there that prediction can improve artificial intelligence or creativity in computing systems? The work of Hawkins and Hofstadter in computational intelligence is quite well known (e.g., Hofstadter 1995, Hawkins and George 2006). A number of researchers, including the authors, have applied Meyer's (1973) ideas of expectation and the later development of these by Narmour (1990) to computational creativity. We will detail our own excursions in this area later in the paper. The success of these investigations to date and the continued commitment to the field by these researchers provides some cause for optimism. So what techniques do we have to pursue the possibilities of prediction in computational music systems?

A common mathematical tool used to implement predictive processes into computational systems is probability or, more precisely, Bayesian statistics. The relevance of this approach to modelling human cognitive skills is well made by Baum (2004) who maintains that even beyond an effective mathematical underpinning "The Bayesian view provides ... a perspective [that is] useful for talking about probabilities and decision theory: how one should make decisions so as to maximise expected prospects" (2004:105). Application of this approach in the domain of music has been the basis for work in music perception by Eerola et al. (2002), Huron (2006), Temperley (2007) and Larson (2012) whose works were based largely on computational analysis, but did explore computational generation in a limited way. Given these building blocks it seems the construction of predictive experimentations in real-time musical metacreation software should be productive.

Later in this paper we outline our attempts so far to introduce prediction into interactive music systems, however we will first define prediction more precisely for our purposes and explore the use of prediction in existing interactive music systems more broadly.

Causal Prediction

Prediction, in general, is a very wide topic, encompassing a number of mathematical, psychological, social science and engineering disciplines—including inferential statistics, forecasting, probability theory, actuarial studies and many more. In the broadest sense, the term 'prediction' is used to describe the estimation of unknown information from known information, usually with reference to accumulated experience, observation or data.

This general definition of prediction glosses over an important characteristic, whether or not a prediction is *causal* - i.e., whether the thing being predicted has yet to occur, or has already occurred but has not yet been (or cannot be) observed. Our interest lies in *causal* prediction; the prediction of *future* events based on currently available information. Our discussions above regarding prediction's pivotal role in perception and cognition also refer to causal predictions.

The importance of clarifying whether prediction is causal or not becomes evident in in real-time interactive music systems, where there is a desire to synchronously perform complementary musical material. In this case accurate assessment of future states allows for coordinated action to be taken in a timely fashion, rather than after-the-fact. In the building of, and performing with, interactive music systems, we have found this distinction to be important.

Prediction in Existing Interactive Music Systems

Despite the strong theoretical basis for the role of prediction in human intelligence outlined above, its application to real-time computational metacreativity has, to date, been relatively sparse. For example the following systems do not, so far as we can tell, use prediction as we have described it here: *Cipher* (Rowe 1992), *The Continuator* (Pachet 2002), *Beginner's Mind* (Ciufo 2005), *Frank* (Casal 2007), *_Derivations* (Carey 2012), *Agent Designer Toolkit* (Martin and Bown 2013), *Swarm Music* (Blackwell 2007), *Self-karaoke Machines* (Eldridge 2008), *Decision Trees* (Bown 2011) and *Odessa* (Linson 2012), along with many others.

The *OMax* system of Assayag and Dubnov is described in the language of prediction, but in this case they are using the term in a non-causal way. However, they note that "one of the drawbacks of [our existing] methods is lack of responsiveness to changes in musical situations that occur during performance" and propose "an anticipatory machine improvisation system" (Cont, Assayag, and Dubnov 2006:2).

There are however some existing systems that employ causal predictive processes. There are a number of systems that utilise predictive beat-tracking, including *B-Keeper* (Robertson & Plumbley 2007) and *DrumTrack* (Collins 2006). The use of predictive beat-tracking has also been successfully employed for temporal prediction in score following systems (e.g. Dannenberg 1984). Perhaps the most sophisticated application of prediction and proactivity in an interactive music system to date—the *Shimon* system (Hoffman & Weinberg 2011)—is a robotic marimba player, with physical manifestation including four multiply-articulated 'arms' and an artificial eye. The physical limitations of the robotic arms mean the system has latency between decision and action, so that "the system also uses anticipatory action to enable real-time improvised synchronization with the human player" (2011:133). Their predictive model uses Bayesian inference to combine beat tracking and visual cues from the human player.

Towards the use of Prediction in our Work

In this section we briefly outline some of our preliminary explorations of predictive processes in real-time interactive music systems. We then outline a proposed extension of a generative melody generator and how it might employ prediction.

Beat Tracking

The efficacy of prediction in beat-tracking is attested by both perceptual theory (London 2004) and engineering practice (Collins 2006). Our *Jambot* system uses predictive beat tracking for proactive improvisation - meaning that it will play notes at predicted beats without needing the human improviser to 'confirm' this beat hypothesis by also playing a note there. The *Jambot's* beat tracking employs a multiple-parallel-hypothesis architecture, with several plausible tempo and phase estimates continuously tracked, and assigned confidence levels - details of the implementation can be found in (Gifford 2011:97-121).

Attaching confidence levels to the Jambot's beat predictions assists in managing prediction errors in live performance - by switching between proactivity and (musically safer) reactivity at a confidence threshold (Gifford 2013).

Prediction of the beat also impacts upon the *Jambot's* onset detection algorithms. Following Jones' theory of dynamic attending (Large and Jones 1999) the *Jambot* adjusts the critical signal-to-noise ratio required to register an onset according to the predictive probability of an onset occurring, given the hypothesised metric context. This notion of 'attentional energy' was operationalised as a gaussian distribution over time, centred on the predicted beat, and with peak height determined by the relative metric strength of the predicted beat; again details can be found in (Gifford 2011:118).

Extended Now

Typically predictive processes involve decisions made with incomplete data, and the earlier the predictions are made, the less complete is the available data. This is the basis for our notion of the Extended Now—where 'now' is the time that action is required, say to have a generated sound coincide with an appropriate human action. 'Now' is 'extended' into the immediate past and future in which available information can result in different levels of predictive or reflective confidence.

The practical musical applications of this need for timeliness and decision making with incomplete data include the performance of elisions - where one segment starts at the same time (often on the same pitch) that the other part plays the last note of its segment. Ideally, the elision could occur between the human and computer performers. An even more demanding requirement is for the computer to perform an anacrusis, where its material starts before the coincidental 'now'. The information available within the Extended Now changes through this period and thus so does the certainty of performance decisions. Based on this perspective our CIM (Brown, Gifford, and Voltz 2013) software makes use of an accumulating evidence process to predict human performer actions and assist in timely computer actions.

Tracking Closure

The notion of closure, or stability, plays an important role in Meyer (1973) and Narmour's (1990) theories of music perception. In particular the variation of periods of tension and release or stability and instability are suggested to provide momentum and define structure in music and thus to be important in to music's sense of emotion. Based on Narmour's theory and our own statistical corpus analysis we have developed a real-time closure measure that tracks the level of closure across several musical parameters, including intervallic movement, harmonic stability and metric context (Brown, Gifford, and Davidson 2012).

Tracking the closure level allows a system to make opportunistic decisions at moments of particular closure levels and to track and project patterns of closure that may anticipate musically salient moments, such as cadence points or segment boundaries.

Future Prospects for Melody Generation

Given the examples outlined above there are some possible next steps to extend the role of prediction in our processes. The features of closure as a predictor of melodic structure have been a part of algorithms we have developed for melody generation based on psychological theories of expectation (Gifford, Brown, and Davidson 2013). These techniques rely, in large part, on the projection of phrase continuations based on gestalt theories of perception including good continuation, proximity, and goal orientation.

To date we have used closure in an opportunistic way in these processes for segmentation even while the individual constituent elements of the closure measure have been used for assessing likely next steps as the melody proceeds. Similar processes could be used to track an improvising partner's performance and to predict the likely next steps. Based on these assessments the computer could preemptively change its behaviour to reach various compositional goals, including synchronisation, harmonisation, counter melody or rhythmic interlocking.

More concretely, our Motivator software employs these melody analysis and generation processes to create reinterpretations and variations on an existing melodic phrase. These can be used as stimuli for a computer-assisted compositional process (Gifford, Brown, and Davidson 2013). The Motivator system works, in part, by identifying structural tones within a melody and generating variations that maintain structural tones but vary the melodic material between them. At present these tones are identified through analysis or manually by the user. An extension to this system could allow the system to predict future structural tones (typically these are several beats apart) and utilise existing process to generate material between them.

The efficacy of such an approach may be difficult to evaluate in the case of Computer-Assisted Composition. As such we propose a further extension, where this melody generation is part of a real time duet interaction system, and the future structural tone estimation is informed by the recent activity in both parts. In this scenario, prediction could be evaluated based on the ability to anticipate coordination with the live human performer; including matching of harmonic progression, phrase boundaries, and so on.

Regarding beat tracking, there are, echoing Collins (2006), significant opportunities for predictive entrainment in systems beyond that already present. Given that temporality is key to our application of prediction for proactive generation, identifying regularities in pulse and tempo of a performance is critical to assisting the computer to entrain to the human interactor in many musical contexts.

Much more could also be made of the approach of continual assessment based on accumulating evidence that is a feature of what we describe as the Extended Now. It is not hard to imagine how various thresholds of evidence could be set at different points in the ongoing musical stream based on the tolerance of particular musical contexts for predictive systems to be more or less accurate. Also the combination of this approach with the maintenance of multiple musical interpretations within the computer music system, an approach we have previously described as a Chimeric Architecture (Gifford and Brown 2009), could produce multiple candidate predictions from which a weighted selection could be made.

Conclusion

In this position paper we propose that the design of interactive computer music systems can benefit from aspects of prediction. This suggestion continues our research agenda to apply insights gained through studying human musical abilities to contribute to musical metacreation.

We have outlined some of the influential literature on prediction, anticipation and expectation in music that show that these are significant aspects of human musical capacity. It seems reasonable to think that they can also be useful, therefore, in interactive computer music systems.

We have examined a number of existing interactive music systems with regard to their use of prediction, and conclude that prediction has been infrequently used to date. We go on to provide some examples of our early work into the use of some predictive processes in our own systems and reflect on ways that these could be extended.

Our experience of adding even modest predictive capabilities to interactive music systems is that these provide a substantial increase in the perception of musical agency to the system and we encourage designers of future systems to explore the benefits of similar approaches.

References

Baum, E. 2004. *What is Thought?* Cambridge, MA: MIT Press.

Blackwell, T. 2007. Swarming and music. In Miranda, E., and Biles, J. eds., *Evolutionary Computer Music*, 194–217. London: Springer.

Bown, O. 2011. Experiments in Modular Design for the Creative Composition of Live Algorithms. *Computer Music Journal*, *35*(3): 73–85.

Brown, A. R., Gifford, T., and Davidson, R. 2012. Tracking Levels of Closure in Melodies. In *Proceedings of the 12th International Conference on Music Perception and Cognition.* Thessaloniki, Greece: ISMPC.

Gifford, T., Brown, A. R., and Davidson. R. 2013. Amplifying Compositional Intelligence: Creating with a Psychogically-Inspired Generative Music System. In *Proceedings of the 2013 ICMC Conference,* Perth, Australia: ICMA

Brown, A., Gifford, T., and Voltz, B. 2013. Controlling Interactive Music Performance (CIM). In *Proceedings of the Fourth International Conference on Computational Creativity.* Sydney, Australia.

Bruner, J. 1967. *A Study of Thinking.* New York: John Wiley and Sons. (Original work published 1956).

Bruner, J. 1979. *On Knowing: Essays for the left hand*. Cambridge, Mass.: Harvard University Press.

Carey, B. 2012. Designing for Cumulative Interactivity: The _derivations System. In *Proceedings of NIME '12*. Ann Arbor, Michigan: University of Michigan.

Casal, D., P. and Morelli, D. 2007. Remembering the Future: An Overview of Co-evolution in Musical Improvisation. In *Proceedings of the International Computer Music Conference*, Copenhagen: ICMA.

Ciufo, T. 2005. Beginner's Mind: an Environment for Sonic Improvisation. In *Proceedings of the International Computer Music Conference*, 781-784. Barcelona: ICMA.

Collins, N. M. 2006. Towards Autonomous Agents for Live Computer Music: Real-time Machine Listening and Interactive Music Systems. Ph.D. diss. Faculty of Music, University of Cambridge.

Cont, A, Dubnov, S. and Assayag, G. 2006. A Framework for Anticipatory Machine Improvisation and Style Imitation. *Proceedings of Anticipatory Behavior in Adaptive Learning Systems.*

Cont, A., Dubnov, S. and Assayag, G. 2007. Anticipatory Model of Musical Style Imitation using Collaborative and Competitive Reinforcement Learning. In *Anticipatory Behavior in Adaptive Learning Systems.* Springer.

Dannenberg, R. B. 1984. An On-Line Algorithm for Real-Time Accompaniment. In *Proceedings of the 1984 International Computer Music Conference*, 193-8. Paris: ICMA.

Eerola, T., Toiviainen, P., and Krumhansl, C. L. 2002. Real-Time Prediction of Melodies: Continuous Predictability Judgements and Dynamic Models. In *Proceedings of the 7th International Conference on Music Perception and Cognition*, 473–476, Australia: Causal Productions.

Eldridge, A. 2008. Collaborating with the Behaving Machine: Simple Adaptive Dynamical Systems for Generative and Interactive Music, D. Phil. diss. University of Sussex, Sussex.

Gifford, T. 2011. Improvisation in Interactive Music Systems. Ph.D. diss., Faculty of Creative Industries, Queensland University of Technology, Brisbane, Australia.

Gifford, T. 2013. Appropriate and Complementary Rhythmic Improvisation in an Interactive Music System. In *Music and Human-Computer Interaction,* 271-86, Springer Series in Cultural Computing 1. London: Springer.

Gifford, T., and Brown, A. R. 2009. Do Androids Dream of Electric Chimera? In *Improvise: The Australasian Computer Music Conference,* 56–63. Brisbane, Australia: ACMC.

Gifford, T., Brown, A. R., and Davidson, R. 2013. Amplifying Compositional Intelligence: Creating with a Psychologically-Inspired Generative Music System. In *Proceedings of the ICMC Conference,* 357–360. Perth, Australia: ICMA.

Harvey, Andrew C. 1990. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Hawkins, J., and Blakeslee, S. 2004. *On Intelligence*. New York: Times Books.

Hawkins, J., and George, D. 2006. *Hierarchical Temporal Memory: Concepts, Theory, and Terminology*. Numenta Inc.

Hoffman, G. and Weinberg, G. 2011. Interactive Improvisation with Robotic Marimba Player. *Autonomous Robots* 31(2):133-53.

Hofstadter, D. R. 1995. *Fluid Concepts and creative analogies: Computer models of the fundamental mechanisms of thought.* London: Allen Lane, The Penguin Press.

Hofstadter, D., Sanders, E. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.

Huron, D. 2006. *Sweet Anticipation: Music and the psychology of expectation*. Cambridge, Mass.:MIT Press.

Larson, S. 2012. *Musical Forces: Motion, Metaphor and Meaning in Music*. Bloomington: Indiana University Press.

Linson, A., Dobby, C., and Laney, R. 2012. Improvisation without Representation: Artificial Intelligence and Music. In *Music, Mind, and Invention Workshop: Creativity at the Intersection of Music and Computation*, 30–31. Ewing, New Jersey.

London, J. 2004. *Hearing in Time.* Oxford: Oxford Uni. Press.

Martin, A., and Bown, O. 2013. The Agent Designer Toolkit. In *Proceedings of Creativity and Cognition 2013*. Sydney: ACM.

Meyer, L.B. 1973. *Explaining Music: Essays and Explorations*. Berkeley Calif.: University of California Press.

Narmour, E. 1990. *The Analysis and Cognition of Basic Melodic Structures*. Chicago: University of Chicago Press.

Pachet, F. 2002. The Continuator: Musical Interaction with Style. In *International Computer Music Conference*, 211–218. Göteborg, Sweden: ICMA.

Robertson, A. and Plumbley, M. 2007. B-Keeper: A Beat-Tracker for Live Performance. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression.*

Rohrmeier, M. A. and Koelsch, S. 2012. Predictive information processing in music cognition: A Critical Review. *International Journal of Psychophysiology* 83: 164-175.

Rowe, R. 1992. Machine Listening and Composing with Cypher. *Computer Music Journal*, *16*(1): 43–63.

Temperley, D. 2007. *Music and Probability*. Cambridge, Mass.: MIT Press.