Musical Metacreation: Papers from the 2012 AIIDE Workshop
AAAI Technical Report WS-12-16

Autonomy in Music-Generating Systems

Oliver Bown
Design Lab,
University of Sydney, NSW, 2006, Australia
oliver.bown@sydney.edu.au

Aengus Martin
Faculty of Engineering,
University of Sydney, NSW, 2006, Australia
aengus.martin@sydney.edu.au

Abstract

The word 'autonomy' is often used in the discussion of software-based music-generating systems. Whilst the term conveys a very clear concept – the sense of self-determination of a system – attempts to formalise autonomy are at an early stage, and the term is subject to a range of interpretations when practically applied. We consider how the evaluation of music-generating systems will be enhanced by a clearer understanding of autonomy and its application to music. We discuss existing definitions and approaches to quantifying autonomy and consider, through a series of examples, the information that is required in order to make precise formal judgements about autonomy, and the identification of relevant levels at which the principle of autonomy applies in music. We conclude that automated measures can supplement human evaluation of autonomy, but that (a) automated measures must be supported by sound reasoning about the features and timescales used in the measurement, and (b) they are improved by a having knowledge of the internal working of the system, rather than taking a black box approach. We consider multi-dimensional representations of system behaviour that may capture a richer sense of the notion of autonomy. Finally, we propose an approach to automatically probing music systems as a means of determining an autonomy 'portrait'.

Introduction

The word 'autonomy' is often used in the discussion of software-based music-generating systems to capture the sense in which we desire such systems to create musical output *on their own*. However, little discussion has been had about what exactly autonomy means in the context of music creation, how it might be measured, and whether it is indeed a central requirement of music-generating systems, or even a desirable property at all. In this paper we discuss the basic issues involved in quantifying autonomy, leading to an approach to measuring autonomy in music-generating systems, and interpreting and reasoning about such results. Our consideration of autonomy in various abstract examples is presented alongside discussion of how we attribute autonomy in human music creation. We propose that mathematical measures of autonomy are meaningful only when supported by reasoning about the choice of observed variables and timescales used, that systems have to be actively probed in order to gain an accurate representation of their autonomy, and that whilst autonomy is an important quality, it is not a quantity that ought to be maximised. Although we look at specific statistical procedures we do not, in this paper, propose an actual applicable measure of autonomy in music-generating systems.

Defining Autonomy

Autonomy refers to the self-determination of a system, leading to a number of different formal definitions. Seth (2010), following a dynamical systems point of view, considers autonomy in terms of a system's influence on its own future, as compared to external influences on that future. A system can be called autonomous if its own state history contributes to a better prediction of its future when combined with the state history of any external factors. Alternatively, a system can be called autonomous in light of its ability to take care of itself, that is, to control various internal and external elements in order to maintain certain variables within acceptable bounds (Ashby 1960). Alternative perspectives on autonomy include goal-directedness, cognitive states and adaptiveness. For example, Luck and d'Inverno (1995) define an agent as an "instantiation of an object together with an associated goal or set of goals" (Luck and d'Inverno 1995, p.265) and an autonomous agent as an agent with a set of motivations that lead to the generation of goals: i.e., the agent's goals are self-determined.

Variations such as Luck and d'Inverno's can be seen as more specific applications of the notion of self-determination.

A weaker but perhaps more common notion of autonomy in software systems is that of a system that is self-correcting, given a predefined function. A driverless car, in this respect, is understood as being autonomous despite being controllable and highly responsive to external factors, and having no such thing as a 'motivation'.

In this paper we adopt the notion of autonomy as 'self-determination', understood in terms of the relationships between the state of a given system and the state of any external systems that may have an influence on it. However, by looking at systems at different time-scales and using different measurable features we will attempt to integrate as far as

Copyright © 2012, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

possible the range of interpretations found in the literature.

Quantifying Autonomy

Two methods for calculating autonomy as self-determination are considered by Seth and his colleagues (Seth 2010; Barnett, Barrett, and Seth 2009). Both are derived themselves from measures of causality. The first, Granger causality (Granger 1969) models a time-series using regression. Any deviation of the time series from the model is tested to see if it caused by another time series. The extent to which the second time-series improves the model of the first determines the Granger causality of the one time series by the other. Seth (2010) defines G-autonomy as the extent to which a series' self-determination is significant, that is, whether the system is better modelled by including the system's own history in its set of causal factors. Seth (2010) shows that G-autonomy measures for well-known elementary models from artificial life, an evolved predator-prey model, and a flocking model, correspond well to our intuition.

An alternative measure of causality is Transfer Entropy (TE) (Schreiber 2000). Entropy is a measure of uncertainty in a system. TE determines the extent to which a system's uncertainty is reduced by observations of another system. Alternatively, this can be understood as the information provided by one system about the future state of another system. Both G-causality and TE are limited by our ability to model any given system with their respective approaches: regression and statistical analysis. As we shall discuss, since the temporal modelling of music is significantly more complex than that of simple artificial life models, the application of either method is highly dependent on choices of features, temporal resolution and other factors.

Two reasons for considering TE over G-causality are that (i) statistical models of music have proven to be successful in recent work (Pearce 2011), and (ii) statistical models also give us other information pertaining to the complexity of a system.

From Schrieber (2000), the Transfer Entropy from a system J to a system I is given by:

$$T_{JI} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) log \frac{p(i_{n+1}|i_n^{(k)}, j_n^{(l)})}{p(i_{n+1}|i_n^{(k)})} \tag{1}$$

where i_n is the state of I at time n and $i_n^{(k)}$ is the combined state of I for all time steps from n back to $(n-k+1)$, i.e., the state over the window $(i_n, i_{n-1}, \ldots, i_{n-k+1})$.

TE gives us a measure which is higher if J has a clearer influence on I. If we score a low TE, this may either be because I is indeed not influenced by J or that there is an influence but it is not apparent. The latter could be because we are looking at the wrong features, or the wrong temporal resolution (for example if we have no way of telling that one event happened before another we cannot expect to identify causality), but given that the measure is based on the predictability of I, it could also be because we can obtain a perfectly good model of I even though it is influenced by J. This can happen if J is very predictable and I is directly determined by J. Equally, if I is extremely unpredictable, then neither the history of I nor the history of J will provide information about its future state and TE will also be low.

Bertschinger et al. (2008) provide discussion of a number of similar information theoretic definitions of autonomy, largely based on the mutual information between two processes, similar to TE. Their definitions take into account the related factors non-heteronomy, self-determination, closure, and causality and form the basis for a rich set of possibly useful measures. Since we do not attempt to apply autonomy measures to music performance in this paper we will not cover these specifics. However, a number of the concepts brought up by Bertschinger et al. appear in our discussion.

As we have said, before being useful these information theoretic definitions require significant decisions to be made about what data is extracted from specific interaction scenarios. The following section discusses such musical scenarios.

Autonomy in Musicians and Music Software

In a range of research initiatives, software is tasked with the production of musical output. These various projects span a range of goals and perspectives with respect to autonomy. Three reasonably clear distinctions can be made between the goals of such projects: (i) systems that are tasked with the production of final musical outputs, versus systems that are used as idea-generators, from which a music creator will choose output; (ii) systems that generate variation but do not employ any selection process versus systems that engage some form of selection process; (iii) systems that compose music offline versus systems that perform live with other (usually human) musicians.

In this paper we will consider the notion of autonomy across this range of scenarios, but with particular attention paid to the last of these distinctions, for which the notion of autonomy can be seen to apply at two different levels.

Performance Autonomy

The first author's research is concerned with Live Algorithms (Blackwell and Young 2004; Blackwell, Bown, and Young 2012). These are are systems that perform interactively with other musicians, rather than composing offline, that generate final outputs, rather than sample material for selection, but that may or may not employ selection processes in determining their output.

Blackwell's definition of an autonomous musical agent in (Blackwell, Bown, and Young 2012), which should be understood as an aspirational goal of Live Algorithms research rather than a description of existing Live Algorithms, is an agent which goes beyond mere automation in determining responses to musical input, by deriving novel but relevant responses. The novelty requirement, if fulfilled, moves the origin of the system's behavioural repertoire from the system's design to the system itself.

We can examine the essence of the term autonomy by comparing two contexts, driverless cars and improvising musical agents, using Seth's minimal formalisation of self-determination, and the mathematical notion of influence given in TE. This will illustrate how the contextual elements that we take into account, including the choice of timescales and appropriate feature selection, are highly significant in the application of the term.

We can say that a driverless car is autonomous compared to a regular car, with respect to the driver. This can be quantified by examining the relationship between two time-series: the state of the car and the state of the driver. If the driver time-series is a good predictor of the car-time series then this is evidence that the car lacks autonomy, whereas if the car time-series makes a significant contribution to predicting its own future states when added to the driver time-series, then this is evidence of self-determination and therefore autonomy. When the driver of a regular car turns the steering wheel or hits the brakes, he or she clearly influences the state of the car and as such the driver's actions could be used to faithfully predict the car's response. The 'driver' of a driverless car may be eating lunch or playing chess while the car turns, accelerates and brakes. In this case the driver's actions will not be good predictors of the car's future state. Rather, with respect to the driver, at least, the car's own past state is a better predictor of its own future state and the car can therefore be considered autonomous (see Figure 1).

Compare this to an automatic hand drier (i.e., one which senses the presence of your hands rather than relies on a push-button). We know that the sensor is simply a drop-in replacement for the button, serving exactly the same operational function, and yet the argument might be made that the automatic hand drier is behaving autonomously, observing its environment and taking appropriate actions, unlike its passive button-operated counterpart. From the perspective of predictability and causality, however, this distinction is irrelevant.

Whether actual time-series analyses corroborate such determinations of causality will depend on a number of factors; the points in time the system is being observed, the temporal resolution, the measured values and the system configurations. For example, the 'driver' of the driverless car punches in coordinates, and these are a good predictor of the future state of the car. At a coarse time-scale, taking into account this input information, the car is no longer autonomous. Equally, naïve measurements of the state of the driver of a regular car, his or her foot position, for example, may be too coarse grained for us to identify the variables that influence the car. In such cases, although we suspect that the driver controls the car, we would lack the data to verify this.

The same analysis can be applied to an improvising musical agent. Again, the autonomy can be discussed in terms of its relationship to another entity, for example a human musician, and in comparison to a suitably comparable entity, such as an electronic musical instrument or a tape-player, where an equivalence between interaction contexts can be established. The state of the electronic musical instrument, like the regular car, could be successfully predicted by information pertaining to the state of its controller, a human musician. A musical agent, by comparison, would be less predictable on the basis of information pertaining to a human improvising partner, even though we should expect the system to respond in some manner to the human, and therefore

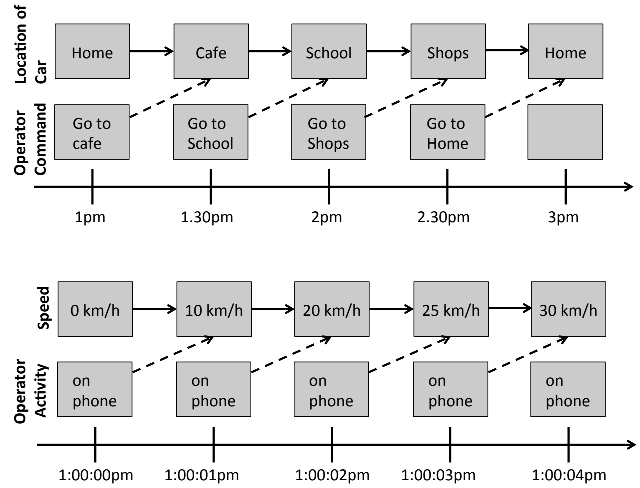

Figure 1: Different sampling timescales and feature measurements pertaining to the causal relationships between driver and 'autonomous' car. Top: the driver's commands are a good predictor of the car's future location. Bottom: the driver's activity state during a journey is not a good predictor of the car's speed.

in some sense be influenced (or controlled or manipulated). A tape player playing back pre-recorded music, meanwhile, would be largely unaffected by any musician playing along to it. This poses an affront to our instinctive understanding of autonomy, to which there are two answers.

Firstly, our intuition is that the tape player is not autonomous because the tape player is entirely predictable, and we can quantifiably show this if we play back the tape enough times in order to build an accurate model of its output. If a sufficiently high-order statistical model of the tape-player is made it would demonstrate low entropy. Not only is the tape-player's own history a better predictor of its own future output than any other entity's history is of it, but it is a perfect predictor. Our notion of autonomy as self-determination could be modified to incorporate the caveat that a completely predictable system is not autonomous.

Bertschinger et al. (2008) extend this condition, in principle, to significantly more complex systems that include adaptive behaviour: "if one knows what is optimal for the system in a certain environment, one might predict the action, behaviour or state of the system from the environment... Thus in case of an adaptive system we require that it is capable of pursuing different objectives in the same environment" (Bertschinger et al. 2008, p.334). However, since the content of a tape might be enormous and hugely diverse, we cannot be sure that 'cheat' systems will be more predictable than 'genuine' systems, even if we impose this requirement.

The alternative, as discussed with cars, is to choose a different time-scale and involve different variables that reflect our intuition that the tape-player is not autonomous. Viewed from the point of view of an operator we could argue that the audio signal that it outputs is irrelevant and that its essential operational variables are (i) whether or not it is playing, (ii) which tape is loaded, (iii) the volume setting, and so on. Such variables are equally good at describing the audio output of the system, but now reveal instead a lack of autonomy of the tape-player, which does not choose what tapes it plays, set itself playing, or control its own volume. This question of different frames with which to view autonomy is also discussed by Bertschinger et al. (2008) with reference to a study of gliders in the Game of Life. Whilst gliders demonstrate measurable autonomy in their movement through open space, through their being critical self-predictors of their own future state, most interactions between gliders and other objects result in the death of the glider. Since the non-existence of the glider cannot be handled by the particular frame for measuring glider autonomy (readings of the state of the glider) the problem is conspicuous.

Creative Autonomy

We can also view the Live Algorithm in terms of a nested hierarchy of goals and their related actions, and consider autonomy with respect to each goal. Here the reasoning applies equally to systems that compose music offline. Like the driverless car, but unlike humans and animals, the system can be seen as fulfilling certain goals that are actually the goals of its designer. At this top level, the system does not derive its own goals from essential motivations and so does not qualify for Luck and d'Inverno's definition of an autonomous agent. However, an advanced Live Algorithm might fulfil an ultimate musical objective by, for example, deriving its own 'conceptual spaces' (Gärdenfors 2004) within which it identifies subgoals that ultimately satisfy the programmer's motives.

Whilst we can therefore build systems that respond to this advanced notion of autonomy, there is a very real danger of setting up hoops to jump through that do not lead to any actual technical progress. In a computer program which dynamically constructs complex data structures, for example, we may choose to refer to various objects or functions as goals and motivations, and may arbitrarily divide systems into subsystems so that certain of these subsystems score on a measure of autonomy with respect to others. As we have seen, obtaining a high quantitative measure of autonomy may prove to be trivial. Measures of creative autonomy should therefore be backed up with reasonable arguments about the relevance of the contextual factors presented.

We can also consider autonomy in terms of musical influences over longer timescales. In this respect autonomy may capture the notion of originality in creative work. Using the same notion of predictability, if an existing body of music is a good predictor of the output of a new composer then, as with the earlier examples, we would declare that the composer lacks autonomy. But as has been widely discussed in the literature of creativity, e.g., (Boden 1990), too much novelty implies low creativity. We expect music to be original but to clearly place itself in the context of existing music, thus be partially predicted by it.

Human Evaluation

As with the limitations of data capture and complexity that apply to mathematical methods, humans make judgements about qualities such as autonomy based on limited information. A source of inspiration for judging the quality of a software system that has been designed to create original music, or engage in live improvised performance with humans, is the Turing Test (Turing 1950). Alan Turing's famous test, although a directly usable format for testing machine intelligence, was philosophically motivated, providing the important clarification that there cannot be any objective basis for believing that human intelligence has an essence that is unobtainable for machines.

The same sentiment can be applied to music. However, as Ariza (2009) explains, "musical Turing Tests do not actually conform to Turing's model" (Ariza 2009, p. 49). "Music, as a medium remote from natural language, is a poor vessel for Turing's Imitation Game" (Ariza 2009, p. 66). In conversation, we can interrogate our subject in order to understand their thinking. Language is the special human faculty that allows us to accurately convey meaning and makes this possible. Music, as far as we know, does not share this property. Cross (2008), for example, argues for music's ambiguity and "floating intentionality" as essential to its particular social function. Music may express emotion, but a convincing format for interrogating a system on emotional grounds or via emotion has yet to be proposed. As well as the weakness of the medium, Turing's test is particularly poorly applied, as occasionally happens, when the 'interrogation' consists of listening to a piece of music and trying to decide whether it was composed by a computer based only on the content of the music. A recent example of this is given in *The Guardian* newspaper's 'Turing Test' for the Iamus music system[1], for which no additional information can be found about whether, for example, the system was primed with human-composed primitive elements. Although, in this and other existing cases, we may be able to tell the system apart from human composers due to something not sounding quite right, instances where we were unable tell the system apart from human composers would not be evidence of human-like intelligence, due to this limitation of music as a medium.

Human-like autonomy is implicated both as the element a judge is interested in detecting in such tests, and the element that eludes us. We are interested in determining autonomy because it is a key component of 'creative composition', as discussed in the previous section. However, autonomy eludes detection through interaction because of (a) a lack of information with which to make a judgement, particularly in crude listener tests, and (b) a lack of expertise at detecting autonomy in novel systems: it is not actually something we do, since we take for granted that other humans are autonomous and have a template for how that autonomy is manifest. This implies that judging systems requires at least a process of interrogation akin to the original Turing Test, and that automated support tools for measuring autonomy

[1] http://www.guardian.co.uk/science/poll/2012/jul/01/musical-turing-test-audio-clip-computer, accessed July 15th 2012.

may supplement human judgements, along with a technical or at least rough conceptual understanding of the systems being judged. Thus as knowledge of a person's life history may affect our judgement of their musical output, so might knowledge of the working of a software system.

Other human evaluation methods of interactive software systems have been investigated. One example is the work of Hsu and Sosnick (2009) who devised a system of questionnaires for both improvisers and audience members with a view to evaluating and comparing two of their systems. There were specific questions relating to the responsiveness of the system as well as more high-level ones regarding the performance experience as a whole.

Collins (2006) used interviews with improvising musicians, based on the contextual inquiry model, to evaluate a selection of interactive software systems. The main results reported concerned the subjective experiences of the musicians (e.g. comfort, stimulation) and the extent to which they attributed agency to the system. Such tests can be combined with objective measures of autonomy, and a critical understanding of the operation of the system, to build a better understanding of when and why humans perceive creative agency.

Conclusion

Autonomy is clearly an important and desirable property in software that creates music, being essential to the notion of 'creation'. But, defined as 'self-determination', the autonomy that we seek in autonomous music systems is not something that should be maximised to the point of freedom from influence. Other properties will be useful to complement autonomy: agency, complexity and intelligence. Furthermore, our discussion also suggests that a singular measure of autonomy itself may not actually be as interesting as the various measures that can be considered to constitute it, such as TE and other factors discussed by Bertschinger et al. (2008): non-heteronomy and closure. For example, Figure 2 shows a proposed sketch of what various systems may look like on a plot showing the combined entropy of input, X, and response, Y (x-axis), and transfer entropy from X to Y (y-axis). Different systems appear as loci, their measured position on the plot varying according to what input they are fed (see figure caption).

Such multi-dimensional representations may underlie a richer notion of autonomy. In such cases we are essentially probing one software system with another in order to obtain a better sense of the system's autonomy in a range of contexts. This may be useful in categorising systems and searching for interesting forms of behaviour. For example, in Figure 2, a range of inputs may be better at distinguishing a smart musical agent from a dumb instrument, where clearly a single measurement will not offer much clarity. Bertschinger et al. (2008) propose 'interventions' in the interaction between a system and its environment that would provide more precise information about the causal and self-determining behaviour of the system than could be understood by simple observation. In our view, the probe system could generate input patterns to pass to the subject system, and would analyse the results to determine their influence.

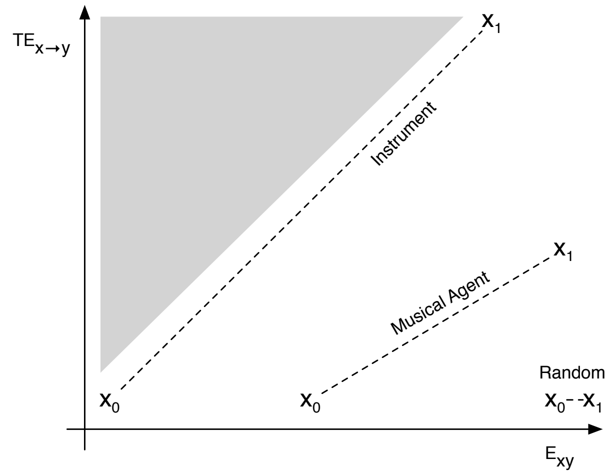

Figure 2: A sketch of possible relationships between overall entropy (E) of the combined states of the XY-system (x-axis) and transfer entropy (TE) from X to Y (y-axis) for different systems under a range of input conditions, where X is the input to the system (e.g., provided by the experimenter) and Y is the output of the system. Each dotted line represents a possible locus within this E-TE space occupied by a given system. Each locus is a line bounded by two extremes, low entropy, X_0 and high entropy, X_1 inputs. A 'dumb' reactive system, such as a musical instrument, demonstrates high TE, but this is invisible under low input entropy conditions. A completely random system never exhibits TE, i.e., external influence. A musical agent may contribute entropy to a system despite low entropy in the input and would only exhibit moderate external causality (TE). The grey triangle is not accessible since systems with low combined entropy will produce low TE.

The probe system could be considered as a simulation environment in which the subject system is analysed. Assuming the target system could be operated precisely by the probe (run it in the same thread, reset it to specific states, and assume control of any random number generation built into the system) interventions could include comparing different inputs in identical scenarios, or the same inputs applied to different internal states in the system. The result would be a portrait of a system rather than a specific measure. Figure 2 is an example of what such a portrait may look like, although the relationship between system entropy and transfer entropy is just one proposed representation and is not likely to be the most interesting.

Our discussion has also elaborated upon two major caveats to consider when attempting to measure autonomy in music-generating systems.

Most importantly, the thing being measured and the features and time-scales that are used in the measurements can lead to inappropriate measures for autonomy, both too high and too low. Ultimately, sound reasoning needs to be used in deciding what time-scales and features constitute a good measure of autonomy, as with the example of the tape-player. With the limits of current music analysis, this is the main reason why measuring autonomy is not presently feasible. In the long term we expect that criteria for interpreting autonomy relevant to different musical contexts can be defined. For example, Blackwell's (2012) elaboration of autonomy in Live Algorithms provides important detail about the relevant context, although it remains a challenge to formalise the features necessary for measuring this.

Secondly, in order to make informed judgements about autonomy we may need to know about the design of systems and not just view them as black boxes. With humans, there is the reasonable and innate assumption that others are like ourselves (i.e., autonomous), but our discussion of human evaluation outlines why we cannot simply use human-ness as a benchmark for the measure of machine autonomy. Indeed, we may have more success if we turn this on its head and discover examples of machine autonomy that can act as benchmarks for understanding human cultural behaviour.

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.

Ashby, W. R. 1960. *Design for a Brain*. John Wiley & Sons. inc.

Barnett, L.; Barrett, A. B.; and Seth, A. K. 2009. Granger causality and transfer entropy are equivalent for gaussian variables. Technical Report 87.10.Mn, 87.19.L, 87.19.lj, 87.19.lo, 89.70.Cf, University of Sussex.

Bertschinger, N.; Olbrich, E.; Ay, N.; and Jost, J. 2008. Autonomy: an information theoretic perspective. *Biosystems* 91(2):331–345.

Blackwell, T., and Young, M. 2004. Self-organised music. *Organised Sound* 9(2):137–150.

Blackwell, T.; Bown, O.; and Young, M. 2012. Live algorithms. In McCormack, J., and d'Inverno, M., eds., *Computers and Creativity*. Springer-Verlag.

Boden, M. 1990. *The Creative Mind*. George Weidenfeld and Nicholson Ltd.

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.

Cross, I., and Woodruff, G. E. 2008. Music as a communicative medium. In Botha, R., and Knight, C., eds., *The Prehistory of Language*. OUP.

Gärdenfors, P. 2004. *Conceptual Spaces: the geometry of thought*. MIT Press.

Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3):424–439.

Hsu, B., and Sosnick, M. 2009. Evaluating interactive music systems: An hci approach. In *Proceedings of New Interfaces for Musical Expression (NIME).*

Huron, D. 2006. *Sweet Anticipation*. MIT Press.

Luck, M., and d'Inverno, M. 1995. A formal framework for agency and automation. In *Proceedings of the First International Conference on Multiagent Systems*.

Malafouris, L. 2007. At the potter's wheel: An argument for material agency.

Pearce, M. T., and Wiggins, G. A. 2004. Rethinking gestalt influences on melodic expectancy. In Lipscomb, S. D.; Ashley, R.; Gjerdingen, R. O.; and Webster, P., eds., *Proceedings of the 8th International Conference on Music Perception and Cognition*, 367–371.

Pearce, M. 2011. Time-series analysis of music: Perceptual and information dynamics. *Empirical Musciology Review* 6(2):125–130.

Schreiber, T. 2000. Measuring information transfer. *Physical Review Letters* 85(2):461–464.

Seth, A. 2010. Measuring autonomy and emergence via granger causality. *Artificial Life* 16:179–196.

Turing, A. 1950. Computing machinery and intelligence. *Mind* 59(236):433–460.