A Cognitive Mechanism for Spontaneous Musical Creativity

Geraint A. Wiggins
Centre for Digital Music
School of Electronic Engineering and Computer Science
Queen Mary University of London
London E1 4NS, United Kingdom
Email: geraint.wiggins@eecs.qmul.ac.uk

***Abstract*—Creativity is an important aspect of humanity, but is shrouded in mystery and mysticism. Only recently have scientific approaches in AI focused on the study of creativity *per se*, as opposed to simpler problem solving [1]. I propose a model of creativity that merges Minsky's Society of Mind [2] with Baars' Global Workspace Theory [3]. In particular, the model simulates the experience of spontaneous inspiration.**

I. Introduction

I propose a cognitive mechanism for regulating non-conscious attention and attentive access to non-conscious information in a Choral Society of Mind [4], [2] formed around a Global Workspace [3]. It involves an anticipatory approach to cognitive process that enables new, useful distinctions to be made between different classes of cognitive information, that allow specification of a hypothetical mechanism to account for spontaneous non-conscious creativity, or "inspiration". I begin by noting the value of music in the study of creativity, and clarifying what kind of creativity I aim to explicate. Next, I present background: the current work draws together several disparate areas of research, on perception, consciousness, and computational creativity. In the central section, I analyse the affordances of the proposed perception simulation with respect to information perceived by a listener, and show how they categorise various aspects of information salience. Finally, I suggest how the whole may regulate attention.

II. Music and the study of creative cognition

My interest is in *musical* perception and creativity because music holds a unique position among creative practice: 1) it cannot be stopped and looked at in detail, so a) is heavily reliant on memory, and b) is, therefore, *entirely* subjective; and 2) it has no denotation, beyond self-reference and onomatopoeia. So music is a *self-contained* cognitive construct, and can be studied as such. This is not to deny musical embodiment; embodiment is in close relation with mind. Music is universally human, without apparent bio-evolutionary benefit [5]: studying it is likely to elucidate the human condition.

III. Two kinds of creativity

I propose an account of the creativity sometimes called "inspiration", that happens spontaneously, without conscious reasoning. It occurs in everyday language [6], and in Mozart's composition, according to his introspection [7]. It differs from creativity performed to order, e.g., when a student harmonises a given melody, which is conscious, planned, problem-solving. These two end-points form a spectrum between conscious creation in the planning of a formalist composer, the spontaneous but partly planned cooperation of the jazz trio, and the spontaneous whistling in the street of a postman. Non-polar positions on this spectrum entail a *combination*: there is no smooth transition between the extremes, but a mixture containing both explicit problem-solving and implicit inspiration.

IV. Background: ideas behind the proposed model

The Choral Society of Mind: I present a model of musical creativity that is like an improvising chorus, "singing" fragmentary ideas, with singers paying varying attention to each other at different times. Singers pick up and develop lines, or sing in unison to attract other singers to music they select. This is built over a model of musical listening that matches Minsky's "society of mind" agent model [2]. Minsky's model has agents with particular capabilities, which are learned, and so does this one. Some of the agents are specialised to learning and predicting sequences of percepts in time. The motivation for this is evolutionary: organisms that can estimate the immediate future have better survival chances than those which cannot. Similarly, organisms that are sensitive to change [4], and that can detect new differences in their environment are more likely to survive than those that cannot.

Statistical modelling of (musical) expectation: Animal learning includes more than static association [8]. An organism needs a mechanism for avoiding danger, in terms of *sequential* data; e.g., if I eat a plant, and I *later* get sick, I learn the sequential association and back-chain on the contrapositive: if I do not eat the plant, I will not be sick. But this simple mechanism is not subtle enough [9]: evolution entails that an organism breed, but if it dies, it neither learns, nor transmits its learning capacity. An effective strategy lies at the meta-level: if an organism knows that it is in unpredictable circumstances, it can be cautious, prepare for flight, and attend more to its surroundings. Huron [9] argues that this process is exapted into the aesthetics of music; some empirical work supports the claim [10]. Self-evidently, there is a mechanism to allow uncertainty to affect behaviour in humans and other animals, that precedes explicit reasoning: we feel nervous in uncertain situations, and the feeling serves to heighten our attention to appropriate sensory inputs and to prepare for flight.

Learning must include *generalisation*: tension alone cannot lead to fear at the bared fangs of a previously-unknown animal. This accords with proposals [11] that perceptual learning systems evolved to detect and quantify similarities and differences between perceived entities in the world, placing new observations appropriately between previously-experienced referents.

The modelling proposed here is statistical [12], [13], [14], but I do not claim that the mind/brain works exactly thus; rather, this abstract level provides a strong theoretical framework with appropriate mathematical background, in information theory [15]. Specifically, Pearce [13] has built a model of melodic listening that has been shown to model human musical expectations very well [16]. It can be used, via naïve sampling, to create minimally acceptable music [17]. A key feature of the model is *entropic weighting* between models of multiple musical features [18], where predictions carrying more information are favoured over ones carrying less. The model reliably predicts neurophysicoloical effects which are consciously reportable [19]. Information content estimates from the model predict subjective musical salience [10].

The current model envisages very many such agents, contributing to and sampling over a highly-structured, multi-dimensional, learned model of the agents' musical culture. The outputs of sampling constitute the new ideas that arise from "inspiration". This begs the question of how the agents interact, and how their outputs are selected for presentation to consciousness, which is the topic of the next section.

Global Workspace Theory: Bernard Baars [3] introduces a theory of consciousness called Global Workspace Theory (GWT). There is not space here to describe this wide-ranging theory in full, so I summarise the relevant important points. The theory posits a framework within which consciousness can take place, based around a multi-agent architecture [2] communicating via something like an AI blackboard system [20], but with particular constraints, which I outline below.

Baars models the non-conscious mind as a collection of expert generators, like Minsky's agents, performing tasks by applying algorithms to data in massive-parallel, and *competing for access* to a Global Workspace via which (and only via which) information is exchanged—this aspect is different from the Society of Mind. Information must cross a notional threshold of importance before it is allowed access to the Global Workspace. The Workspace is always visible to all generators, and contains the information of which the organism is conscious at any given time. However, it is capable of containing only one "thing" at a time, though what that "thing" might be is variable. The Global Workspace is highly contextualised; meaning contained therein is context-sensitive and structured; and such contexts can contain goals, desires, etc. Aside from further discussion of the "threshold" idea, below, this is all that is needed to understand the purpose of the competition mechanism proposed here. Baars mentions the possibility of creativity within this framework in passing, effectively equating entry of a generator's output into consciousness with the "Aha!" moment [21]. However, he does not develop this idea, beyond noting that a process of refinement may be implemented as cycling of information into the Workspace and out again. To my knowledge, it has not been addressed elsewhere in the related literature.

Baars proposes that information integration may take place in stages, via what one might (but he does not) call local workspaces, integrating information step by step in a sequence, rather than all in one go as it arrives in the Global Workspace. This information integration approach has been extended by Tononi and Edelman [22], who propose information-theoretic measures of information integration as a measure of consciousness of an information-processing mechanism. Baars has embraced the information-theoretic stance, too, and these authors have jointly proposed that it is time to begin implementing a conscious machine based on their ideas [23].

The Threshold Paradox: Baars [3, pp. 98–99] addresses what he acknowledges is a problem for his theory. He posits a *threshold* for input access to the Global Workspace, crossing of which is thought of in terms of recruiting sufficient generators to produce information that is somehow coordinated, or synchronised between them: it must be (metaphorically) "loud" enough to be "audible" in the Workspace. However, in terms of the Global Workspace alone, there is no means of doing this: generators can only be synchronised (whatever that means) via the Global Workspace, and so they are faced with the beginning artist's dilemma: you must be famous to show your work, but you must show your work to become famous. Baars presents two possible solutions to the paradox, but both are presented somewhat half-heartedly, leaving a gap in the theory. I will propose a mechanism to fill this gap, below.

Models of Creativity: Creativity studies have considered numerous approaches, sociological, affective and cognitive [21], [24], [25], [26], [27], [28]. There have also been attempts to formalise [29], [30] and quantify [31] creativity. As mentioned above, Baars [3, §6.2.4] quoting Mozart, outlines a high-level creative loop in GWT. The mechanism described here is a substantially more detailed model, than any of the above, to the degree that it can be implemented as a computer program. As already mentioned, it addresses the Threshold Paradox of GWT [3], and forms the boundary between the *incubation* and *insight* phases of Wallas' model [21]. Guilford's notions of *divergent* vs. *convergent production* are implicit, arising from the probabilistic nature of the model, though not in phases as he suggests [25]. The *flow* model of creativity [26] seems to relate in that it proposes states that create the conditions necessary for the model to function. Boden's [27], [28] general framework accommodates the proposal, also.

V. A MODEL OF SPONTANEOUS MUSICAL CREATIVITY

I propose a model consisting of multiple statistical sampling agents, operating over a learned perceptual model of music [13]. At each point in *listening*, agents compute a) the likelihood of each note heard, in terms of the preceding sequence and the learned model; b) the likelihood of each possible next note; and c) the *entropy* [15], [32] of the distribution resulting from b. From a and b, we calculate the *information content* [32] of the respective notes.

The model of *creativity* reverses this. The central idea is that agent-generators *continually* generate fragments of musical information from their learned models, anticipating incoming stimuli (musical or otherwise). This is motivated by the evolutionary need to anticipate the world: an organism that anticipates steals the march on one that cannot. So my musical creativity model is a specialisation of a general survival mechanism. Note that the agent-generators' models are constructed so that they can generate structures not previously encountered, by various kinds of generalisation [11], [13], and thus it is possible to infer futures from novel encounters.

Next, we propose that the Global Workspace is sensitive to information-theoretic qualities [15] of the ideas being generated. MacKay [32] distinguishes between *information content* or *unexpectedness*, h, an estimate of the number of bits required to describe an event, e, given a context, c:

$$h(e \mid c) = -\log_2 p(e \mid c),$$

and *entropy*, H, which is defined as an estimate of the *uncertainty* inherent in the distribution of the set of events $\mathcal{E}$ from which e is selected, given the context, c:

$$H(c) = -\sum_{e \in \mathcal{E}} p(e \mid c)\, h(e \mid c) = -\sum_{e \in \mathcal{E}} p(e \mid c)\, \log_2 p(e \mid c).$$

H is maximised when all outcomes are equally likely, and zero when a single outcome is certain. Both of these quantities are useful in the model, at different times.

First, consider h_n, the unexpectedness of a partial model of actual on-going experience in a particular state, n. If the experience is likely (in particular, if it is *readily predictable* from what has gone before), it is not unexpected, so h_n is low; if it is unlikely, it is unexpected, so h_n is high. Unlikely predictions are infrequent in the Workspace, and so have low volume, but high h_n draws attention. I call this the *recognition-h* case: it explains why unexpected things are noticed.

Now, consider, h_{n+1}, the unexpectedness of a predicted situation. It is maximally unlikely that a *prediction* will be made including a percept that has not been encountered before, and, as above, we would therefore expect h_{n+1} to be very high. However, if such a *prediction* is made, we would not want it to be too loud. Excess of such predictions, or even repeated occurrence of a single one, in general, would lead to a state of constant anxiety. Of course, in a simplistic frequentist account, predictions introducing novel percepts or concepts cannot arise; this is why I include generalisation and/or interpolation in the theory, as above. I call this the *prediction-h* case. It may explain why surprising predictions are more likely to draw attention than unsurprising ones.

Over-active prediction-h is mitigated by the mechanism above, where prediction is probabilistic and additive across predictors. There are two opposing forces here, one changing inversely relative to the other; because they are co-occurrent, their effects should multiply. Therefore, the overall outcome audible in the global workspace can be estimated by multiplying the probability, p, of an event (estimating the likely number of generators predicting it) by h (estimating the volume at which they are predicting). The resulting likelihood is illustrated by Figure 1. This creates a bias away from predictions which are either very likely or very unexpected, reducing the power of the very unlikely or the very obvious to attract attention. This explains why unlikely possibilities do not overwhelm the acting organism with choice, and accounts for Wundt's *hedonic curve* in creative preference [33].

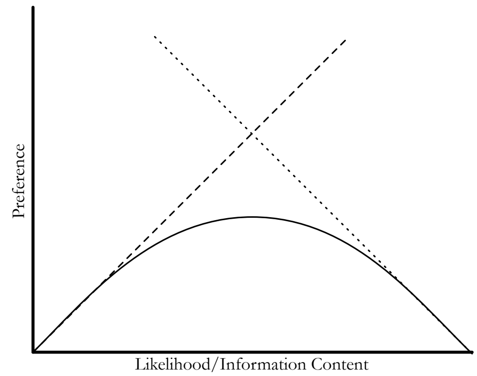

Fig. 1. Illustration of the interaction between likelihood and unexpectedness. The overall likelihood (solid) is formed by the multiplication of two monotonic functions: the unexpectedness of a generated item (dashed) and the number of generators likely to agree on it, according to its likelihood (dotted).

Note the difference between recognition-h and prediction-h in the context of the Global Workspace. Agent-generators may generate structures of either kind, and the two compete for attention. Thus, clear and present danger or benefit outweighs predicted likelihoods, because the distribution of *potential* predictions is over a much wider range of possibilities than that over *actual* perceptions, and therefore, comparatively, probability mass is spread more thinly. Conversely, for example, likely but unexpected predicted benefits can outweigh less seriously dangerous present circumstances—thus, prioritising an unusual positive opportunity emerges mechanistically.

H, the expected value of the information content of a distribution, is different from h, which deals with specific situations. It is characterised as the *uncertainty* inherent in a distribution [32]. Unlike h, H really only has meaning in the predictive context: once one knows which possibility of a range is the right one, only information content is really relevant. However, a certain predicted outcome is more useful than one which is uncertain: H measures this difference.

I propose, therefore, that, in the predictive generators, higher H also predicts lower volume, so that less certain generated outputs are de-emphasised. This, I call *prediction-H*. It may explain how it is possible to *feel certain* about intuitions (as opposed to be convinced by reasoned argument). It prevents the Global Workspace from being flooded with unsupported predictions, allowing secure predictions to shine through. A particularly interesting point is this: an unlikely prediction, with sufficient prediction-h to be audible, in the absence of other explanations, will have low prediction-H, and so will not be suppressed by this final mechanism.

No straightforward diagram (cf. Fig. 1) can be drawn of the effect of prediction-H on the overall likelihood of a generator taking over the Global Workspace, because the numbers depend heavily on the multidimensional distributions from which the various Hs are calculated. However, it is possible to specify a volume value for each idea, V, which is estimated by the following, for either kind of h, above:

$$V = \frac{p \times h}{H}.$$

V can be used to adjudicate between candidate ideas: the highest one at a time has access to the Global Workspace.

VI. Conclusion

I have outlined a simple mechanism by which statistically likely and information-theoretically rich novel structures can emerge from a multi-agent system furnished with high-quality models of a domain of knowledge. This mechanism for choosing access to consciousness simulates inspiration by managing admission of (partial) ideas to conscious awareness, and drawing attention to them. I should note that it is possible that such a mechanism is one of Baars' own proposals; however, if so, it is not clearly specified as such.

Note that this mechanism can apply to any statistical model available to the generators, so it need not be restricted to music (as it is in the system components summarised in the next section). In principle, the same idea can work with any model from which statistical likelihoods can be computed. This means, for example, that it can account for the generation of sentences, and therefore possibly internal speech. If internal speech is equated with essential thought, as commonly, then the current approach can account for general creative thought and for the emergence of particular thoughts into consciousness as intuition. Prediction-H can account for the feeling of certainty associated with thoughts and intuitions.

As for the Threshold Paradox: it is not in the threshold, but in the formulation of the Global Workspace as requiring one.

Acknowledgments

This work was funded by EPSRC grant EP/H01294X, 'Information and neural dynamics in the perception of musical structure'. I thank my colleagues in ISMS, especially Marcus Pearce, whose doctoral work underpinned my thinking.

References

[1] A. Cardoso, T. Veale, and G. A. Wiggins, "Converging on the divergent: The history (and future) of the international joint workshops in computational creativity," *AI Magazine*, vol. 30, no. 3, pp. 15–22, 2010.

[2] M. Minsky, *The Society of Mind*. New York, NY: Simon and Schuster Inc., 1985.

[3] B. J. Baars, *A cognitive theory of consciousness*. Cambridge University Press, 1988.

[4] M. Minsky, "Music, mind and meaning," *Computer Music Journal*, vol. 5, no. 3, pp. 28–44, 1981. [Online]. Available: http://www.ai.mit.edu:80/people/minsky/papers/MusicMindMeaning.html

[5] S. Pinker, *The Language Instinct*. Perrenial, 1995.

[6] H. Plotkin, *Evolution in Mind*. Cambridge, MA: Harvard University Press, 1998.

[7] E. Holmes, *The Life of Mozart: Including his Correspondence*, ser. Cambridge Library Collection. Cambridge University Press, 2009.

[8] D. Dennett, *Consciousness explained*. Boston: Little, Brown and Co., 1991.

[9] D. Huron, *Sweet Anticipation: Music and the Psychology of Expectation*, ser. Bradford Books. Cambridge, MA: MIT Press, 2006.

[10] G. A. Wiggins, "Cue abstraction, paradigmatic analysis and information dynamics: Towards music analysis by cognitive model," *Musicae Scientiae*, vol. Special Issue: Understanding musical structure and form: papers in honour of Irène Deliège, pp. 307–322, 2010.

[11] P. Gärdenfors, *Conceptual Spaces: the geometry of thought*. Cambridge, MA: MIT Press, 2000.

[12] D. Conklin and I. H. Witten, "Multiple viewpoint systems for music prediction," *Journal of New Music Research*, vol. 24, pp. 51–73, 1995.

[13] M. T. Pearce, "The construction and evaluation of statistical models of melodic structure in music perception and composition," Ph.D. dissertation, Department of Computing, City University, London, UK, 2005.

[14] G. A. Wiggins, M. T. Pearce, and D. Müllensiefen, "Computational modelling of music cognition and musical creativity," in *Oxford Handbook of Computer Music*, R. Dean, Ed. Oxford University Press, 2009, ch. 19, pp. 383–420.

[15] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–56, July and October 1948.

[16] M. T. Pearce and G. A. Wiggins, "Expectation in melody: The influence of context and learning," *Music Perception*, vol. 23, no. 5, pp. 377–405, 2006.

[17] ——, "Evaluating cognitive models of musical composition," in *Proceedings of the 4th International Joint Workshop on Computational Creativity*, A. Cardoso and G. A. Wiggins, Eds. London: Goldsmiths, University of London, 2007, pp. 73–80.

[18] M. T. Pearce, D. Conklin, and G. A. Wiggins, "Methods for combining statistical models of music," in *Computer Music Modelling and Retrieval*, U. K. Wiil, Ed. Heidelberg, Germany: Springer Verlag, 2005, pp. 295–312. [Online]. Available: http://www.doc.gold.ac.uk/ mas02gw/papers/cmmr04.pdf

[19] M. T. Pearce, M. Herrojo Ruiz, S. Kapasi, G. A. Wiggins, and J. Bhattacharya, "Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation," *NeuroImage*, vol. 50, no. 1, pp. 303–314, 2010.

[20] D. D. Corkill, "Blackboard systems," *AI Expert*, vol. 6, no. 9, pp. 40–47, 1991.

[21] G. Wallas, *The Art of Thought*. New York: Harcourt Brace, 1926.

[22] G. Tononi and G. M. Edelman, "Consciousness and complexity," *Science*, vol. 282, no. 5395, pp. 1846–1851, 1998.

[23] G. M. Edelman, J. A. Gally, and B. J. Baars, "Biology of consciousness," *Frontiers in Psychology*, vol. 2, 2011.

[24] A. Koestler, *The Act of Creation*. London: Hutchinson & Co., 1964.

[25] J. Guilford, *The Nature of Human Intelligence*. New York: McGraw-Hill, 1967.

[26] J. Getzels and M. Csikszentmihalyi, *The Creative Vision: A Longitudinal Study of Problem Finding in Art*. New York: Wiley, 1976.

[27] M. A. Boden, *The Creative Mind: Myths and Mechanisms*. London: Weidenfield and Nicholson, 1990.

[28] ——, "Creativity and artificial intelligence," *Artificial Intelligence Journal*, vol. 103, pp. 347–356, 1998.

[29] G. A. Wiggins, "A preliminary framework for description, analysis and comparison of creative systems," *Journal of Knowledge Based Systems*, vol. 19, no. 7, pp. 449–458, 2006. [Online]. Available: http://dx.doi.org/10.1016/j.knosys.2006.04.009

[30] ——, "Searching for computational creativity," *New Generation Computing*, vol. 24, no. 3, pp. 209–222, 2006.

[31] G. Ritchie, "Assessing creativity," in *Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences*. Brighton, UK: SSAISB, 2001, pp. 3–11. [Online]. Available: http://www.dai.ed.ac.uk/homes/graeme/papers/aisb01.ps

[32] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press, 2003. [Online]. Available: http://www.inference.phy.cam.ac.uk/mackay/itila/

[33] E. H. Margulis and A. P. Beatty, "Musical style, psychoaesthetics, and prospects for entropy as an analytic tool," *Computer Music Journal*, vol. 32, no. 4, pp. 64–78, 2008.