Stimulating creative flow through computational feedback

Daniel Jones, Oliver Bown, Jon McCormack, François Pachet, Michael Young, Rodney Berry, Iris Asaf and Benjamin Porter

September 28, 2009

Abstract

This report summarises the discussion and experimental work produced by the authors at the 2009 symposium *Computational Creativity: An Interdisciplinary Approach*, Dagstuhl Leibniz-Zentrum für Informatik. It outlines the motivation for using computational techniques to stimulate human creativity, briefly summarising its historical context and predecessors, and describes two software studies produced by the group as base-line exemplars of these ideas.

> "Already at the very beginning of the productive act, shortly after the initial motion to create, occurs the first counter motion, the initial movement of receptivity. This means: the creator controls whether what he has produced so far is good."
>
> Paul Klee, *Pedagogical Sketchbook* [6, p33]

Within the field of computational creativity, theoretical emphasis is frequently located on one side or other of an established dichotomy: either studying and formalising the human creative process through analysis, or producing new systems which exhibit such notions of creativity through synthesis. There is, however, a third relationship between creativity and computation which is arguably, to this date, under-theorised: namely, using computational systems to provoke, augment and reshape the human creative experience through a succession of ongoing interactions.

This working group was formed to discuss these processes under the aegis of *stimulating creativity through feedback*, where the 'feedback' in question is a continuing response to a participant's creative acts through computational means. The key motivation is to open up the field of productive possibilities by encouraging potentially inaccessible or non-obvious avenues of creation.

1 Synopsis

The standard approach in such a model uses an input-processing-output architecture. A participant is engaged in some kind of creative act – for example, drafting an architectural layout on a virtual artboard. Each mark made by the participant is translated into a symbolic form and parsed by a computational

Dagstuhl Seminar Proceedings 09291
Computational Creativity : An Interdisciplinary Approach
http://drops.dagstuhl.de/opus/volltexte/2009/2223

component, which, based on this and its previous states, generates an output, usually in the same medium as that used by the participant. In this case, the system may produce a series of new marks on the artboard, in some relationship to those previously drawn. It is the nature of this relationship that provides the most significant questions and potential challenges.

This approach is analogous to the PqF architecture proposed by Blackwell [2] in his research on Live Algorithms for Music. According to Blackwell, a well-designed system to autonomously respond to musical input from a human performer should have three components P, f, Q:

- listening (P)
- generation of response and/or new ideas (f), and
- articulation by means of [*creative*] gestures (Q)

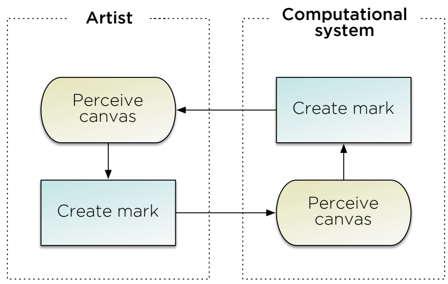

Figure 1: Feedback between human participant and computational system

The immediate analogy is that of an artistic mirror, reflecting the artist's marks under some degree of transformation to heighten their awareness of their own style. On one hand, this may alert them to tropes and habits that they have adopted, which can then be broken or modified; on the other, if this 'mirror' is sufficiently distorted, it may offer them new directions or opportunities which may not have been adopted without external input. The distortion is introduced by the operations performed in the f stage: a Q response may be generated by closely imitating the marks made in P, or by using external material to create an output which may be difficult to relate to P.

Finding the optimal level of distortion is perhaps one of the key challenges in this domain. The expected result should roughly follow the shape of the Wundt curve (Figure 2): the payoff increases as the novelty (that is, distortion relative to the input) increases, up to a peak, after which point it drops off to become aesthetically unrelated and nonsensical[1].

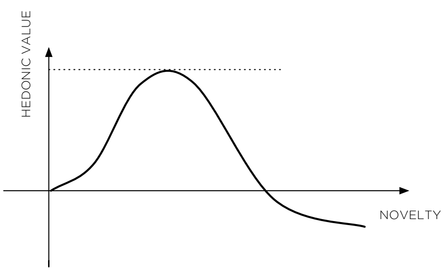

Figure 2: The Wundt curve

An instructive example, paradigmatic within the field, is Pachet's Continuator [8]. The Continuator is a software system which uses a Markov model of a pianist's performance to generate new melodic sequences, reflexively demonstrating their rhythmic, harmonic and stylistic properties. Control alternates between pianist and continuator:

[1]Though nonsensical responses may be appealing in certain scenarios, and so this measure should be taken as subjective. We assume a position that seeks a small degree of novel direction.

the pianist plays a sequence, the Continuator responds with its own output, and the pianist responds once more.

With its Markovian architecture, the Continuator requires a minimum of predefined musical knowledge, and functions upon pure probabilities based on the frequency of events in the ongoing corpus collected through its 'listening' phases. It is, therefore, stylistically agnostic, and can act as a creative mirror in a demonstrably broad sense. By introducing multiple Continuator agents, or playing with a corpus previously created from another musician's performances, it can alternatively be used to extend performance techniques in a non-linear fashion. The author suggests that this has the capacity to open up new musical modes which are not ordinarily possible: for example, as a "musical amplifying mirror" [8, p7], multiplying and layering the pianist's performances indefinitely.

The Continuator is elegant in its uncomplicated yet aesthetically flexible design, pushing a mimetic strategy just far enough to create a wide range of interesting responses. In doing so, a process of iterative feedback is formed, as the user recognises emergent patterns and proceeds to build upon and redirect them.

It is our view that the cumulative building process described here is an entry route to a state which has broad ramifications for this research field: that which is described as 'flow', a fully-engaged creative situation wherein each mark or note follows intuitively. Yet despite this intuitiveness, flow is closely linked to unpredictability, newness and innovation, which drive the artist to explore a creative space.

1.1 Flow

Csikszentmihalyi talks of 'flow' in terms of an optimal state of experience, and as an increase in the intensity or complexity of consciousness [3, p74]. In a given practice – let us say, playing the piano – the experience of 'flow' is to be found when the level of challenge is well-matched to the player's level of skill. If an overly difficult piece is attempted, the player is frustrated by the inability to perform it adequately; if the piece is too easy, it will be found trivial and quickly grow boring. A channel of optimality is thus found along the line of intersection of the two[2] (Figure 3).

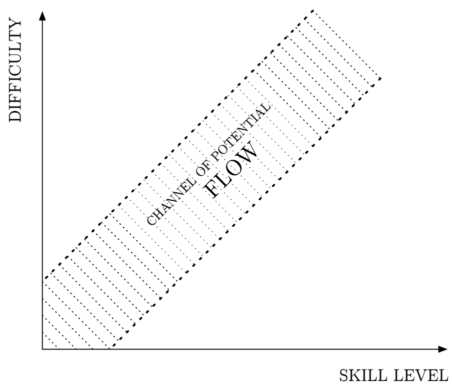

Figure 3: Region of potential flow states

For a computational creative stimulant to encourage 'flow', it must move with the player's set of skills and expectations, encouraging goals to be set just beyond their level of comfortable ability. Furthermore, by aiming towards the 'flow' state, the emphasis moves

[2]Csikszentmihalyi notes, however, that this is a necessary but not sufficient criterion for a flow experience; "A professional athlete might be 'playing' football without any of the elements of flow being present: he might be bored, self-conscious, concerned about the size of his contract rather than the game." [3, p76]

to the *process* rather than the final outcome: the payoff is, as Csikszentmihalyi asserts, this ongoing 'optimal state' itself.

This class of activity is tied closely to scenarios of improvisation, be it playing jazz piano, participating in a dinnertable conversation, or playing a game of football. The ruleset and goals may be more or less defined, though a participant should be able provide some judgment of the scenario and their engagement with it. Of crucial importance is an uncertainty as to what should come next, and the continual process of decision-making that fills this space of indeterminacy. In a flow situation, of course, the level of exterior focus should be such that this decision-making goes fully unnoticed.

Smith and Dean [11] describe such an "exploratory" nature of improvisation in both auditory and visual fields, positioning the work as (partially, we would qualify) "self-generating" – marks can suggest other marks, creating forking paths to be selected and followed by the artist.

Jazz percussionist Eddie Prévost also advocates an exploratory strategy, citing Thelonius Monk and his endeavours for his music "to find other places" [9]. Prévost suggests that the aspects of chance in an improvised musical performances are opportunities to make unforeseen 'errors' which can subsequently be followed and investigated. He recounts a tale in which Monk, frustrated with an improvised performance, complained that he had "made all the wrong mistakes" [9] – indicating the existence and appeal of *correct mistakes*, which may aid us in this creative search.

To follow this path computationally, then, we are effectively designing for *serendipity* [1]: tacitly encouraging or in(tro)ducing "correct mistakes" as a route to unforeseen discoveries and new creative terrain.

2 Context

Outside of the computational realm, there is a vast ancestry of strategies to intentionally challenge and provoke creative action[3]. Csikszentmihalyi recounts an ethnographical report of the Shushwap Native American practice of uprooting and relocating its village every 25-30 years – in doing so, introducing novel, chaotic challenges and ensuring a continual enrichment of cultural cycles.

More recently, the Surrealist practice of automated writing and the Oulipo group's exercises in constraint [7] have offered new creative routes to writers (paradoxically, often through *restricting* the parameters of their output). The improvised painting of the Cobra group drew up a manifesto describing the process of "finding" a painting through its production, seeking an art which is "spontaneously directed by its own intuition" [11, p108]. Later, the American abstract expressionists adopted practices such as action painting, aleatoric and combinatorial techniques to surrender direct authorship of their works [11, p109]. A broader approach is taken by Eno and Schmidt's 'Oblique Strategies' cards [4], which claim to suggest escape routes from creative deadlock via koan-like prompts.

The computational feedback approach also lies somewhere on a continuum with generative art [5], in which the artist sets up a system with a given set

[3]By chance, the symposium that produced this paper took place 50 years after the like-named *"Interdisciplinary Symposia on Creativity"*, Michigan State University (1959), whose focus was on "creativity and its cultivation".

of rules for its ongoing autopoiesis, either computationally or otherwise[4]. A purely generative work would involve no subsequent intervention after it has been set in motion; a work with no generative elements would therefore have a null ruleset, and so may, say, generate a direct mirror of its input. The class of systems that we are interested in lies somewhere between the two, producing an ongoing output which is neither a direct mirror of the artist's input, nor does it disregard this input and follow only its internal logic. Its output will inevitably be mediated through some predetermined structure or ruleset – such as Pachet's Markov chains [8] – but should always follow from the artist's (and, in some cases, the computational system's) previous marks.

3 The Studies

Following the delineation of what constitutes a feedback system to stimulate creative flow, and the theoretical dissection of what is at stake when pursuing such goals, two systems were produced with the intention of creating tangible embodimentss of these ideas. Rather than building a large-scale framework to create aesthetically complex outputs, a more theoretically instructive approach was adopted: to create a number of systems whose complexity was *minimally sufficient to demonstrate and explore the qualities in question.* In doing so, we hoped to gain further insights into these qualities and their *a priori* conditions.

3.1 Study #1: Group Improvisation in Colour and Form

The first study[5] was an investigation into the process of group improvisation in the visual domain. Using a projection screen and a MIDI controller, an interface was devised in which up to 9 participants can move, resize and recolour a series of primitive shapes located on a white canvas. Rather than the turn-based approach of the Continuator, transformations are immediate, so all of the participants can immediately perceive any changes made by the others.

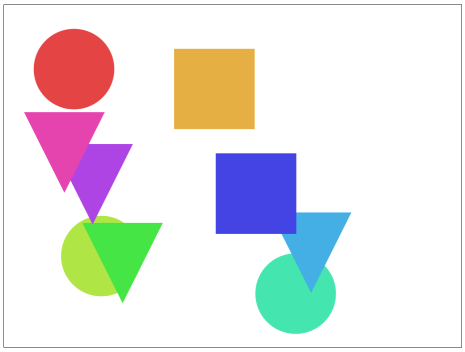

Figure 4: Screenshot from Study #1

It should also be noted that this approach enables prior marks to be unmade and modified; unlike in musical improvisation, where previously-produced sounds cannot be erased, the entire canvas remains fluid and subject to change. As a result, there is the continual possibility that a radical move can render the pictoral plane sufficiently unstable to require further radical changes to become balanced.

[4]Galanter provides an example of primitive tiling techniques following a set of internal generative rules following from the selection of tiling shapes. We take this as an appealing baseline example of what constitutes generativity.

[5]theoretically developed by McCormack, Jones, Pachet, Berry, Asaf and Porter, with technical implementation by Jones

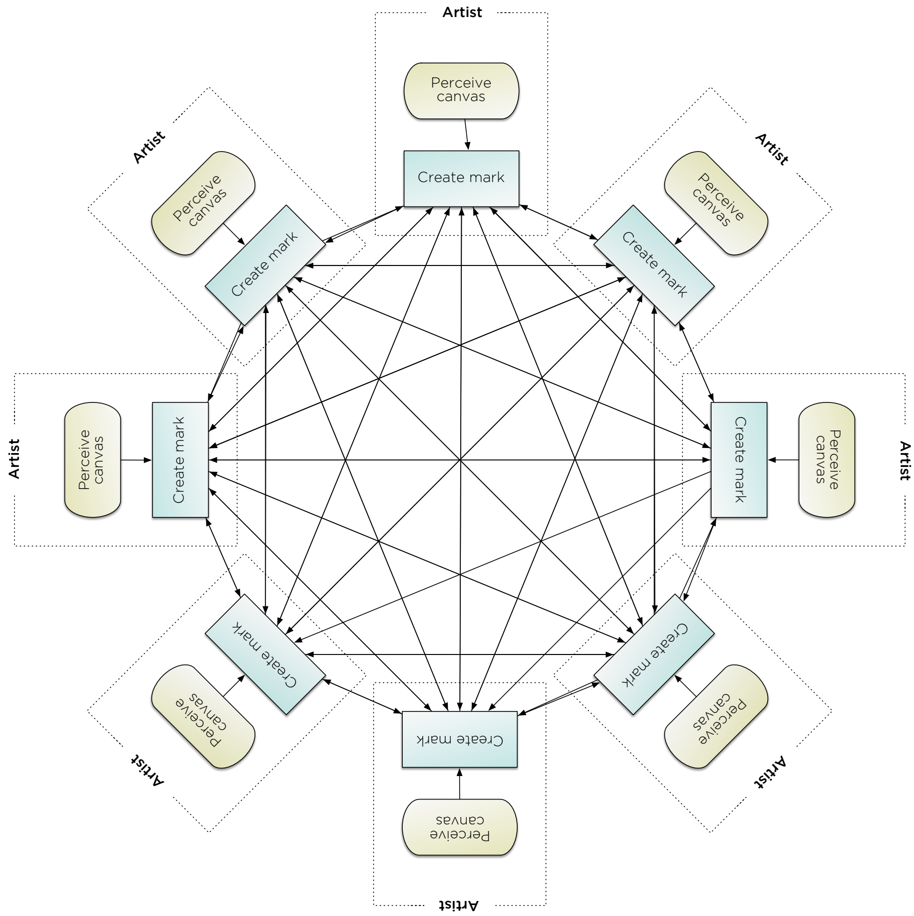

Figure 5: Multi-participant feedback

The first striking property of the system to emerge through early experimentation is its gamelikeness: it gives rise to undirected, free play, whose state and trajectory are constantly under reassessment. Each fresh scenario begins with very wide, chaotic modifications; as tendencies are recognised and participants gradually attune to their peers' behaviours, the plane begins to coalesce, and changes become finer. If participant-group A initially tends towards closely aligning similar shapes, and participant-group B tends towards scattering their different shapes widely, an emergent behaviour may be a few different clusters of forms. Different properties may take the fore in each scenario, and participants may change their behaviour between scenarios.

Of particular interest is how the final outcome emerges from the continual interplay of judgments and actions between the web of participants (see Figure 5). Smith and Dean put forward the notion that, in such a closely coupled situation of group improvisation, individual trajectories are assimilated into a merged group-subject, in which the contributions of each artist become indistinguishable to the point where it is difficult to say who has done what [11, p35].

They continue:

> The improvising process often makes the improvisor concentrate on small elements because (s)he cannot see the overall whole. This means that a multiplicity of elements is likely to build up, especially where a number of different improvisors are interacting with each others' small units. The total effect is likely to be very different from that which develops when a creator starts out with a whole in mind. [11, p33]

In his discussions of group improv in theatre, R. Keith Sawyer also emphasises the importance of this "decentralized mindset" [10]. Many of his accounts focus on emergent dialogues between actors, whose subject matter is honed through each new statement. We suggest that the notion of a dialogue can be extended to the non-verbal exchanges seen in scenarios such as this visual improvisation: each act is a response to the cumulative set of prior statements, creating a well-formed exchange (though, arguably, with less straightforwardly formulated syntactics and semantics than spoken language).

A natural challenge to this study is that it lacks any significant computational element; it is merely a dynamic visual framework, with no element of analysis or augmentation. We view it, however, as the first step into an investigation of visual improvisation. Several next steps suggest themselves: for example, a performative scenario could be recorded and subsequently analysed to gain further insight into the developing interactions between participants.

A second approach, closer aligned to our original goals, would be to introduce some computationally-controlled participants, whose behaviour could mimic or avoid that exhibited by the human players to encourage different responses. This is more akin to the behaviour of our next software study.

3.2 Study #2: Eluding Expectations in Rhythmic Cycles

The second study[6] investigated alternative modes of interaction to that of mirroring and, considered their relationship to the notion of flow. Bown and Young proposed that notions such as conflict, struggle and negotiation were also relevant, and developed the idea of *negotiation* as an alternative mode of interaction in human-computer improvised performance alongside *shadowing* (system tracks performer and synchronously shadows certain aspects of performer's action, perhaps with transformation), *mirroring* (system reflects back certain aspects of performer, such as style in the case of the Continuator) and *coupling* (system and performer are engaged in very loosely defined process of mutual influence). Negotiation is proposed as defining an equal relationship between participants (neither participant is source or respondent, as in the case of shadowing or mirroring), where each participant has goals that are expressed in terms of the composite outcome of the musical activity, and which may not be fixed: i.e., both actions and goals can be negotiated.

They developed a prototype system to explore these ideas: Clap-along, a duet system for human-computer interaction, in which human and computer participants produce continuous, synchronised 4-bar clapping patterns in 4/4.

[6] developed by Bown and Young

The musical context of Clap-along is as minimal as could be conceived: a fixed tempo, a fixed metrical structure, and single sound events quantised to beats.

In any loop instance n there is a human clapping pattern (H_n), a computer pattern (C_n) and a composite of the two patterns (R_n). A feature set, F_n, is extracted from R_n and compared by the system to a target feature set T_n, and this comparison is the basis for the next iteration. R_n is the reality of the current state, and T_n represents an expectation that is unknown to the human performer.

At the end of each 4-bar pattern, the system takes the composite of the two rhythms, R_n, and calculates a feature set F_n that forms a minimal internal representation of the musical output. The four features used in the prototype version were:

- density: number of claps as a fraction of the maximum possible number.
- homophony: number of coincident claps as a fraction of the maximum possible value.
- position weighting: normalised average position in the cycle over all claps.
- clumping: the average size of continuous clap streams as a fraction of the maximum possible value.

In this multi-dimensional feature space, the system calculates the Euclidean distance between F_n and the target feature set T_n. If the distance exceeds an arbitrarily defined threshold, this is deemed to indicate a significant musical difference between reality and the expectation.

If the distance between F_n and T_n is under the threshold – i.e. reality is sufficiently close to the expectation – the system introduces random variation to its expectation to produce T_{n+1}, with mutation along each feature dimension drawn from a Gaussian distribution.

The human performer is invited to negotiate with the system in a possibly comparable way. The negotiation occurs both in the feature space and in the foreground surface of actual rhythms. This is because the performer's actions and immediate descriptions of interaction can only be formulated with reference to the musical surface, in order to gradually develop a deeper understanding of negotiation within the machine's feature space. This offers a considerable challenge to the performer.

As intended, the behaviour of the system offered no obvious indication of its interactive capacity. That is, there was no way to verify through interaction that the system was engaged in musically significant process. This evoked a sense of awkwardness clearly contradictory to flow, but still of interest. The performer tended to be reluctant to join in with the computer-generated clapping until a pattern had been identified. In addition, once the performer did join in, this dramatically changed the value F, the computer's representation of the music, causing it in turn to vary its output in radically different ways. Clapping the exact same pattern repeatedly (including not clapping at all) would cause the system to slowly evolve its output until its expectation was satisfied, at which point its expectation would drift gradually, or occasionally make a larger leap. In fact, since the features specified were not independent of each other, it was not guaranteed that a given expectation could actually be reached, in which case the system would get stuck. This introduced an unexpected new scenario into the interaction: the performer could actually assist the system out of its stuck state using different patterns.

These scenarios, whilst perhaps closer to thought experiments than real psychological tests, are interesting from the point of view of thinking about human-computer musical interaction from first principles. The experiment offered further ways to look at how different performers go about negotiating with these kinds of simple 'black-boxes' under different scenarios, and how strategies to deal with the computer's behaviour are developed. In all cases, however frustrating or rewarding, these procedures manipulate the expectations and actions of human and machine performers alike, engendering an unresolved negotiation that may have intrinsic creative value.

4 Summary

In his *Pedagogical Sketchbook*, Paul Klee talks of the chain of the 'productive' and 'receptive' [6] act and counter-act that continually take place through a creative act. We believe that computational means may be of use to influence this see-saw of creation and judgment, opening up novel possibilities without the need to construct a formal model of creativity. Such a system should not be referred to as *itself* creative; rather, it operates as a tool to augment our own creative behaviours. We feel that this delineation may prove constructive for future research.

References

[1] Paul André, M. C. Schraefel, Jaime Teevan, and Susan T. Dumais. Discovery Is Never by Chance: Designing for (Un)Serendipity. In *C&C '09*, Berkeley, California, 2009. ACM.

[2] Tim Blackwell and Michael Young. Swarm Granulator. In G R Raidl et al, editor, *EvoWorkshops 2004*, pages 399–408, Berlin, 2004. LNCS 3005, Springer-Verlag.

[3] Mihaly Czikszentmihalyi. *Flow: The Psychology of Happiness.* Random House, London, 1992.

[4] Brian Eno and Peter Schmidt. *Oblique Strategies: Over 100 worthwhile dilemmas by Brian Eno and Peter Schmidt.* Apollo, London, 1975.

[5] Philip Galanter. What is generative art? complexity theory as a context for art theory. In *GA2003 6th Generative Art Conference*, 2003.

[6] Paul Klee. *Pedagogical Sketchbook.* Faber and Faber, London, 1972.

[7] Harry Matthews and Alastair Brotchie. *The Oulipo Compendium.* Atlas Press, London, 2005.

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

[9] Eddie Prévost. *Minute Particulars: meanings in music - making in the wake of hierarchial realignments and other essays.* Copula, London, 2004.

[10] R. Keith Sawyer. Group creativity: musical performance and collaboration. *Psychology of Music*, 34(2):148–165, 2006.

[11] Hazel Smith and Roger Dean. *Improvisation, Hypermedia and the Arts since 1945*. Harwood, Amsterdam, 1997.