Essays in Philosophy

A Biannual Journal

Vol. 8, No. 1, January 2007

Book Review


_The Creative Mind_ is a lively, wide-ranging, accessible discussion of human creativity and how computers can help us understand it. First published in 1990, the second edition (2004) adds an introductory ten-page overview of Boden’s position (“In A Nutshell”) and an Epilogue surveying recent developments, such as new programs for musical composition, architectural design, drawing, proving mathematical theorems, and connectionist networks that can be evolved rather than designed. The one significant change in the body of the book is the added discussion of Douglas Hofstadter’s analogy program, COPYCAT.

Boden’s basic position, then, is unchanged from the first edition. She is well aware that many readers will doubt or flatly deny that computers can tell us anything about creativity, on grounds that they have no purposes, desires, or values, and can do nothing but what their human programmer tells them to. This was the view of one Lady Lovelace, and Boden responds to such broadsides by distinguishing what she calls the four “Lovelace questions”: 1) Can computational concepts help us to understand human creativity?; 2) Could a computer (now or in the future) appear to be creative?; 3) Could a computer ever appear to recognize creativity?; 4) Could a computer ever really be creative? She is inclined to answer ‘no’ to the last question, which she considers a complex philosophical issue rather than a matter of empirical investigation (Ch. 12). But most of the book concerns the first three questions, and above all the first, to all of which she answers ‘yes’. Her primary concern is to demonstrate that, and specifically how, our thoughts about human creativity, as difficult as they may be to make clear and as tempting as it can be to cast them in the form of suggestive but vague metaphors, can to a large extent be rendered precise and testable by formulating them in computational terms and implementing them in functioning computer programs. Moreover, two facts lend credibility to the notion that computational theories such as Boden’s really are theories of human creativity and not just accounts of the achievements of fancy mechanical devices: a) relatively recent connectionist programs have a degree of psychological and neural plausibility far beyond that of earlier, linear computer programs, and b) specific programs in a wide variety of domains have been able to mimic ever more impressively both the products and (arguably) some of the processes manifest in human creativity.

More specifically, Boden’s argument that computational concepts can help us understand creativity begins by distinguishing three basic types of creativity: (i) (re)combination of familiar items within some domain to form new ones; (ii) exploration of an established conceptual space to discover new
and perhaps quite unexpected possibilities within that space; (iii) transformation of a conceptual space itself to arrive at things (ideas, artifacts, etc.) that were not even possible within the earlier space. (She also distinguishes P-creativity from H-creativity: a valuable idea is P-creative if the person who produced it could not have done so before; it is H-creative if, in addition, no one has ever produced that idea before. But the latter question involves historical issues and is not a major concern of the book, which is in effect about the more basic notion of P-creativity.) Although Boden notes more than once that genuine creativity of any sort must produce not mere novelties, but novelties that are of some interest or value, her focus throughout is on computational aspects of the topic. In keeping with that approach, Boden finds the familiar metaphor of “conceptual space” apt but vague; reformulated in more precise computational terms the most striking instances of creativity can be said to involve results that, relative to some previously established system or structured style of thought or production, could not possibly have arisen: they were computationally impossible, and became possible only through a change in the principles or rules of the underlying generative, computational system itself.

Boden is quite sympathetic toward other approaches, and discusses (Chs 1, 2) some classical theories of creativity (e.g., Poincaré’s four stages of a) preliminary conscious work, b) non-conscious incubation, c) conscious insight, and d) evaluative work, or Koestler’s theory of bisociation of matrices), along with such oft-told tales of creativity as Coleridge’s imagery in Kubla Khan or Kekule’s discovery of the cyclical structure of benzene rings, of which she gives an unusually detailed and instructive account. She finds much of value in these accounts and attempts in Chapters 3-6 to incorporate what is useful in them (and especially in Koestler’s main idea) into a computational theory of creativity. Fundamentally this requires postulating for each domain considered an underlying set of computational principles that define a space of possibilities (like a set of grammatical rules defining an infinite space of grammatical sentences or the rules of chess defining infinitely many legal moves). On this basis one can view instances of creativity as the result of explorations or transformations of that space, or simply new combinations of familiar elements within that space.

The most urgent question about the most striking form of creativity (“transformational”) is that of how the mind could possibly manage to transform its own conceptual space so as to go beyond what was even possible before. Boden’s answer (Ch. 4) draws on the developmental psychologist Annette Karmaloff-Smith’s notion of “representational redescription”. Roughly, this is our capacity to represent explicitly and in detail activities that had previously been carried out as an implicitly represented routine (e.g., the motor routines involved in riding a bike, or a four-year-old’s largely implicit representation of how to draw a human being)—and subsequently, to re-represent those representations, and so on to higher levels of representation. Forming explicit representations of our own representations allows us to explore or transform them, and thereby to alter fundamentally our thought or action. (It may be noted in passing that the question of how we are able to formulate maps of our own conceptual spaces is not itself explored.) Some present-day computer programs do in fact include representations of their own procedures, and can on that basis alter those procedures.

One may describe associative, exploratory, and transformative creativity either in “classical” or connectionist style networks (Ch. 5), and Boden discusses a wide array of actual computational programs implementing various instances of creativity (Ch.6-8)—e.g., the discovery and interpretation of analogies, musical composition, discovery or proof of mathematical theorems,
formulation of hypotheses in organic chemistry, story telling, composition of haikus, musical improvisation, etc. Here connectionist or PDP (parallel distributed processing) models suggest how the mind is able to go beyond the “cut and paste” combining of ideas to produce a merging or blending of old ideas in a new structure. (One might hope that a future edition will apply this important idea to the familiar contest between associationistic and organic/“genius” theories of literary creativity, since one major claim of the later—as propounded by Goethe, Coleridge, and others—was precisely that true creativity was not a “brick and mortar” or merely combinational process, but brought together and transformed old ideas into something more radically new.) A further interesting implication of all this is that the generation of creative ideas in the arts takes place in the same basic way(s) as in the sciences.

These discussions range from the interesting to the fascinating, and are in general presented with admirable lucidity. (One possible exception: the explanation of connectionism itself is perhaps a bit too breezy.) In some domains, such as drawing, generating rules for analyzing organic molecules, or diagnosing soybean diseases (this is not so easy!), the results have been impressive; in others, such as storytelling, results have been very limited or even comically bad (e.g., early TALE-SPIN’s story of Henry Ant and Bill Bird). But even here much progress has been made, and the attempt to figure out how computers might better spin a tale has generated interesting ideas about how the mind might in fact represent human emotion, motivation and action. Probably most impressive are programs that can “train themselves up” or even “evolve” by changing their own ground rules or heuristics—either in light of evaluations of results provided from outside, or in accord with built-in principles of evaluation, or randomly. Such programs (e.g., “EURISKO”, or the mighty “ID3 algorithm”) have achieved impressive results in art, mathematics and science, and should be taken into consideration by anyone who is tempted to simply assume that computers “can only do what they are told to do.” Although the question of whether even these programs are genuinely creative is complex and controversial, they support Boden’s primary claim that by generating, clarifying, and putting into practice a range of substantial ideas about creativity—including the reflexive mapping of thought, evaluation of ideas, and transformation of conceptual spaces—a computational approach can be extremely helpful in thinking about human creativity.

The remaining chapters move quickly through a series of interesting but less central questions. Chapter 9 argues that a scientific account of creativity need not imply either determinism or predictability (nor, as chaos theory shows, does determinism entail predictability-by-humans), and discusses a number of important ways in which chance or random or serendipitous factors can enter into creativity in combination with constraints, where the latter remain central to creativity. Chapter 10 defends the view that creativity is neither a single nor special capacity, but an aspect of intelligence in general, involving a host of familiar abilities (memory, perception, noticing, comparing, etc.). Cases of exceptional genius (Mozart) are not yet understood, but might well involve unusual enhancement of some perfectly common capacity such as short term memory, whose enlargement can demonstrably have a major impact on the achievements of computer programs, including those for musical improvisation. A final chapter (apart from the Epilogue) considers various possible reasons for answering ‘no’ to the fourth Lovelace question, “Can computers really be creative?” All of these arguments apply to intelligence in general, not just to creativity in particular; for example, creativity requires consciousness, and computers cannot be conscious; computers are all syntax and no semantics, so that their symbols and operations are all meaningless to them. Boden finds these arguments inconclusive and points out that even if accepted
they would in no way call for a negative answer to the other three Lovelace questions. The last portion of the Chapter addresses fears that the scientific explanation of creativity would “dehumanize” it, dispelling its beauty and mystery, and our awe. Livingston Lowes’ remark on his own study of Coleridge’s imagery, quoted by Boden and applied to creativity in general, seems a sufficient answer: “I am not forgetting beauty. It is because the worth of beauty is transcendent that the subtle ways of the power that achieves it are transcendently worth searching out.”

Of course, a further mark of a useful approach is that it stimulates useful questions, including any about its own limits or plausibility. We close with a few such questions about Boden’s computational framework itself, and about the role of affective/emotional/aesthetic factors in creativity. (We make no distinctions in this review among such factors.) First, with regard to any purported computational system or structured style of thought, we would add to the four Lovelace questions the question posed by R. Crumb’s creation, Boingy Baxter: “Is dis a system?” For as Boden points out, the theory has no application to a given activity (story-telling, divination, mathematics) unless that activity is engaged in by humans on the basis of computationally expressible principles or rules. But it is also worth emphasizing that this is not guaranteed by the fact that it or its products can be described (as opposed to actually being generated) in a rule-based way. Further, even if some activity is in part rule-generated, it may involve other factors (e.g., chance, randomness, serendipity, as discussed in Boden’s Ch. 9) that are not; and it may be that the interaction of computational and non-computational processes is the core of at least some types of creativity. In such cases a computational description could take us only to the threshold of creativity.

Pressing the issue further, Boingy will also want to ask, “Even supposing this is a system, how do we know we have correctly identified its true underlying principles?” With regard to Boden’s example of Western tonal music, different styles of structural analysis are available. Is there any one correct system, either musically or psychologically, for the structural analysis of this sort of music? Boden points to Longuet-Higgens’ formal description of a conceptual space of musical keys and their relationships (and mentions elsewhere his study of factors involved in expressive performance). Clearly many features, including harmonic structure, melody, rhythm, dynamics, and timbre, will interact in critical ways to produce music. Fred Lehrdahl and Ray Jackendoff’s A Generative Theory of Tonal Music (MIT, 1983) is an admirable attempt to come to grips with much of this complexity; other highly sophisticated systems of musical analysis exist as well. But even if we consider simply the structural analysis of music, how do we know when we have identified the right one for purposes of understanding musical creativity—again, assuming there is a right one? Lehrdahl and Jackendoff’s work derives from “Schenker analysis”, whereas others favor a type of approach pioneered by Schoenberg; and David Epstein (Beyond Orpheus, Oxford, 1987) proposes a kind of synthesis of the two. Of course what one needs for a theory of creativity is not just a coherent way to construe music (Milton Babbett’s term), but a musical analysis that provides a plausible and reasonably detailed psychological representation of the (perhaps implicit) principles in accordance with which it is composed. But as soon as we get beyond the barest essentials, where are we to find confirmation of one type of analysis rather than another? From an empirical point of view a crucial question will be whether one system rather than another leads to distinctive, testable predictions about how composers compose, as well as about what they can compose relative to that system. At present, computer programs can at least give us some information about what does or can happen in simple situations, and about how programs can be made more sophisticated and their
products more interesting. But when we approach the level of real-world complexity, what more can we say than that many extant systems of tonal music analysis are consistent with the same musical “surface” (e.g., the score of a Mozart symphony); and what can we really say about whether that surface was actually produced on the basis of one system rather than another? Similar questions will arise with regard to creativity in most if not all domains within the arts and sciences.

Mr. Baxter will also want to know how one is to identify and distinguish between “exploratory” and “transformative” creativity. One difficulty (noted by Boden) is that there may be a series of steps between the familiar and the apparently revolutionary, so that the latter seems less like a fundamental transformation than simply the latest in a series of tweaks. To return to the musical example, Schoenberg’s change from a traditional tonal system to a twelve-tone one might seem a clear case of fundamental creative transformation. But as Boden indicates, tonal music became very highly chromatic, shifting keys often and with great freedom as regards key relationships, before Schoenberg (and in Schoenberg’s early work as well). So did he just take the final step in the process of relaxing traditional key relationships, and is his twelve-tone system not so radically creative after all? Put the other way round, we could describe in principle any advances by Classical or Romantic or Expressionist composers toward tonal freedom as changes in previous rules or constraints, and thus as a series of transformations. It’s just a matter of expressing changes in compositional practices as changes in underlying rules. Here perhaps the depth of the principles involved will save the day. This remains to be seen, however, and depends at a minimum on finding and establishing in detail the hierarchical properties of the right generative system.

Another option would be to respond that twelve-tone composition does not simply remove traditional constraints, but imposes new rules of its own, e.g., that each tone in the “row” must be sounded before any are repeated—a rule Schoenberg himself honored in the breach as well as the observance. This would raise the more general question of whether transformational creativity requires the introduction of new rules, or whether it can consist simply in dropping some basic constraint(s)—or both. Perhaps it would be best to say that these represent three interestingly different sorts of transformation, all of which make possible what had previously been computationally impossible. But it is far from evident that this would make possible a tolerably clear distinction between tweaks, explorations, and transformations. After all, some additions to the rules will be relatively trivial. So we may find ourselves returning to the criterion of depth, and to the problem of specifying and justifying a particular hierarchical system.

Meanwhile the question would remain of whether the distinction between an exploration and a transformation—or for that matter, between either of those and a combinational novelty—will give us even a rough measure of creativity. Could a Mozartean tweak be more creative than the latest modernist revolutionary transformation (complete with manifesto)? If so, perhaps the criterion of value or interest (duly recognized by Boden) should not be regarded simply as one requirement for admission to candidacy, as it were, with computational distinctions then determining which things of value are genuinely creative and in what way; perhaps creativity is more centrally the creation of value or beauty or interest or the like, and different sorts of computational innovation are important instruments of such creation, and tweaks can sometimes be more valuable and creative than transformations.

The case of twelve-tone music might suggest that yet a further, phenomenological, sort of criterion of creativity is at work, one having to do simply with the way a piece sounds (or a painting looks,
etc.). Suppose for a moment that neither the computational nor the evaluative criterion justifies our distinguishing between a chromatic tweak (of romantic music) and the twelve-tone transformation (of chromatic music), and yet listeners and composers alike tend to experience the later but not the former as a radically creative break with its musical past. Perhaps this is a case in which a major change in auditory gestalt results from a change in underlying principles of composition that are, from a computational point of view, not proportionately great. In fact the question can be applied to any two transitions: possibly the change in psychological impression produced by a new style in music or painting will be much more (or less) radical than the changes in the underlying generative principles that made the new style computationally possible. Of course even if the two sorts of change are found (through psychological experimentation?) to be directly proportional, it would complicate the account of creativity considerably to include the factor of audience response. And if we did, would this set artistic creativity apart from the mathematical and scientific, at least in one significant respect?

But even if we decline to bring in the audience we may want to consider a further psychological criterion on the side of the creator (composer, mathematician, painter, physicist, etc.), namely, the strength of the psychological entrenchment of a set of rules, or the power of the hold it has on the creator’s thought or imagination. The last hundred years have seen a great many “revolutionary” changes, and once the artistic community gets used to the fact of such change, transformative innovations can seem unimpressively easy, even normal. Each new revolutionary is just doing the same thing as everybody else—“making it new”, as they say in the modern theater. And yet in terms of altering computational possibilities to produce something of value, each new “breakthrough” may be just as transformative as that of the first great, and truly revolutionary, revolutionary. But he or she had to do something much more difficult and, we may be inclined to say, more genuinely creative, than those who came after. Only that pioneering effort is, as Boingy would put it, “in the realm of great achievements”.

Further, we note that so long as the older and newer generative styles or systems have something basic in common, there will always be a clear sense in which they are both versions of a more fundamental system. So perhaps one ought really to view them in that way rather than by seeing the second as a transformation of the first. Or perhaps better yet would be to view them in both ways. This is not a merely logical possibility: Anton Webern, the strictest of the “big three” Viennese twelve-tone composers, argued precisely that twelve-tone music was based on the same fundamental principles of composition as a Bach cantata. (Webern, *The Path to the New Music*, Theodore Presser Co., 1953). A computational approach might be able to help us describe clearly the way in which different systems overlap; but how could it tell us whether twelve-tone music was a radical break with the tonal past or, as Webern argued, constituted at bottom a different development of the same root principles and just appeared on the surface to be radically different? And in the latter sort of case, should we regard both the older tonal system and the serial system as creative developments of a more fundamental system—developments achieved neither by the latter altering or dropping some rules of the former (even if describable in that way), nor by either system changing or dropping any of the more fundamental common principles, but in each case through specification of additional, more specific, rules within a shared underlying framework of very elementary principles (e.g., both are built on a foundation of twelve consecutive half-steps, both aim largely at the creation of tension and release, both employ structural repetition, extension, inversion, alteration, etc.). A computational approach can certainly help us formulate any
computational principles involved in dealing with such issues, but cannot by itself tell us which description is musically or psychologically correct.

A further challenge arising from within the computational framework itself is that however we modify an older system to arrive at a new one, we would like our account of the process to reveal the nature of the constraints (if any) on our selection of initial principles to be dropped, altered, replaced, or added, and on any changes we make. The way in which we represent and re-represent the relevant rules will be a critical factor, but presumably that representation will be equally consistent with a range of potential changes without explaining the occurrence of any one change in particular. Then, too, one will want to know whether these constraints are themselves best described computationally—and whether they are themselves activated or guided by further computational processes. This will surely cause Boingy to wonder, “Where does the ascent through levels of constraints stop?”

We close with a question concerned more directly with the apparently non-computational factors that keep cropping up. Boden focuses on what she sees as the cognitive/computational basis of creativity, but she clearly acknowledges the role of further factors—principally, that creativity must produce something of value or interest—and touches occasionally on accounts giving a substantial role to emotion or aesthetic sensibility. But it may be that these other factors should be given a larger or sometimes even central role. Through the persona of a Priestess of certain Higher Mysteries (Diota in the Symposium), Plato speaks of a passion for “(pro)creation” in beauty—whether of beautiful mathematical theories, civic constitutions, epic poems, or philosophy; this, she says, is a great motivator of creativity. But in addition to providing motivation, aesthetic judgments might be involved more essentially in the creative process itself. Poincare (quoted by Boden) describes the conscious “aesthetic sensibility” of the geometer or other creative person as noticing, among the many combinations of ideas proffered by the unconscious mind, certain particularly harmonious and hence potentially fruitful ones. (This is the third stage in Poincare’s theory, that of “insight”.) Whatever one thinks of Poincare’s own development of the notion, other mathematicians and scientists have spoken of being guided by aesthetic perception of beauty and elegance in a theory as a whole or in a particular proof. Einstein’s search for a “beautiful” unified theory in physics is perhaps the most widely known example, but there are many others, including the ultimate creator, who, we are told, paused after each step in his creative work to see “that it was good”. So we have a kind of aesthetic/affective judge at work, as well as a motivator.

The point is more obvious and can easily be pressed further in the fine arts. Aaron Copeland describes much of his own compositional experience as involving at every step a combination of a rather intuitive production of ideas with “critical-aesthetic” judgment. (Music and the Imagination, 1952) The phrase ‘at every step’ suggests a more thoroughgoing and co-equal role for affect and/or aesthetic sensibility. Such a role might be intermittent, alternating temporally with more straightforwardly computational processes. But an interestingly different possibility would be that of a more or less continuous aesthetic monitoring of one’s work and a shepherding of its progress toward an aesthetic goal. Something along these lines is in effect suggested by J.S. Mill (in “What is Poetry?”), who thought of emotion as exercising an organizing and unifying influence over more simply associative processes involved in artistic creativity. Here the emotional factor is involved not just in the “why” but also the “how” of the creative process: without it, the result would not achieve the integration or unity essential to its being a worthwhile work of art. But in principle the
relationship between aesthetic or emotional apprehension and computational processes may sometimes be even closer than that, amounting to a fusion of the two, such that for at least some stages of the creative process it would be wrong to suppose that these factors retain their own identities, psychologically or neurally. In sum, there are a number of ways in which affect or aesthetic judgment might well play a major role, and even the central role, in the creative process itself, and which are not in any obvious way captured by a theory of exploration and transformation of computational spaces.

* * *

None of these questions amounts to an objection to a computational approach, which as Boden claims can be very helpful in thinking about human creativity. They are posed rather as indicating directions for further investigation. Perhaps in the end they are all answerable from a broadly computational perspective. Who knows, perhaps even emotion, aesthetic sensibility and consciousness may eventually yield to such treatment, and Boden’s three types of creativity will fit neatly into a comprehensive computational account of the human mind. But if so, that day is not yet at hand.

Richard Patterson
Katherine Thomas
Emory University