ELSEVIER

Artificial Intelligence 79 (1995) 161–182

Artificial Intelligence

Response

Modelling creativity: reply to reviewers

Margaret A. Boden

University of Sussex, School of Cognitive and Computing Sciences, Brighton BN1 9QH, UK

The reviewers of *The Creative Mind* (henceforth *TCM*) have raised a host of interesting points. Most fall into seven groups: the definition of creativity; the distinction between H-creativity and P-creativity; the role of the social context; the role of evaluation; the four Lovelace questions; specific computational mechanisms used for modelling creativity; AI-models as aids to creativity; and the treatment of music in *TCM*. I shall group my replies accordingly.

1. Definition of creativity

Most of the reviewers point out, quite rightly, that my definition of creativity in *TCM* was vague. To some extent, this was unavoidable: the notion of positive value or interest, which is essential to the concept of creativity, cannot be given any general definition and is open to considerable disagreement in individual cases. But I should have been clearer about the differences between what I called "combinational", "exploratory", and "transformational" creativity.

For instance, I did not make it sufficiently clear (despite my examples of Dickens in Chapter 3 and Mozart in Chapter 10) that mere exploration of conceptual spaces is often, and understandably, counted as creative. Exploration of a space (conceptual or terrestrial) can sometimes show us regions that were previously unsuspected, and boundaries lying in surprising places. In short, it can offer surprises comparable to the surprises provided by transformational creativity. Not all exploration does this, of course. But even simple induction—which, as Roger Schank points out, is the antithesis of transformational creativity—can explore a space with sometimes surprising (and valuable) results (*TCM*, Chapter 8). Most current AI-programs explore rather than transform their conceptual spaces, but even so they may come up with apparently creative surprises. As David Perkins remarks, however, the term applies in its richest sense with the creative process involves changing the rules, not just working within them.

Also, I often used the distinction between a (superficial) tweaking and a (fundamental) transformation in a way that was intuitive rather than analytic.

0004-3702/95/$09.50
SSDI 0004-3702(95)00074-7

Thus I described Kekulé's move from string-molecules to ring-molecules as a change that transformed the relevant conceptual space more fundamentally than the (later) move to rings comprising five atoms, not all of which need be carbon. A chemist specializing in aromatic compounds might see things differently, taking ring-molecules for granted and seeing the basic gulf as lying between (for instance) benzene and pyridine. In general, it would not always be easy to "place" specific AI-processes, such as Schank's heuristics for tweaking explanation patterns, within my scheme. Distinguishing tweaking from transformation is not straightforward, since it depends not only on defining the dimensions of conceptual spaces but also on individuating them. Because of the hierarchically structured nature of many such spaces, this is not a cut-and-dried matter, but requires us to distinguish styles within styles.

In Chapter 4, for example, I described post-Renaissance Western music as a single conceptual space, which took several centuries to explore and which was eventually transformed by the move into atonal music. But this space undeniably consists of a number of sub-spaces (baroque music, jazz) and sub-sub-spaces (Bach, Vivaldi). Likewise, a scientific space (such as chemistry) will usually contain many theories or hypotheses, only some of which are linked in a clear hierarchical fashion. The difficulty of judging the boundaries of distinct conceptual spaces is highlighted by work in the history of art and science, yet such work is itself necessary to help us judge the creativity (both the novelty and the interest) of the ideas under consideration. This difficulty of definition bedevils all those who discuss creativity in general, including AI-workers who seek to model it.

I plead guilty, then, to the complaint that my definition of creativity was problematic. However, I do not accept Scott Turner's charge that I lay too much stress on the hypothesis of representational redescription, or RR (*TCM*, Chapter 4). I do not claim that it is the sole factor underlying creativity, rather that it is a precondition for the construction, exploration, and transformation of conceptual spaces. In a recent publication Karmiloff-Smith herself says as much, and claims that animals lack creativity because they lack RR [13]. Nor do I claim that RR happens only in children (I cite relevant evidence in adults learning to read, or play the piano). But the point is that it appears to happen in all children, and it appears to be essential if the child is to be flexible and imaginative in his/her skilled behaviour. Self-mapping of conceptual spaces is necessary for creativity in general: as Kenneth Haase remarks, if we form and follow a map of the bead-game, our explorations will be more fruitful than if we played the game without it.

Turner says that creativity is not due to a "singular" process of conceptual restructuring, but to conscious, incremental problem-solving processes. That is often true. The types of example Turner has in mind presumably include the scientific problem-solving modelled by Herb Simon and his colleagues [14]. But Grandpa's jokes? And Coleridge's imagery? These are not generated by conscious problem-solving processes. I agree with Turner that they, and creative ideas in general, are produced by normal everyday processes: that is what

Chapter 10 (and much of Chapters 5 and 6) is about. In my account, conceptual restructuring is not a question of a sudden Gestalt-switch, but a focussed (sometimes conscious) change in one or more dimensions of the pre-existing structure. So Kekulé, for instance, did not "discard" his previous space: on the contrary, he transformed it into another, closely related, one. As I argue in *TCM*, if we were to discard our conceptual spaces whenever thinking something new, the novel thoughts would be unintelligible—so would be neither valued nor preserved.

This leads me to hesitate also over Haase's description of a radical conceptual change as one which throws out the lines and landmarks of the previous map and replaces them with new ones. This may be very largely true, but some aspects of the previous space need to be retained if we are to orient ourselves intelligibly. (Some aspects may be "retained" even in being dropped or negated: the obvious absence of an expected dimension can act as a reminder of it.)

Turner says that my definition of creativity lays me open to the devil of determinism. But he does not refer to Chapter 9, which was wholly addressed to this problem. And he and Ashwin Ram (and colleagues) complain that my definition of transformational creativity in terms of apparently impossible ideas leads to inconsistency. However, I explain in *TCM* (Chapter 1) that to say that someone "could not" have produced the idea before can be sensibly asserted only by reference to a specific conceptual space. If we try to apply this notion to the entire resources of the person's mind (or the computer's program), of course we get an inconsistency.

The same applies to computer programs. A program which uses (for example) genetic algorithms, or AM's concept-transforming heuristics [15], is changing its space in significant ways. It could not do so without the algorithms provided by the programmer; but that is just to say that computers—like people—need a set of general computational processes besides the domain-specific spaces that they deal with, to which we must implicitly refer when we speak of creativity. One might want to argue that some of these processes are domain-specific, so should be counted as lying within the conceptual space itself. This recalls Schank's comment that the distinction between rules and meta-rules is highly problematic.

Schank sees my definition as implying that creativity is a matter of degree. I am very happy with that, if it means that creativity is not an all or none matter. But, as explained in Section 4 (below), I prefer to say that creativity is multi-dimensional—which in turn means that the relevant dimensions need to be specified in any argument about the "degree" of creativity of a given idea. For the same reasons, I agree with Haase that the single distinction "creative" is too coarse to describe the inter-representational relations involved in complex generative systems.

Ram's complaint that my definition of creativity presupposes a particular computational account of mental life is perhaps prompted by the fact that I did not sufficiently stress either exploratory or combinational creativity. In fact, I assume that many different types of process, and many different types of representation, are involved. And Ram's comment that new methods of inference and control (such as those used for case-based reasoning) can also lead to creative novelty is well-taken. It is difficult enough to enable computer programs to change their representations. At present, most cannot do so: they explore, rather than transform, their conceptual spaces. My hunch is that it will be even more difficult to enable them to come up with novel—and fruitful—forms of inference and control. But note the term "enabling" here: Ram rightly notes that constraints (in minds or programs) limit what structures can be generated—but they enable generation, too. They are not mere prohibitions, but positive guidelines.

Whereas my definition concentrated on creative ideas (leading me to ask how a single novel structure could have originated from the conceptual resources already present), Perkins broadens the discussions to creative systems—which I see as covering several related ideas and conceptual spaces. I found much of interest in his review, not least the notion of "adaptive" novelties gradually changing a conceptual system. The biological metaphor of adaptation fits many examples (Schank's work on goals and explanation, for instance). But the criteria of "fitness" may suddenly alter. Especially in the arts, the values (selection pressures) may shift unpredictably to very different ones, because the system's environment contains other people, whose judgments—and even whims—may influence what counts as adaptive (see Section 4, below).

Perkins reminds us that striking novelty may emerge from mundane underpinnings. He compares psychological creativity with the emergence of adaptive novelty from blind biological and social processes. Many examples of the emergence of complexity from simple processes are being modelled by work in artificial life (Perkins mentions the Tierra system, for instance [20]), which—despite emphatic denials by some members of the A-life community—overlaps AI in various ways [5]. In *TCM* I cited Howard Gruber's [8] study of the development of Darwin's ideas over many decades, but did not sufficiently stress the gradualness of this development. As Perkins says, we should consider the slow growth of novel ideas, as well as sudden insights and deliberate (conscious) problem-solving. This requires not just subtle psychologizing, but also careful historical work, like Gruber's.

Further helpful remarks from Perkins concern how to describe the inherent difficulty of a given search space. The search spaces of interest, he points out, must be complexly structured, otherwise simple search and hill-climbing would suffice. What sorts of difficulty inhibit us from moving to new, and fruitful, locations within the space, and how can they be best overcome? Perkins' discussion of the rarity, isolation, oasis, and plateau problems are highly suggestive. I see them as related to Haase's intriguing questions about the "thickness" of interesting results within a space. Some measure of the likelihood that a system will be "struck", or "liberated", on certain dimensions or at certain points would be very helpful in understanding creative systems. Possibly, AI-modellers might adapt some of the techniques used in A-life to model fitness landscapes. They would also need to define (domain-general? domain-specific?) heuristics for overcoming the four problems identified by Perkins.

Perkins objects that my definition of creativity does not include evolution, because evolution does not change the rules. It is true that the most general rules, variation and natural selection, remain unchanged. (Likewise, the basic GAs in an evolutionary program remain unchanged). But evolution clearly changes the space, altering the constraints influencing the potential variation and selection at any given stage. There is a difference between the mutations and recombinations that can arise, and the results that can survive, from a bat-genome and a fish-genome. In a highly abstract sense, both bats and fish were within the possibility-space of biological evolution right from the beginning. But their emergence became less improbable as their evolutionary precursors appeared. Similarly, the general resources of the human mind include the possibility of painting "Guernica" and writing "Hamlet"—but one had better learn something about paint, war, guilt, and indecisiveness if one wishes to do either of these things. In other words, taxonomic and morphological categories in biology are comparable to conceptual spaces (as Perkins himself suggests).

2. H-creativity and P-creativity

The distinction between H-creativity and P-creativity is taken up by several reviewers, sometimes in a way that obscures the fact that "H" implies "P". For example, Ram says that H-creativity "typically" results from P-creativity, which implies that occasionally it does not. However, since H-creativity is defined in terms of P-creativity (*TCM*, p. 32), H-creative ideas are a sub-class of P-creative ones. Similarly, Ram says that "most examples of H-creativity must in the end be explained in terms of some individual's P-creative act". For "most", read "all": here, too, we have a necessary truth.

This is not to deny (what is mentioned in *TCM*, Chapter 3) that the individual's H-new idea may be dependent in various ways on the closely-related ideas of other individuals in his/her reference group. In AI-modelling to date, such influences have been ignored (it is difficult enough to model what's going on within a single mind). They are not precluded in principle: AI-models of cooperation, or distributed knowledge-bases, might contribute to our understanding of creativity. Even so, insofar as our question is the psychological query "How did the idea arise?", the answer can only be in terms of individual minds, functioning in varying degrees of independence.

Both RAM and Schank suggest that the best research strategy in AI-modelling is to focus on the everyday psychological processes underlying P-creativity. Analogy, for example, is highlighted by Ram as a wide-ranging phenomenon which contributes significantly to creative thinking, and Schank shows that reminding (*TCM*, p. 241) is a source of novel ideas too. I agree. Indeed, this was the central theme of Chapter 10 on "Elite or Everyman?". Perkins' psychological experiments cited in that chapter provide illustrations of the role of everyday processes in creativity. Not surprisingly, his experimental subjects did not come up with any ideas that will go into the history books: experimental studies, barring extraordinary coincidences, are confined to P-creativity.

It does not follow, as Schank seems to believe, that AI-modellers should ignore examples of H-creativity. Schank claims that an understanding of creativity will come from modelling normal P-creative insights, rather than from providing rational reconstructions of scientific breakthroughs, as Simon's group do (Chapter 8 of *TCM*; [14]). Likewise, he complains that I forget my own P/H distinction, in mentioning many H-creative cases (as well as more mundane instances) throughout the book. But H-creative examples are necessarily P-creative too. Moreover, they are intriguing to AI-researchers and psychologists, as also to the general reader. To attempt to model, or write a book about, creativity without addressing any examples of H-creativity would not be persuasive. Where Schank (like Ram) is correct is in saying that, in modelling the mechanisms of creativity, we should focus first on everyday psychological processes shared by us all.

If this (in conjunction with the input of the "historical" data) results in a convincing rational reconstruction of famous H-historical cases, so much the better. But in the first instance, at least, that is not necessary. As Seymour Papert remarked years ago, AI should avoid the superhuman human fallacy: if one cannot model Newton or Shakespeare, it does not follow that one cannot model significant aspects of the minds of lesser mortals.

Questions about the "how" of H-creativity primarily concern individual minds, as we have seen. But if our question is historical-sociological: "Why was the idea greatly valued after it had arisen?" then social factors are not only relevant, but predominant. Moreover, these are not systematic, but highly contingent. That is why, as stressed in *TCM*, we cannot hope for any systematic explanation of H-creativity—whether psychological, sociological, or historical. There are three reasons for this: many non-psychological facts are relevant, not all of which can in practice be discovered; these facts are not all of the same type, so do not fall under any unifying theory; and the values used to ascribe the honorific "creative" to novel ideas differ across societies, and change unpredictably within social groups.

This does not mean that psychologists who study H-creativity are wasting their time. Many studies, most recently those of Howard Gardner [7], have found evidence that certain motivational and personality factors are typical of H-creators. To put it bluntly, most H-creators are so determined (not to say obstinate) and self-directed (not to say selfish) that they exploit, and often damage, those closest to them in highly unfortunate ways. As remarked in *TCM* (Chapter 10), motivation and self-confidence have at least as much to do with someone's ability to H-create as cognitive factors do.

Perkins stresses the importance of psychodynamics, which cognitive science has largely ignored. To be sure, the question of how it was possible for a specific creative idea to arise in someone's mind is distinct from the question of what motivated the person in coming up with it—and sticking to it. It is thus reasonable for a cognitive psychologist to decide to focus on the first question rather than the second. But, as Perkins reminds us, the second requires attention also. (Computational theories of motivation are still thin on the ground. The most interesting, I think, is that developed by Aaron Sloman and Luc Beaudoin [1, 24]—but even this theory cannot significantly enrich our commonsense intuition that H-creativity depends on strong motivation.)

We also need to answer Perkins' question as to why some people (Turing, for example) seem to H-create repeatedly over many years. As I remarked in *TCM* (p. 36), a P-creative idea which misses being H-creative only by some historical accident (missing the priority by a few weeks or months) is very good grounds for a "bet" that the person concerned will come up with H-creative ideas later. (Roger Lustig's comment that "Turing's work [for the award of his Fellowship] was H-creative for all practical purposes" thus merely restates my point.) However, just what it is which enables someone to H-create over and over again is highly obscure.

Ram suggests that in such individuals the usual search space may be interestingly different or expanded, so enabling creative thought using the very same mechanisms that usually generate more mundane ideas. This is intuitively plausible. Perhaps the difference is partly a matter of extra constraints being developed within the widely-known constraints: such additional distinctions would allow for additional possibilities in exploring, and transforming, them. And perhaps this is a large part of what is going on when, as we say, the creator is spending years "acquiring domain expertise"—for expertise is necessary, though not sufficient, for constant creativity. (One example might be the musical "rules" characteristic of certain areas of Bach's music mentioned in *TCM*, Chapter 5.)

As for P-creativity in general, there is considerable psychometric (and even some neurophysiological) evidence suggesting that specific ways of using and generalizing concepts are characteristic of creative individuals [6]. But studies of motivation, personality, and conceptualization can only take us so far: they can tell us nothing about why certain people, at certain times and places, decided that an H-novel idea was not only novel but also valuable.

Because of the unclarities involved in identifying H-creative ideas, Haase suggests that we replace the binary P/H distinction with several analogues, varying with the size of the community over which the novelty is defined. This is a helpful suggestion, which can contribute to our psychological as well as our historical understanding. Detailed historical studies can sometimes distinguish distinct groups and sub-groups within what to the undiscriminating eye looks like one scientific community (see [22] and Section 3 below).

An interesting example of a potentially creative "community" suggested by Haase is the human–computer duo. As Haase points out, AM's apparent creativity is dependent on human judgment in various ways. Not only did Doug Lenat, in writing the program, insert the generative and evaluative heuristics he expected to be most fruitful, but he frequently influenced AM's ongoing performance by "naming" some of its novel structures, thus leading the program to focus on them rather than others. Lenat reports that AM (and even EURISKO) have come up with a few valuable H-creative ideas (minor theorems about maximally divisible numbers, for example). But—at least in the present state of the art—the more the human "partner" is involved, the more this is likely to happen. Programmed "agents" could help human creators in various ways, sometimes leading to the generation of ideas that would probably not have arisen from either human(s) or computer alone [4].

Haase focusses on the embedding of creativity in various social contexts, and on motion across the boundaries between different spheres of novelty. I agree with his conclusion (arrived at by some other reviewers also) that we need far more words than "creative" and "novel", to describe innovations in such a complex social-conceptual space. To arrive at a more powerful and discriminating vocabulary for describing creative change, we shall need to study a wide range of examples drawn from different conceptual spaces, each mapped (defined) in detail. It is an open question to what extent the same vocabulary will apply to all cases, but my bet would be that significantly different distinctions will need to be drawn in differing domains.

3. Social context

Several reviewers stress the importance of the social context—not only in the crucial evaluation phase, which decides whether the new idea is accepted and promulgated, but also in the generation phase itself. Ram, for instance, makes the important point that innovation often arises when ideas from one culture are applied in another. There are many historical examples of outbursts of creativity when new cultural groups (such as traders or refugees) come into a settled society.

The generative importance of the social context is also evident if we consider peer-groups of expert individuals. Ram mentions, as examples, Maxwell's location in Cambridge and Faraday's links with continental physicists. A superb essay on the role of such social contexts in both the generation and the evaluation of creative ideas is Simon Schaffer's "Making Up Discovery" [22]. Schaffer shows that the social construction of a scientific "discovery", even of something as apparently unproblematic as the discovery of dinosaurs, may take years of theoretically informed (and nationalistically biassed) negotiation within the relevant scientific community.

Ram points out that some social factors can be figured into the account without our being required to produce a computational model of culture. For instance, to produce computer modelling equivalents of the injection of unfamiliar ideas from other cultures would require us to enlarge the data-base by adding a different conceptual space, or additional entries to the semantic network available to a connectionist system. But, for this to be fruitful, the generative processes already available to the model would have to be powerful (general) enough to be able to work on these newly-added structures as well as the "native" structures already provided. The more different the cultures in question, the more difficult this would be.

The simulation of the sorts of peer-group negotiation described by Schaffer would however be more difficult. It involves not only the cooperative interchange of ideas to aid generation and modification of the nascent scientific discovery in question, but also the cooperative—and, to some degree, competitive—evaluation of the ideas that result. The role of evaluation is currently underplayed in AI-models of creativity, often being left implicit rather than made explicit as a separate process within the model. To simulate the social agreements on evaluation that are so important in human creativity is, as yet, beyond the state of the art. So too is the phenomenon, also mentioned by Ram, of someone's taking a very long time to develop a creative idea. Ram mentions Maxwell's taking nine months to develop the notion of electrostatic action, and one might add Darwin's idea of evolution by natural selection, which took almost a lifetime to mature [8].

4. Evaluation

The problem of evaluation in assessing—and modelling—creativity is put in a nutshell by Ram, who remarks that the distinction between an idea's being novel and its being interestingly novel is problematic.

As suggested in Section 3 above (and in Chapter 3 of *TCM*), evaluation is a social matter that can be affected by all sorts of intellectually extraneous factors, such as nationality, fashion, rivalry, and commercialism. This fact is well illustrated by Gary Taylor's [25] scholarly work on "the Shakespeare industry", cited in *TCM* (p. 34), and also by Schaffer's work on scientific discovery, mentioned above.

One may, of course, recognize the strong social determinants of evaluation while also adopting certain evaluative criteria oneself. That is, one does not necessarily have to accept relativism in agreeing that "creative" is, to a large extent, a socially bestowed honorific. (I am therefore puzzled by Lustig's complaint that I insulted Taylor by calling him a relativist. I did not use the term relativist as an insult, nor name Taylor as a relativist. I deliberately spoke of "the relativist" in general, in pointing out that relativist critics of literature or science cannot use the term "creative" in its usual sense, because they accept only descriptive, not evaluative, differences between scientific theories, or works of art.)

Another difficulty identified by Ram lies in the fact that (as argued at length in Chapter 10 of TCM) the psychological processes responsible for creative thinking are not unique to special individuals, but are possessed by us all. It is therefore unclear how to distinguish those thoughts which are creative from those which are not—in a given individual, irrespective of past history. In other words, even assessing the degree of P-creativity is problematic.

This is largely correct. But I prefer to avoid speaking of the "degree" of creativity, since to do so implies a continuous spectrum along which individual thoughts are to be ordered. To the contrary, a main theme of *TCM* is that creativity is multi-dimensional. Some relevant dimensions are easily conceived of as continuous, others are not: certain structures simply cannot be generated from particular (untransformed) conceptual spaces, for instance. But even if all the dimensions were continuous, they are to a large extent orthogonal. Ideally, therefore, one should gloss attributions of creativity in terms of the specific respect (or respects) in which the thought is creative. This would involve detailed comparisons that would place the putative novelty, relative to a variety of other ideas. Unless one is in the business of handing out Nobel prizes or the like, it does not matter if one cannot decide which of two ideas is "more" creative. Indeed, it may be misleading (as well as socially unproductive) to make the attempt.

Haase argues that the real work facing AI-modellers (and psychologists) of creativity lies in thinking carefully about the boundaries and interactions of conceptual spaces, and requires terms more precise (and less loaded) than "creative" or "novel". I fully agree that we need a more precise terminology, and many detailed investigations of specific examples. I cannot agree, however, if by "less loaded" Haase means "non-evaluative". As I argued in Chapter 3 of *TCM*, positive evaluation is an essential aspect of our concept of creativity, and as such is unavoidable in any reference to or study of it. If we were to study merely the generation of new ideas (granting that "new", also, can be problematic), we would not be studying creativity, as this term is normally understood. Moreover, we should still require evaluative criteria (of mathematical interestingness, for example [15]) to sort out the wheat from the chaff—an evaluative distinction, if ever there was one.

This is not to deny that, as Schank remarks, more discussion concerning the criterion of value would have been helpful in *TCM*. There are, as he says, many kinds of value. His list of values focusses on explanatory creativity (in everyday life, as well as in science), and he says that we need a richly detailed theory of the kinds of goals and evaluative criteria that may inform this type of creativity. I have no quarrel with that. But there are other kinds of creativity, which have nothing to do with explanation, or even with "reminding". A musician composing a sonata, or improvising a jazz-blues, is doing something which Schank's account does not cover.

Also, we need to remember that the "goal", even assuming that one can sensibly identify one, typically becomes more well-specified as the creative thinking proceeds, largely as a result of a dialectic interaction with the successive stages of thought. Lenat's AM (which, by the way, is not a model of explanatory creativity) attempts to model this feedback-led specification to some extent, but AI-modelling is not yet adequate to capture the subtleties of this important aspect of human creativity (well described in informal terms in [9]).

To cap it all, we have to remember that evaluation can vary with time and place. Any AI-model able to recognize creativity in its own and others' ideas (thus satisfying the third Lovelace question) would unavoidably be using evaluative criteria open to change and challenge of unpredictable kinds. It could be accepted by people as a satisfactory model of creativity—or even as a useful tool for coming up with creative ideas—only to the extent that the people concerned shared the same values within the domain concerned.

Haase's reminders of the cases of Mendel and McClintock are apt. The interest in these cases lies not only in how the new ideas arose, but in why (as we now judge) they were not recognized for many years as being valuable. The answers are complex, involving not only the expected resistance to maverick scientific concepts but also (for example) lack of publication in widely-accessible journals, and prejudice against women. In this context, Haase mentions the "possible injustice [which] resonates powerfully with our own most basic experiences of not being understood". This remark reminds us of the exceptionally strong motivation that must drive H-creators—partly to acquire the necessary expertise, and partly to persist in their own evaluation of their maverick ideas.

As Schank reminds us, the person who encounters someone else's creative idea may evaluate it largely because they have already "internalized" the conceptual spaces and evaluative criteria of the creator. In the haiku-program described in *TCM* (Chapter 7), for instance, the "meaning" is provided by the interpretative skills of the human reader, not the (non-existent) semantic powers of the automatic writer. This is partly why the third Lovelace question is relevant to the topic of computer creativity. To interpret sensibly, and to value acceptably, both require complex cognitive capacities.

The impossibility of a scientific theory of evaluation in general (because values can change for unpredictable, and sometimes seemingly perverse, reasons) does not rule out the possibility that some values may be very common, or even universal. Criteria of symmetry and simplicity, for example, are commonly used in the valuation of works of art or scientific theories. To be sure, some degree of asymmetry and disorder are often welcomed (think of Robert Herrick's poem on "A sweet disorder in the dress . . ."), and baroque complexity may be preferred over classical order. But that symmetry and simplicity are so often judged to be relevant may be grounded in very general facts about human information processing. Similarly, some people argue for the "biophilia hypothesis", claiming that humans share universal aesthetic preferences grounded in our evolutionary ancestry: for instance, certain types of landscape, and even tree-shape, appear to be favoured by people from vastly different societies, living in vastly different terrains [10]. Identification of shared evaluative biasses, however, is hampered by the fact that these are only tendencies—which may be swamped by strong cultural influences. In other words, even if they exist at all they are not strong enough to make us insist that they should be included as a matter of course in AI-models of human creativity.

5. The four Lovelace questions

The four Lovelace questions itemized in Chapter 1 are:

LQ-1. Can computational concepts and models help us to understand human creativity?

LQ-2. Could a computer at least appear to be creative, its performance being comparable to human creativity?

LQ-3. Could a computer appear to recognize creativity (in its own ideas, or those of other systems)?

LQ-4. Could a computer be "really" creative (as opposed to merely producing apparently creative performance whose originality is wholly due to the programmer)?

I am therefore bemused by Turner's claim that I do not address "the question of whether computers can be creative". The last three of the four Lovelace questions address this issue, and LQ-2 and LQ-3 are discussed throughout the book. As for LQ-4, this is discussed at length in Chapter 11. I "sideline" it in the preceding chapters because, as a philosophical rather than a scientific question, it lies outside the main issue (LQ-1): whether the concepts and processes underlying apparent-creativity in computer models can help psychologists to formulate helpful theories about human creativity.

A further reason for delaying consideration of LQ-4 until the final chapter was that it is an extremely complex question, involving at least four types of philosophical argument. As outlined in Chapter 11 of *TCM*, these arguments concern brain-stuff, consciousness, semantically empty programs, and the practical implications of human community. Each of these arguments is controversial, and they are to a large degree independent of each other. Turner seems to confuse them, in claiming that if one accepts the naturalist view of "mind as computer", and if computers cannot in theory be creative, then the human mind must be more than just a physical mechanism.

Functionalism—or "mind as computer program, brain as computer"—is just one of many naturalist views of the mind/brain, and is philosophically controversial. Many people argue, for example, that consciousness cannot be captured in functionalist, or computational, terms. In *TCM* (pp. 278–281) I argue that although this may perhaps be true, it is not enough to render computational theories of human psychology, including creativity, irrelevant. The "perhaps" is needed here not merely because of our neuroscientific ignorance, but primarily because we do not yet sufficiently understand the concept of consciousness at the philosophical level. In view of this unclarity, we cannot be sure that if computers cannot be conscious (and if consciousness is necessary for creativity) then the human mind/brain must be partly supra-physical.

Some naturalist views of mind/brain do not accept "mind as computer program". To take just one example, well-known to the AI-community: the philosopher John Searle [29] sees the mind as a natural product of the brain, and claims that the (as-yet unknown) properties of brain-stuff which underlie mental processes are not possessed by metal and silicon. I show in Chapter 11 (pp. 271–273) that Searle offers no argument for this empirical hypothesis (which may conceivably be true), and that his associated claim—that it is intuitively obvious that neuroprotein can support mentality, whereas metal and silicon cannot—is certainly false. However, neither Searle nor I wish to claim that the brain is more than just a physical mechanism. Moreover, Searle explicitly allows that some computer could, in principle, produce performance precisely the same as the behaviour of people. This covers all cases of human creativity, including Shakespeare and Beethoven. Whether Turner would take this to be an argument "for" or "against" computers being creative is unclear, as he does not distinguish the four Lovelace questions.

The fourth type of philosophical argument against regarding computers as creative (*TCM*, pp. 281–284) has nothing to do with whether or not the mind/brain is "more than just a physical mechanism". Rather, it concerns an essentially moral-political decision about whether to accord moral and epistemological status to computers, as we all do to other human beings (though Nazis did not do to Jews, nor slave-owners to slaves), and as some humans do also to some animals. Just as the case for treating animals with respect or compassion does not depend on, or imply, the view that animals are more than physical mechanisms, so the decision to accord some moral or epistemological status to computers would not imply this either. Conversely, someone who refuses to regard computers as creative, so as to avoid certain practical (moral) implications, is not thereby claiming that human brains are not purely physical mechanisms.

Whereas Turner accuses me of ignoring the question whether computers can be creative (let us assume he means "really" creative, as in LQ-4), Schank complains that I have wasted too much space on it. He says the question has no practical importance "unless in a philosophy class". Science-fictional scenarios apart, I agree. This is why, as explained above, I postponed LQ-4 until the final chapter. But it does not follow that it should not be discussed at all. Quite apart from its intrinsic interest (and the regularity with which it is raised by people outside AI), LQ-4 needs to be discussed in order to demonstrate its difference from the first three Lovelace questions. In my experience, even AI-professionals sometimes confuse it with the others.

Schank himself seems to assume that, because the evaluative aspect of attributions of creativity is inescapably subjective (there being no "acid test" for creativity), no answer to LQ-4 can be given. But the evaluation of creativity is irrelevant to some philosophical arguments about it (such as the "non-human" argument, *TCM*, pp. 281–284). In resisting Schank's dismissal of LQ-4, I am not saying that I agree with Searle (whose views Schank rejects). In Chapter 11 of *TCM* I make it clear that I do not. The point, here, is that LQ-4, and Schank's claims relating to it, require philosophical consideration.

I agree with Schank that people do not always fully understand each other, so that to insist on perfect understanding in an AI-model is unfair. I agree, too, that (as noted on page 278 of *TCM*) understanding is not an all-or-none matter. But it does not follow that a purely behavioural criterion of understanding is acceptable. Schank tells us that his educational programs are felt by the students to understand their questions to some degree. He concludes that his programs clearly do embody a degree of understanding. This does not follow. His students could be wrong, in two different ways.

First, they may mistakenly attribute capacities to the user-friendly system which it does not, in fact, possess (but which some more advanced program might possess). Early mistakes of this kind with respect to ELIZA were legion, and I have argued elsewhere [2, pp. 463–472 and 496–497] that user-friendliness should not be overdone, lest this sort of over-trusting attitude be uncritically adopted by non-specialist human users.

Second, Schank ignores all the philosophical controversy concerning the Turing Test. This (thankfully) is not the place to rehearse these arguments. But it must be said that most philosophers (myself included) do not accept the Turing Test as an adequate criterion of understanding. We would be mistaken, for example, in attributing understanding to a computer whose passing the Turing Test was due to a giant look-up table, as opposed to a set of systematic and interacting generative processes. (Similarly, a novel notion unconnected in any way with the domain in question could not properly be termed a "creative" contribution to that domain—even if, as a matter of fact, it prompted someone else to see the connections.) The fact that Schank's students can be "fooled" into believing that his program understands them is therefore irrelevant.

Schank might counter this argument by pointing to his claim that AI-models of thinking (understanding, creativity) need not fit our intuitions about how humans achieve these things. Up to a point, I agree: some of our intuitions may be wrong; and in any event, Martians (and computers) might be different. But that is not to say that the internal mechanisms are wholly irrelevant. One example showing they are not is the giant look-up table, mentioned above. Another is the geometry-program discussed in *TCM* (pp. 104–110), and mentioned by several reviewers. This program is not creative in proving the "base-angles equal" theorem, because its method involves no transformation of the conceptual space of geometry (nor any analogue of the visual understanding of 2D- and 3D-space possessed by Euclid and Pappus).

In his final paragraph, Schank says "the right question isn't whether a mechanistic explanation of creativity is possible, it's how". But "the right question" for whom? My book was aimed not only at AI-workers and sympathizers, such as computational psychologists, but at the far wider audience of people who are convinced that computer models must be utterly irrelevant to the psychological understanding of human creativity. That is, I was not preaching only to the converted. I therefore had to argue for controversial philosophical positions which many AI-researchers take for granted, perhaps without realizing that they are doing so.

When Schank says "If progress is to be made in modeling human capabilities such as creativity, the focus must be on research strategies in AI, not on philosophical discussions", I agree with him—provided he means "the main focus". That is why Chapters 5–8 of *TCM* focus specifically on AI work, and why Chapters 1–4 offer definitions and examples to be used as background for considering the relevant AI-research. But the philosophical discussion in Chapters 9 and 11 is necessary too. (One point made in Chapter 9 is that, contra Schank, science does not consist only in finding predictive "laws of nature": it seeks to make natural phenomena intelligible, by showing how they are possible, how they are structured, and how they are related.)

6. Specific computational mechanisms

Many examples of computational mechanisms are mentioned in the reviews, and in *TCM* itself. Here, I shall focus on just a few.

One promising example is case-based reasoning, discussed in Ram's review (and used by Turner in his work on story-generation). I see this work as lying within the tradition of heuristic problem-solving described by George Polya [19]. Ram focusses on how we can—and how AI-models might be able to—analyze and reformulate problems, and how we can adapt old solutions to new problems. This adaptation may require us to combine past solutions or experiences, which in turn demands powerful methods of storage and access.

To illustrate the subtle problems facing us here, let me give an example. Some years ago, I was marking a pile of philosophy dissertations. Several were on a highly abstract, and much-discussed, philosophical topic: the nature of representation. The essays stated the problem, defined the jargon, laid out the competing philosophical positions, provided citations to the specialist literature, and attempted a summary evaluation of the various positions. So far, so good—but also, so boring. Then I came upon a dissertation formatted like a theatrical script. There were four characters: Winnie-the-Pooh, Piglet, Rabbit, and Christopher Robin. The characters chatted to each other in plain language and simple sentences, mentioning various homely topics—honey-pots, and the like. In fact, without this being explicitly stated, they were discussing the nature of representation, each one putting forward a coherent philosophical position and defending it against the objections of the others. And, amazingly, the positions "fitted" the different personalities of these familiar fictional friends (Pooh, for instance, was the bluff commonsense realist, Rabbit the devotee of formalism). Naturally, the author got a First. But more relevant here are the mechanisms of indexing, access, comparison, and control involved in noting—and constructively developing—the analogy between A.A. Milne's nursery-world and a convoluted philosophical controversy.

Ram's insistence that case-based reasoning normally involves multiple processes interacting is well-taken. However, I suspect that this AI-work is more useful in theory than in practice. In other words, Ram's work can help us to understand the sort of psychological processes going on in my student's mind when he wrote his dissertation. Some future case-based reasoner might even be able to mimic this particular example, if the relevant conceptual dimensions were specifically provided by the programmers. (As Haase points out, current "creative" programs, such as BACON, normally have the creative insight implicitly coded in the initial formulation of the problem.) But whether an AI-program could generate such creative comparisons spontaneously is another matter, about which I am highly sceptical. The knowledge-base involved is just too great, and too diverse.

This is not to say that Ram and other AI-researchers are wasting their time. As I noted in Chapters 9 and 11, science aims for understanding, not—or not necessarily—for prediction, or even for detailed post hoc explanation. If the processes of conceptual comparison and control implemented in AI-programs help us to see, in general terms, how human performance is possible, that is a considerable advance. This is especially true of areas such as creativity, where many people assume that no scientific account can ever be given.

The same caveat applies to Schank's work on explanation patterns (another form of case-based reasoning). He mentions some helpful general heuristics for adapting familiar patterns to less familiar cases. But it's not clear to me how these heuristics might link Pooh and Piglet to distinct theories of representation. He admits as much himself, in saying that we are not seeking to reverse-engineer actual cases, but merely looking for AI-models that could account for creativity in principle.

I am uneasy at Schank's claim that probably 90% of P-creativity is the creative explanation of ill-understood or anomalous phenomena in one's world. It was to remind readers of the very wide range of (explanatory and non-explanatory) human creativity that I deliberately referred in *TCM* to a host of dissimilar examples: music, poetry, novels, needlework, painting, choreography, sarcasm, car-mechanics, even advertising. Schank's approach is well-suited to the explanatory cases, but not to the others.

I am uneasy, too, at Schank's account of my discussion of connectionism in Chapter 6. Associative creativity is, in my terms, combinational creativity. It is thus different from the exploration and transformation of structured conceptual spaces, and even from analogizing which relies on the comparison of conceptual structures.

I do not claim that it cannot be modelled by symbolic systems, nor that such systems cannot model unconscious processes (natural-language processors, Schank's included, do that constantly). But I do think that associative creativity—such as we see in the examples selected from "The Ancient Mariner" discussed in Chapter 6—is more naturally modelled in connectionist terms. That is, I am not appealing to connectionism "as a general-purpose net to catch thoughts and ideas that apparently fall through the cracks of symbolic rule-based models", but as an approach whose features seem to reflect what may be going on in some—only some—examples we normally regard as creative. (I think it interesting that there is a significant resemblance between the intuitions of a poet (Coleridge) and a literary critic (Livingston Lowes) concerning the nature of associative memory, and the behavior of connectionist systems.)

Of course, connectionism is not magic. Moreover, as Schank points out, the human designers of connectionist systems decide on the features or sub-features to be coded. In NetTALK, for example, the phonetic Wickelgren features coded by the input-layer provided, in effect, an analysis of the task-domain that enabled successful learning to take place [21]. And current connectionism is crude. I agree with Schank, when he says that we might benefit from implementing ideas such as Minsky's [17] society in mind.

Also, as I argue in Chapter 6, we need to understand the interplay between associative and structuring processes in creativity—for example, the imagery and the stanza-structure of "The Ancient Mariner". In other words, we should not try to force all examples of creative thinking into one computational box. Schank's application of the (non-connectionist) Birnbaum–Collins [27] model to Kekulé's insight about benzene is helpful, but that is largely because the Kekulé case was relatively goal-directed. I do not see how this approach would apply productively to the examples of Coleridge's imagery cited in Chapter 6. In short, I agree with Schank's comment that connectionism, as we know it now, is suitable for relatively low-level processes and pattern-recognition, but not for creative insight. But I wouldn't (and others normally don't) call Coleridge's images "insights"—a word that suggests a much more goal-directed, problem-solving approach.

Haase remarks that we need a better understanding of how to "liberate" our programs from the confines of the initial representation we gave them. This is perhaps the most difficult question of all. Heuristics for tweaking and transforming conceptual spaces have been implemented by a number of AI-workers, but the "transformations" are not fundamental, and in any event are usually prefigured in the sort of way that Haase (and Schank) complains of. There are two problems here: how to make (sometimes fundamental) transformations, and how to evaluate them, once they have been made. Arguably, fundamental transformations are more likely to arise with the use of genetic algorithms than with structure-specific heuristics. But AI-researchers will hope that the selection at each generation would be done by the program, not the human being.

Moreover, even if the system is self-evaluating, we shall probably find that plausible-generate-and-test (as opposed to random mutation) is needed. And to incorporate that, we—or the GA-system itself—have to decide what counts as plausible. In *TCM* (p. 213) I point out that the ARTHUR poem quoted in Chapter 1 was, in effect, a skit on the use of genetic algorithms. But the human poet was using plausible generate-and-test: because he swapped whole phrases (not individual words) from line to line, all but one of the final lines were, though strange, potentially meaningful.

7. AI-models as aids to creativity

In *TCM*, as Ram and Schank point out, I was concerned more with how creativity is possible than with what guidance can make it more probable. Some relevant remarks were made in Chapter 10, and Seymour Papert's LOGO-led stress on exploration and self-analysis was noted. But I said nothing, there, about how current AI-models of creativity might be used to help the human creator. I have recently argued that such models could provide programmed "agents" for creativity [4].

Such agents (program-parts able to function relatively independently) could help the human to learn the constraints involved in the relevant conceptual space, and also help him/her to explore that space. They could suggest and identify differences, of varying degrees of subtlety, between familiar ideas and novel ones. The evaluation would be the responsibility of the human being, as is the case for example in evolutionary computer graphics [23,26]. Together, human and machine can create novel structures that the human could not have generated alone, and which in some cases are valued sufficiently to be sold in art-galleries.

Sometimes, as Perkins notes, creative solutions are hard to find because the way in which the problem is presented masks those solutions in specific ways. This is illustrated in *TCM* (pp. 103–104) by the two-houses-and-string example, which shows that—in this case—visualization tends to block the (exceedingly simple) arithmetical solution. But visual imagery can sometimes be very helpful, as in Kekulé's case (see especially Chapters 4 and 10). As Ram remarks, the rarity of visual representations in current AI-models of problem-solving is unfortunate. If we are to use AI-agents as aids to our own creativity, agents capable of building, exploring, and transforming visual representations will be needed.

In recommending AI-models as ways of helping or teaching creativity, however, we must not rely only on intuition and anecdote. As I note in *TCM* (Chapter 10), some of Papert's confident claims about the pedagogical uses of LOGO have been questioned by psychologists on the basis of careful experimentation [18]. So Schank's—intuitively plausible—claim that his programs will enable (most or all) students to become (consistently) more creative must be checked by psychological experimentation. AI-workers need to cooperate with empirical psychologists, not seek to supplant them. (More generally, people wanting to understand creativity need to learn from a variety of psychological approaches, which though highly dissimilar may complement each other in significant ways [3, Introduction].)

8. Music in *TCM*

One review, by Lustig, is largely devoted to *TCM*'s treatment of music. Lustig's review is very different in tone from the others, and much less helpful. As a professional musicologist, he is understandably irritated by my self-declared musical amateurism. His irritation, however, has led him into a less than careful, and very much less than generous, reading of my book—and also into an ill-tempered attack on a researcher whom I praised. There are many points at which he fails to take proper account of the spirit, and even of the letter, of my text. In addition, he ignores the rhetorical constraints involved in writing for a general audience.

His remarks on the history of Romanticism, for example, are irrelevant. It is precisely because I know this history is diverse and complex that I was careful to say of the two popular attitudes, not "They are called the inspirational and the Romantic", but rather "I call them the inspirational and the romantic" (p. 4). My use of the lower-case "r" was deliberate. So too was my decision not to refer to the disputes between individual Romantics, who therefore "are never named in the book". I was concerned with today's widely popular view that creativity is a gift restricted to an elite. Although this view has roots in the Romantic movement (which is why I used the word), it would have been inappropriate to offer a lengthy historical disquisition on Romanticism.

It would have been inappropriate also to pepper the text, here and elsewhere, with scholarly references. I chose to keep references to a minimum because I was writing for the general public, for whom the paraphernalia of footnotes and references might be off-putting. Since I took pains to mention as wide a range of examples as possible, the "complete" endnotes that Lustig demands would have filled many pages. Perhaps Lustig believes such rhetorical constraints should not be accepted, that professional academics should write only scholarly tomes for the delectation of other professional academics. If so, I profoundly disagree. Ideas that are potentially of interest to the general public should be communicated to them, not kept within the ivory tower of academia.

Lustig's lack of attention to the rhetorical demands I faced also underlies his charge that I am "determined to plead Mozart's case at any cost". I am not. I deliberately referred to Mozart (as also to Kekulé) in almost every chapter, in order to remind the reader that there are many different sorts of questions to be asked about any individual H-creator. For this purpose, any well-known composer (or scientist) would have done. I chose Mozart for five reasons, which concern the rhetorical structure of my book rather than the specifics of his music.

In my experience, Mozart is the composer most commonly mentioned in discussions (written or verbal) of creativity in general, so is likely to cross the reader's mind anyway. Second, the interpretation of "Amadeus" mentioned in Chapter 1 explicitly glossed Mozart's creativity in "inspirational" terms, which was useful for my argument. Third, Mozart is someone whose work has been very differently valued at different times, which supports my remarks (in Chapters 3 and 10) about the social construction of H-creativity. Fourth, he was a child prodigy, which is relevant to various questions about nature and nurture raised in Chapter 10. And fifth, an oft-cited "letter" purporting to be in Mozart's hand is probably a forgery, which enabled me to show (in Chapter 10) how people who do plead his case may muddy the waters in respect of the nature of creativity. (I forebore then, and I forebear now, to name two recent books—one of which has received enormous attention—that have taken this "letter" on trust: I have no wish to point out other writers' mistakes just to make them look foolish.)

As for the comparison between Mozart and Haydn on page 253, Lustig's criticism again betrays careless reading. As his own quotation shows, I explicitly flagged the fact that this was not my own judgment ("I have heard some musicians say . . ."), and I took care also neither to endorse nor to deny it, since I am not competent to do so. This purpose of this passage was to make the point that such a comparison (whether justified or not for this specific pair of composers) is intelligible: in other words, that much of what we call "creative", and admire greatly, is—in my terminology—exploration rather than transformation. I have acknowledged (in Section 1, above) that I should have made this general point clearer earlier on in the book. But pouring scorn on my lack of musical expertise about Mozart is inappropriate, since I took care not to claim any.

Since I make no claim to musical expertise, I cannot comment on most of Lustig's (almost equally scornful) criticisms of Longuet-Higgins. I suspect, however, that Lustig is forgetting that Longuet-Higgins' aim is to offer a psychologically plausible analysis of harmony. Longuet-Higgins is well aware of the received theory of tonality, in which fifths (and octaves) can identify all the intervals. But this is a purely analytic identification. Interpreted psychologically, this analysis would mean that in order to recognize a major third (which he says is essential in understanding certain passages), one would have to go up by a fifth four times, and then down two octaves. He therefore adds major thirds as a dimension of his (psychological) harmonic space.

As I point out in *TCM*, Longuet-Higgins' work on Bach is not intended as a theory of musical creativity, but of musical perception. In my terms, it counts as "mapping" a representational space. Lustig implies that, as such, it is irrelevant to creativity: "What are we to do with this particular map? . . . How will we arrange it to be musically creative?". But this is to ignore the exploration of a (non-transformed) musical space. For the purposes of computer modelling, someone has to define the space before anyone can explore or transform it. Longuet-Higgins' work has been used (for example) as the basis of a model of jazz improvisation described in *TCM* [11,12].

The charge that Longuet-Higgins' reference to first principles is "bluster" is mistaken, as is the charge that I "blithely accept and repeat" it. Lustig assumes, with his usual lack of generosity, that my statement that Longuet-Higgins based his representation of harmonic space on "first principles" was based merely on my noticing his use of that phrase in the final sentence of the relevant paper. On the contrary, my statement was grounded also on my familiarity with Longuet-Higgins' writings on a wide range of psychological topics, and on various unpublished lectures and informal conversations over many years. His belief that psychological theory—whether concerned with music, language, vision, or anything else—should start with what David Marr [28] called a "computational" analysis of the abstract structure of the task is made explicit in several of his papers [16]. His description of his work on harmony is thus entirely consistent with the rest of his psychological work (which is scrupulously free of bluster).

One last example: Lustig complains that it is absurd to suggest that Bach knew of the musical rules extracted from his music by Longuet-Higgins, or that he thought about them when composing his fugues. But whoever claimed that he did? Not I, and certainly not Longuet-Higgins. What Longuet-Higgins was trying to do was to identify certain musical structures characteristic of (some of) Bach's music, and to define procedures whereby these structures could be perceived by a listener. Just as we are not (usually) consciously aware of following rules when we understand language, so we are not (usually) conscious of following musical rules. Many aspects of linguistic—and musical—generation and understanding lie beyond conscious access. One could, of course, raise questions about whether the procedures and symbolic representations defined by Longuet-Higgins are plausible models of what actually goes on in the (conscious or unconscious) mind. Indeed, Longuet-Higgins himself does so (as I point out on page 92 of *TCM*), in saying that parallel-processing is involved in human perception, though not in his program. But Lustig does not appear to be asking that question.

My disappointment at Lustig's review lies in the lost opportunity for constructive discussion and development of the musical examples I presented. I would have been very interested to know what a musicologist might say about various detailed aspects of Longuet-Higgins' work on fugue (discussed in Chapter 5), or Johnson-Laird's on jazz (Chapter 6). A refusal to address these issues, whether because of an unargued scepticism about computational approaches or because of the musical amateurism of the author reporting them, is unhelpful.

9. Conclusion

Ram and colleagues have offered an excellent summary of our problem. They say that the main interrelated pieces of the creativity puzzle are: inference and the control of inference, knowledge representation, and representational change. I agree. And Perkins has offered a pertinent reminder, and warning: the Gordian knot of creativity is a nest of knots within knots.

Some of these knots have already been loosened by AI-ideas, and may one day be untied by work in AI. Others must be tackled by philosophical analysis, historical research, or psychological experimentation. The set of four Lovelace questions is a conceptual knife designed to cut the Gordian knot of computer-creativity into four stands, to be unravelled in different ways. AI-researchers need to cooperate with colleagues in psychology, philosophy, history of art and science, and even musicology, if they are to find their way within this nest of knots.

References

[1] L.P. Beaudoin and A. Sloman, A study of motive processing and attention, in: A. Sloman, D. Hogg, G. Humphreys, D. Partridge and A. Ramsay, eds., *Prospects for Artificial Intelligence* (IOS Press, Amsterdam, 1993) 229–238.

[2] M.A. Boden, *Artificial Intelligence and Natural Man* (Basic Books, New York, 2nd ed. expanded, 1987).

[3] M.A. Boden, ed., *Dimensions of Creativity* (MIT Press, Cambridge, MA, in press).

[4] M.A. Boden, Agents and creativity, *Commun. ACM* **37** (7) (1994) 117–121; Special Issue on Agents (D. Riecken, ed.).

[5] M.A. Boden, *The Philosophy of Artificial Life* (Oxford University Press, Oxford, 1994).

[6] H.J. Eysenck, The measurement of creativity, in: M.A. Boden, ed., *Dimensions of Creativity* (MIT Press, Cambridge, MA, 1994) 199–242.

[7] H. Gardner, *The Creators of the Modern Era* (Basic Books, New York, 1993).

[8] H.E. Gruber, *Darwin on Man: A Psychological Study of Scientific Creativity* (Wildwood House, London, 1974).

[9] A. Harrison, *Making and Thinking: A Study of Intelligent Activities* (Harvester Press, Hassocks, Sussex, 1978).

[10] J.H. Heerwagen and G.H. Orians, Humans, habitats, and aesthetics, in: S.R. Kellert and E.O. Wilson, eds., *The Biophilia Hypothesis* (Shearwater Press, Washington, DC, 1993) 138–172.

[11] P.N. Johnson-Laird, *The Computer and the Mind: An Introduction to Cognitive Science* (Fontana, London, 1988).

[12] P.N. Johnson-Laird, Jazz improvisation: a theory at the computational level, in: P. Howell, R. West and I. Cross, eds., *Representing Musical Structure* (Academic Press, London, 1993) 291–326.
[13] A. Karmiloff-Smith, *Beyond Modularity: A Developmental Perspective on Cognitive Science* (MIT Press, Cambridge, MA, 1992).
[14] P. Langley, H.A. Simon, G.L. Bradshaw and J.M. Zytkow, *Scientific Discovery: Computational Explorations of the Creative Process* (MIT Press, Cambridge, MA, 1992).
[15] D.B. Lenat, The role of heuristics in learning by discovery: Three case studies, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell, eds., *Machine Learning: An Artificial Intelligence Approach* (Tioga, Palo Alto, CA, 1983).
[16] H.C. Longuet-Higgins, *Mental Processes: Studies in Cognitive Science* (MIT Press, Cambridge, MA, 1987).
[17] M.L. Minsky, *The Society of Mind* (Simon and Schuster, New York, 1985).
[18] R.D. Pea and D.M. Kurland, On the cognitive effects of learning computer programming, *New Ideas Psychol.* **2** (1984) 137–168.
[19] G. Polya, *How To Solve It: A New Aspect of Mathematical Method* (Princeton University Press, Princeton, NJ, 1945).
[20] T.S. Ray, An approach to the Synthesis of life, in: C.G. Langton, C. Taylor, J. Doyne Farmer and S. Rasmussen, eds., *Artificial Life II* (Addison-Wesley, Redwood City, CA, 1992) 371–408.
[21] C. Rosenberg and T. Sejnowski, Parallel networks that learn to pronounce English text, *Complex Syst.* **1** (1987) 145–168.
[22] S. Schaffer, Making up discovery, in M.A. Boden, eds., *Dimensions of Creativity* (MIT Press, Cambridge, MA, 1994) 13–52.
[23] K. Sims, Artificial evolution for computer graphics, *Computer Graph.* **25** (4) (1992) 319–328.
[24] A. Sloman, Motives, mechanisms, and emotions, in: M.A. Boden, ed., *The Philosophy of Artificial Intelligence* (Oxford University Press, Oxford, 1990) 231–247.
[25] G. Taylor, *Reinventing Shakespeare* (Hogarth Press, London, 1990).
[26] S. Todd and W. Latham, *Evolutionary Art and Computers* (Academic Press, London, 1992).
[27] L. Birnbaum and G. Collins, Opportunistic planning and Freudian slips, in: *Proceedings Sixth Annual Conference of the Cognitive Science Society*, Boulder, CO (1984) 124–127.
[28] D.C. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, San Francisco, CA, 1982).
[29] J.R. Searle, Minds, brains, and programs, *Behav. Brain Sci.* **3** (1980) 417–424.