CREATIVITY IN A NUTSHELL

Margaret A. Boden

Clarifying what creativity is the first step towards answering the question: could a computer be creative?

Creativity and computers: what could these possibly have to do with one another? 'Nothing!,' many people would say. Creativity is a marvel of the human mind. But computers, with all due apologies to Mario, Sonic, and friends, are basically just tin-cans. It follows — doesn't it? — that the two are related only by utter incompatibility.

Well, no. Computers and creativity make interesting partners with respect to two different projects. One, which interests me the most, is understanding human creativity. The other is trying to produce machine creativity — or anyway, machine 'creativity' — in which the computer at least appears to be creative, to some degree.

What is Creativity?

First things first. Human creativity is something of a mystery, not to say a paradox. One new idea may be creative, while another is merely new. What's the difference? And how is creativity possible? Creative ideas are unpredictable. Sometimes, they even seem to be impossible — and yet they happen. How can that be explained? Could a scientific psychology help us to understand how creativity is possible? Creativity is the ability to come up with ideas or artefacts that are new, surprising, and valuable. 'Ideas', here, includes concepts, poems, musical compositions, scientific theories, cooking recipes, choreography, jokes ... and so on, and on. 'Artefacts' include paintings, sculpture, steam-engines, vacuum cleaners, pottery, origami, penny-whistles ... and you can name many more.

As these highly diverse examples suggest, creativity enters into virtually every aspect of life. It's not a special 'faculty', but an aspect of human intelligence in general. In other words, it's grounded in everyday abilities such as conceptual thinking, perception, memory, and reflective self-criticism. So it

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

isn't confined to a tiny elite: every one of us is creative, to a degree.

Nor is it an all-or-none affair. Rather than asking 'Is that idea creative, Yes or No?' we should ask 'Just how creative is it, and in just which way(s)?' Asking that question will help us to appreciate the subtleties of the idea itself, and also to get a sense of just what sorts of psychological process could have brought it to mind in the first place.

Creative ideas, then, are new. But of course, there's new — and there's new. Ask a teacher, for instance. Children can come up with ideas that are new to them, even though they may have been in the textbooks for years. Someone who comes up with a bright idea is not necessarily less creative just because someone else had it before them. Indeed, if the person who had it first was Shakespeare, or Euclid, we'd think even more highly of the achievement.

Suppose a twelve-year old girl, who'd never read *Macbeth*, compared the healing power of sleep with someone knitting up a ravelled sleeve. Would you refuse to say she was creative, just because the Bard said it first? Perhaps, if you'd been talking around the topic with her, encouraging her to come up with non-literal ways of speaking, and even putting one or more of the three key ideas into the conversation. Otherwise, you'd have to acknowledge her remark as a truly imaginative one.

What you might do, and what I think you should do in this situation, is to make a distinction between 'psychological' creativity and 'historical' creativity. (P-creativity and H-creativity, for short.) P-creativity involves coming up with a surprising, valuable idea that's new to the person who comes up with it. It doesn't matter how many people have had that idea before. But if a new idea is H-creative, that means that (so far as we know) no-one else has had it before: it has arisen for the first time in human history.

Clearly, H-creativity is a special case of P-creativity. For historians of art, science, and technology — and for encyclopaedia users, too — H-creativity is what's important. And in daily life, we appreciate it too: it really isn't true that 'The old jokes are the best ones.' But for someone who is trying to

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

understand the psychology of creativity, it's P-creativity that's crucial. Never mind who thought of the idea first: how did that person manage to come up with it, given that they had never thought of it before?

If 'new', in this context, has two importantly different meanings, 'surprising' has three. An idea may be surprising because it's unfamiliar, or even unlikely — like a 100-to-1 outsider winning the Derby. This sort of surprise goes against statistics.

The second sort of surprise is more interesting. An unexpected idea may 'fit' into a style of thinking that you already had — but you're surprised because you hadn't realized that this particular idea was part of it. Maybe you're even intrigued to find that an idea of this general type fits into the familiar style.

And the third sort of surprise is more interesting still: this is the astonishment you feel on encountering an apparently impossible idea. It just couldn't have entered anyone's head, you feel — and yet it did. It may even engender other ideas which, yesterday, you'd have thought equally impossible. What on earth can be going on?

The Three Ways of Creativity

'What is going on' isn't magic — and it's different in each type of case. For creativity can happen in three main ways, which correspond to the three sorts of surprise.

The first involves making unfamiliar combinations of familiar ideas. Examples include poetic imagery, collage in painting or textile art, and analogies. These new combinations can be generated either deliberately or, often, unconsciously. Think of a physicist comparing an atom to the solar system, for instance, or a journalist comparing a politician with a decidedly non-cuddly animal. Or call to mind some examples of creative associations in poetry or visual art.

In all these cases, making — and also appreciating — the novel combination requires a rich store of knowledge in the person's mind, and many different ways of moving around within it.

The journalist or newspaper-reader needs a host of concepts

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

about both politics and animal behaviour, and some 'personal' knowledge about the individual politician in question. Cartoonists who depict Ken Livingstone (the first publicly-elected Mayor of London) as a newt are tapping into many different conceptual streams, including gossip about what he keeps in an aquarium in his home. The surprise you feel on looking at the cartoon is largely caused by seeing a human figure with a newt's crest and tail: a combination of ideas that's even less probable than the outsider winning the Derby.

If the novel combination is to be valued by us, it has to have some point. It may or (more usually) may not have been caused by some random process — like shaking marbles in a bag. But the ideas/marbles have to have some intelligible conceptual pathway between them for the combination to 'make sense'. The newt-human makes sense for many reasons, one of which is Ken's famed predilection for newts. (What are some of the others?) And (to return to the example from Macbeth) sleep is a healer, as knitting can be. Even if two ideas are put together randomly in the first place, which I suspect happens only rarely, they are retained/valued only if some such links can be found.

The other two types of creativity are interestingly different from the first. They involve the exploration, and in the most surprising cases the transformation, of conceptual spaces in people's minds.

Exploring Conceptual Spaces

Conceptual spaces are structured styles of thought. They're normally picked up from one's own culture or peer-group, but are occasionally borrowed from other cultures. In either case, they're already there: they aren't originated by one individual mind. They include ways of writing prose or poetry; styles of sculpture, painting, or music; theories in chemistry or biology; fashions of couture or choreography, nouvel cuisine and good old meat-and-two-veg ... in short, any disciplined way of thinking that's familiar to (and valued by) a certain social group.

Within a given conceptual space, many thoughts are possible, only some of which may have been actually thought.

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

Some spaces, of course, have a richer potential than others. Noughts-and-crosses is such a restricted style of game-playing that every possible move has already been made countless times. But that's not true of chess, where the number of possible moves, though finite, is astronomically large. And if some sub-areas of chemistry have been exhausted (every possible molecule of that type having been identified), the space of possible limericks, or sonnets, has not — and never will be.

Whatever the size of the space, someone who comes up with a new idea within that thinking-style is being creative in the second, exploratory, sense. If the new idea is surprising not just in itself but as an example of an unexpected general type, so much the better. And if it leads on to others (still within the same space) whose possibility was previously unsuspected, better still. Exploratory creativity is valuable because it can enable someone to see possibilities they hadn't glimpsed before. They may even start to ask just what limits, and just what potential, this style of thinking has.

We can compare this with driving into the country, with an Ordnance Survey map that you consult occasionally. You can keep to the motorways, and only look at the thick red lines on your map. But suppose, for some reason (a police-diversion, or a call of nature), you drive off onto a smaller road.

When you set out, you didn't even know it existed. But of course, if you unfold the map you'll see it marked there. And perhaps you ask yourself 'I wonder what's round that corner?' and drive round it to find out. Maybe you come to a pretty village, or a council estate; or perhaps you end up in a cul-de-sac, or back on the motorway you came off in the first place. All these things were always possible (and they're all represented on the map). But you'd never noticed them before — and you wouldn't have done so now, if you hadn't got into an exploratory frame of mind.

In exploratory creativity, the 'countryside' is a style of thinking. Instead of exploring a structured geographical space, you explore a structured conceptual space, mapped by a particular style of painting, perhaps, or a specific area of theoretical chemistry.

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

All professional artists and scientists do this sort of thing. Even the most mundane street-artists in Leicester Square produce new portraits, or new caricatures, every day. They are exploring their space, though not necessarily in an adventurous way. Occasionally, they may realize that their sketching-style enables them to do something (convey the set of the head, or the hint of a smile) better than they'd been doing before. They add a new trick to their repertoire, but in a real sense it's something that 'fits' their established style: the potential was always there.

Transforming the Space

What the street-artist may also do is realize the limitations of their style. Then, they have an opportunity which the Sunday driver does not. Give or take a few years, and ignoring earthquake and flood, the country roads are fixed. Certainly, you can't change them. Your Ordnance Survey map is reliable not only because it's right, but because it stays right. (Have you bothered to buy a new book of road-maps within the last few years?) But the maps inside our heads, and favoured by our communities, can change — and it's creative thinking which changes them.

Some changes are relatively small and also relatively superficial. (Ask yourself: what's the difference?) The limits of the mental map, or of some particular aspect of it, are slightly pushed, slightly altered, gently tweaked. Compare the situation in geographical space: suppose everyone in that pretty village suddenly added a roof-extension to their cottage. It may ruin the prettiness of the village, but it won't change the dimensions of the map. At most, the little 'portrait' of the village (assuming that it's that sort of map) will have to be redrawn.

The street-artist, then — or Picasso, in a similar position — has an opportunity. In principle, he (or, as always, she) could do the psychological equivalent of adding roof extensions, or building a new road (a new technique, leading to new possibilities), or even re-routing the motorway.

Re-routing the motorway (in 'real life' as in the mind) is the

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

most difficult of all. The surprises that would engender could be so great as to make the driver lose his bearings. He may wonder if he's been magically transported to a different county, or even a different country. Maybe he remembers a frustrating episode on his last trip, when he wanted to do something but his passenger scornfully said: 'In England, motorways are like this: they simply don't allow you to do that. You want to do it? Tough! It's impossible.'

A given style of thinking, no less than a road-system, can render certain thoughts impossible — which is to say, unthinkable. The difference, as remarked above, is that thinking-styles can be changed — sometimes, in the twinkling of an eye.

Someone skilfully writing a limerick won't find iambic pentameters dropping from their pen. But if you want to write a new sort of limerick, or a non-limerick somehow grounded in that familiar style, then maybe blank verse could play a role. The deepest cases of creativity involve someone's thinking something which, with respect to the conceptual spaces in their minds, they couldn't have thought before. The supposedly impossible idea can come about only if the creator changes the pre-existing style in some way. It must be tweaked, or even radically transformed, so that thoughts are now possible which previously (within the untransformed space) were literally inconceivable. But how can that possibly happen?

Machine-Maps of the Mind

To understand how exploratory or transformational creativity can happen, we must know what conceptual spaces are, and what sorts of mental processes could explore and modify them.

Styles of thinking are studied by literary critics, musicologists, and historians of art, fashion, and science. And they are appreciated by us all. But intuitive appreciation, and even lifelong scholarship, may not make their structure clear. (An architectural historian, for instance, said of Frank Lloyd Wright's Prairie Houses that their 'principle of unity' is 'occult'.)

This is the first point where computers are relevant. Con-

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

ceptual spaces, and ways of exploring and transforming them, can be described by concepts drawn from artificial intelligence (AI).

AI-concepts enable us to do psychology in a new way, by allowing us to construct (and test) hypotheses about the structures and processes that may be involved in thought. For instance, the structure of tonal harmony, or the 'grammar' of Prairie Houses, can be clearly expressed, and specific ways of exploring the space can be tried out. Methods for navigating, and changing, highly-structured spaces can be compared.

Of course, there is always the additional question of whether the suggested structures and processes are actually implemented in human heads. And that question isn't always easy to answer. But the point, here, is that a computational approach gives us a way of coming up with scientific hypotheses about the rich subtleties of the human mind.

Computer Creativity?

What of the second link between machines and creativity? Can computers be creative? Or rather, can they at least appear to be creative?

Many people would argue that no computer could possibly be genuinely creative, no matter what its performance was like. Even if it far surpassed the humdrum scientist or street-artist, it would not be counted as creative. It might produce theories as ground-breaking as Einstein's, or music as highly valued as McCartney's *Yesterday* or even Beethoven's Ninth ... but still, for these people, it wouldn't really be creative.

Several different arguments are commonly used in support of that conclusion. For instance: it's the programmer's creativity that's at work here, not the machine's. The machine isn't conscious, and has no desires, preferences, or values — so it can't appreciate or judge what it's doing. A work of art is an expression of human experience and/or a communication between human beings, so machines simply don't count.

Perhaps you accept at least one of those reasons for denying creativity to computers? Very well, I won't argue with you here (but see Chapter 11 of Boden 2004). Let's assume, for

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

the purpose of this discussion, that computers can't really be creative. The important point is that this doesn't mean that there's nothing more of interest to say.

All the objections just listed accept, for the sake of argument, that the imaginary computer's performance is indeed very like that of human beings, whether humdrum or not. What I want to focus on here is whether it's true that computers could, in fact, come up with ideas that at least appear to be creative.

Computer Combinations

Well, think of combinational creativity first. In one sense, this is easy to model on a computer. For nothing is simpler than picking out two ideas (two data-structures) and putting them alongside each other. This can even be done with some subtlety, using the (connectionist) methods described in Chapter 6.

In short: a computer could merrily produce novel combinations till Kingdom come.

But would they be of any interest? We saw, above, that combining ideas creatively isn't like shaking marbles in a bag. The marbles have to come together because there is some intelligible, though previously unnoticed, link between them which we value because it is interesting — illuminating, thought-provoking, humorous ... — in some way. (Think sleep and knitting, again.) We saw also that combinational creativity typically requires a very rich store of knowledge, of many different kinds, and the ability to form links of many different types. (Here, think politicians and newts again.)

And we don't only form links, we evaluate them. For instance, we can recognize that a joke is 'in bad taste'. In other words: yes, the links that the joker is suggesting are actually there (so it is a real joke). But there are other links there also, which connect the ideas with sorrow, humiliation, or tragedy. The joker should have noticed them, and should have refrained from reminding us of them.

For a computer to make a subtle combinational joke, never mind to assess its tastefulness, would require (1) a data-base with a richness comparable to ours, and (2) methods of link-

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

making (and link-evaluating) comparable in subtlety with ours. In principle, this isn't impossible. After all, the human mind/brain doesn't do it by magic. But don't hold your breath!

The best example of computer-based combinational creativity so far is a program called JAPE, which makes punning jokes of a general type that's familiar to every eight-year-old (see Chapter 12). But making a one-off jest is usually more demanding. Ask yourself, for instance, what Jane Austen had to know in order to write the opening sentence of *Pride and Prejudice*: 'It is a truth universally acknowledged that a single man in possession of a good fortune must be in want of a wife.' (And why, exactly, is it funny?)

Artificial Explorers and Self-Transforming Machines

What about exploratory creativity? Several programs already exist which can explore a given space in acceptable ways. One example is AARON, a drawing-program described in Chapter 7. AARON can generate thousands of line-drawings in a certain style, pleasing enough to be spontaneously remarked upon by unsuspecting visitors — and to be exhibited in galleries worldwide, including the Tate. (The most recent version of AARON is able to paint its drawings, too: see Chapter 12.)

Another is David Cope's 'Emmy', discussed in Chapter 12. This composes music in many different styles, reminiscent of specific human composers such as Bach, Vivaldi, Mozart ... and Stravinsky. Still others include architectural programs that design Palladian villas or Prairie Houses (also mentioned in Chapter 12), and programs that can analyse experimental data and find new ways of expressing scientific laws (Chapter 8).

A few AI-programs can even transform their conceptual space, by altering their own rules, so that interesting ideas result. Some of these ideas were already known to human beings, though not specifically prefigured within the program. (See the discussion of the automatic mathematician, AM, in Chapter 8.) But others are first-time-fresh. 'Evolutionary' programs, for instance, can make random changes in their current rules so that new forms of structure result. At each

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

generation, the 'best' structures are selected, and used to breed the next generation.

Two examples that evolve coloured images (some of which, like AARON's, are exhibited in galleries world-wide) are described in Chapter 12. In each case, the selection of the 'fittest' at each generation is done by a human being, who picks out the most aesthetically pleasing patterns. In short, these are interactive graphics-environments, in which human and computer can cooperate in generating otherwise unimaginable images. These computer-generated images often cause the third, deepest, form of surprise — almost as if a coin being tossed repeatedly were suddenly to show a wholly unexpected design. In such cases, one can't see the relation between the daughter-image and its parent. The one appears to be a radical transformation of the other, or even something entirely different.

Anyone who has watched TV regularly over the past few years, or who has visited museums of contemporary art, will already know that many novel graphic images have been produced by self-transforming AI-programs of this kind. The problem is not to make the transformations: that is relatively easy. What's difficult is to state our aesthetic values clearly enough to enable the program itself to make the evaluation at each generation. At present, the 'natural selection' is done by a human being (for example, the gallery-visitor).

In more well-regulated domains, however, the value-criteria can often be stated clearly enough to allow the evolutionary program to apply them automatically. An early example, a program for locating leaks in oil-pipelines, is mentioned in Chapter 8. Now, scientists are starting to use these techniques to enhance their own creativity. Biochemical laboratories in universities and pharmaceutical companies are using evolutionary programs to help design new molecules for use in basic research and/or medicine. Even the 'brains' and 'bodies' of robots can now be evolved, instead of being designed (see Chapter 12).

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

Values and Creativity

One huge problem here has no special relevance to computers, but bedevils discussion of human creativity too.

I said earlier that 'new' has two meanings, and that 'surprising' has three. I didn't say how many meanings 'valuable' has — and nobody could. Our aesthetic values are difficult to recognize, more difficult to put into words, and even more difficult to state really clearly. (For a computer model, of course, they have to be stated really, really clearly.)

Moreover, they change: who will proudly admit, today, to having worn a beehive hairdo or flared trousers in the 1960s? They vary across cultures. And even within a given 'culture,' they are often disputed: different sub-cultures or peer groups value different types of dress, jewellery, or music. And where transformational creativity is concerned, the shock of the new may be so great that even fellow-artists find it difficult to see value in the novel idea.

Even in science, values are often elusive and sometimes changeable. Just what 'simplicity' or 'elegance' mean, as applied to scientific theories, is something that philosophers of science have long tried — and failed — to pin down precisely. And whether a scientific finding or hypothesis is 'interesting' depends on the other theories current at the time, and on social questions too (might it have some medical value, for instance?).

Because creativity by definition involves not only novelty but value, and because values are highly variable, it follows that many arguments about creativity are rooted in disagreements about value. This applies to human activities no less than to computer performance. So even if we could identify and program our aesthetic values, so as to enable the computer to inform and monitor its own activities accordingly, there would still be disagreement about whether the computer even appeared to be creative.

The answer to our opening question, then, is that there are many intriguing relations between creativity and computers. Computers can come up with new ideas, and help people to

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

do so. Both their failures and their successes help us think more clearly about our own creative powers.

This article is reprinted with permission from pp. 1-10 of M. A. Boden, The Creative Mind: Myths and Mechanisms *(London: Routledge, 2004). All references are to this book.*

Margaret Boden is Research Professor of Cognitive Science, in the Centre for Cognitive Science, University of Sussex.

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press

https://doi.org/10.1017/S147717560000230X Published online by Cambridge University Press