CREATIVITY AND LAW:
HOW CAN THEY LIVE TOGETHER?

Margaret A. Boden
Research Professor of Cognitive Science
University of Sussex

Creativity is the ability to generate ideas that are new, surprising, and valuable.

'Ideas' include concepts, analogies, statutes, stories, theories, music... and artifacts such as tools and houses. 'New' can mean what's psychologically new to *a particular individual* (i.e. P-creative), or what--in addition--is new with respect to *a particular culture*, or even to *the whole of human history* (i.e. H-creative). 'Surprising' can mean different things--including, sometimes, the shock of a seemingly *impossible* idea. 'Valuable' covers many different features--which vary across times, places, and social groups.

In legal thinking, 'valuable' may mean any of the following (and, very likely, more): internally coherent; congruent with current statutes and/or common-law practices; precedented; unprecedented; procedurally feasible; financially affordable; politically acceptable (to government and/or public); in accord with natural justice; in accord with accepted moral standards; fair; simplifying (as in 'tidying,' not 'simplistic'); long-awaited/overdue; elegant; easily intelligible; imaginative; daring; and radical.

Clearly, no single example of legal thought can satisfy all these values maximally. So the legal version of 'horses for courses' will require judicious decisions as to which ones are most relevant in any particular case. This doesn't mean that *no* contradictions can be allowed. People can weigh competing values against each other, sometimes consciously juggling to keep them all, somehow, in the air. Similarly, AI-systems capable of 'multiple constraint-satisfaction' can tolerate logically untidy sets of criteria to some extent. But an AI-model without this ability cannot: for such a model, the competing values must simply be dropped.

Creativity can't come out of nothing: it requires an established knowledge-based, and disciplined self-critical thinking (not least, in assessing the value of the novel idea). Creative ideas can arise in three ways. The three types of creativity (defined below) are *combinational* creativity (CC), *exploratory* creativity (EC), and *transformational* creativity (TC). These are analytic distinctions: a given case of creative thinking may involve more than one type.

Where the law is concerned, we must distinguish between (1) the daily practice of law and (2) the suggestion of new judgments and/or statutes, possibly involving novel insights in jurisprudence. I'll call these *practical law* and *theoretical law*, respectively. The optimal relation between creativity and law will vary, depending on which type of law--and which type of creativity--is involved.

Combinational creativity

(CC) involves putting familiar ideas together in unfamiliar ways.

Examples include: collage in visual art; much poetic imagery; some metaphors and similes (and their visual equivalents, e.g. in cartoons); and analogies in art and science (e.g. the Rutherford-Bohr 'solar-system' model of the atom).

Legal examples of CC include the imaginative use of past cases as analogies of some current case (studied in AI under "case-based reasoning"). For this to count as "creativity," the analogy mustn't be obvious. But the legal values listed above require that the analogy not be *too* imaginative. Poets can get away with outrageous and/or highly tenuous analogies. Scientists are much more limited: their values include the possibility of empirical verification. Lawyers lie somewhere in between the two.

Practical lawyers are more tightly constrained than their theoretical colleagues. For they have to satisfy those values listed above which are relatively straightforward to assess. However, "relatively" is a slippery term, and one must admit that the practice of law can be imaginative indeed.

One memorable example (reported in my newspaper at the time) concerned a case in which a man was accused of stealing banknotes. The bank had provided a list of the numbers of all the missing notes, many of which were found in the possession of the accused. Nevertheless, the prisoner was acquitted. This travesty of justice occurred because the law then stated that any documents accepted in court must have been produced by someone "having knowledge of their contents." The judge argued that the list of numbers was inadmissible as evidence, because it had been produced by the bank's computer--which, like *all* computers, was incapable of knowing anything.

Full marks for ingenuity! (Whether we should assign full marks for the judge's philosophy of AI/mind is a controversial question, which I'll ignore. But I'd be interested to know whether this judgment was thrown out on appeal, and/or how the law was adjusted in order to prevent such a defense being presented again.)

What sort of AI-system could have come up with this ingenious defense? It would have to be one which considered not only the evidence, but the *source* of each piece of evidence. And it would have to be provided with general rules/criteria for assessing different sources. If the source is a witness present at the scene of crime, can psychological research on the fallibility of eye-witness evidence be brought to bear, either in general or in some more specific way? If it's a policeman's notebook, how likely is it that the police are 'fitting up' the accused? If it's an expert witness, what are their qualifications, and what is their past history of court-appearances? And if it's a trusted institution, such as a bank, how far can that trust be relied on?

In the case mentioned above, the key legal criterion was that the person in the institution who prepared the documentary evidence be capable of understanding it. That criterion could conceivably be included within an AI-program, and 'flagged' to a quick-witted human lawyer--who might then think up the 'computers can't know/understand' defence. But I don't see how this could have been done for the first time by the AI-system itself. It's not that this is impossible *in principle*, for every aspect of human thought is--in principle--expressible in computational terms. But it's hugely improbable *in practice*, because we lack the psychological knowledge, the computer-power, and the patience (not to mention the money) to build AI-systems that approximate human minds so closely as

this. However, once someone has presented this defense in court, it could be attached to the relevant criterion for consideration in future cases.

This raises a general point about the usefulness of AI-systems for creative thinking in law (or anything else). One might want them *to come up with the answer*. Of course, the human lawyer (or medic, or ...) must always take the final responsibility of announcing this to be the answer. But the hope, here, is that the AI-program will provide *candidate answers*. Alternatively, one might merely want the AI-system *to help human beings to come up with the answer*, for instance by reminding them of considerations and/or similarities which they might not have thought of. (Given an *interactive* system, the lawyer might be able to prompt, or guide, the program in various ways.)

My own view is that--because of the difficulties already remarked above--there's likely to be more mileage in using AI-models as *aids* to creative problem-solving than as *substitute* problem-solvers. This comment applies for all types of creativity, and for both 'practical' and 'theoretical' law. Indeed, it even applies to relatively 'uncreative' problem-solving in legal contexts, for a judge's *interpretation* of the law will depend on a rich background of world-knowledge and human sensitivities. (So in the battle between those who believe and those who deny that statute law can be fully captured by logic, I side with the deniers: see [Leith 1988; Whitby 1996]).

A large part of the practice of law consists in identifying similar cases, and in arguing the similarity. Arguing the similarity is a matter of rhetoric, and very little AI-research has addressed this. Roger Schank's group at Yale started thinking about it in the late-1970s [Birnbaum 1982], and some recent papers have addressed various social/rhetorical aspects of argumentation [Reed 1997; Reed, Norman, & Gabbay forthcoming]. Moreover, one AI story-writer makes a distinction between the points-of-view of the Characters and the point-of-view of the story-teller--a distinction which is used to determine not what the story (the basic plot) is, but *how the story is told* [Turner 1994]. Possibly, developments of some of these might help lawyers to *present* their supposedly similar case persuasively, after having found it. (Social psychological research on the biasses in the order of presentation [e.g. Hovland 1957] would be relevant, as would cognitive studies of everyday heuristics in decision-making and the assessment of 'statistical' evidence [Gigerenzer & Todd 1999]. But since juries are one thing and judges another, one would also need empirical research on the extent to which *judges* can learn to overcome such biases.)

What 'legal' AI has focused on is *finding* the similar case in the first place. In general, AI models of analogy typically allow one to adjust the desired closeness of the match (less for poets than for scientists). Similarly, models of case-based reasoning in law should allow such latitude. But even if the lawyers can decide just how close they want the analogy to be, and (more to the point) in just which ways, it may be very difficult to express this decision in computational terms. In particular, values (including the 'legal' values listed above) usually aren't easy to specify explicitly. This has nothing to do with the common claim that "Computers can't really have values!" Perhaps they can't. But values can certainly be represented, either implicitly or explicitly, in AI-systems.

A notorious illustration of this was the St. George's Hospital Medical School computer program, designed in the early-1980s to 'match' the medics' choice of applicants to be called for interview [n.a. 1988]. It eventually became clear that the gender-based and racial prejudices of the doctors had been included, represented as

numerical weightings added 'unthinkingly' in order to achieve a closer match to the overall pattern of choices made in previous years by the human assessors. These biases weren't obvious from the program itself, and hadn't been consciously inserted. They were discovered only by chance. (Someone looking over the shoulder of the clerk entering the data on the application forms into the computer asked why she was pressing a certain set of buttons; it turned out that the buttons were coding gender and race, and were linked to prejudicial weightings 'inside' the system.)

Of course, these relatively straightforward values *could* have been included explicitly, as criteria to be considered by the program. So, in principle, could all of the values listed above. But that's easier said than done. Defining "elegance," for example, is a challenge. Philosophers of science have been trying to define elegance, or "simplicity," for years, with scant success. And although AI-modellers of scientific discovery have managed to define one type of (mathematical) "symmetry" in computational terms, they can't define *symmetry in general* [Langley et al. 1987]. This is doubly relevant here, since John Rawls' [1971] influential theory of justice rests on a principle of symmetry--namely, "fairness." As for defining "consonant with natural justice," or "in accord with culturally accepted morals," in computational terms, don't hold your breath! What's meant by "natural justice" has been debated over centuries of work in jurisprudence.

The second and third types of creativity are closely linked, and significantly different from combinational creativity.

Exploratory creativity

(EC) involves exploring, navigating, and testing the potential and boundaries of some pre-existing way of thinking, or 'conceptual space.' Any reasonably complex space is likely to contain points (i.e. potential ideas) that haven't been visited before.

Examples include: story-schemas (such as Goal-Interference-Revenge, or Goal-Problem-Help-Gratitude-Reciprocation); limericks and sonnet-form; fugue and rock-and-roll; the various styles of dance; Impressionist painting and Palladian architecture; cricket, chess, and noughts-and-crosses; benzene-chemistry; Euclidean geometry ... and 'punk' clothes and jewelry.

Transformational creativity

(TC) involves changing one or more dimensions of the current conceptual space, with the result that *things can now be thought which were impossible to think before.* Some changes are mere 'tweaks'; others are more fundamental. (One might want to classify the 'tweaks' as EC, not TC.)

Examples include: skirts designed with many layers instead of just one; the move from tonal to atonal music; a new style of painting, such as impressionism or cubism; the move from 'string-molecules' to 'ring-molecules' in chemistry; non-Euclidean geometries, constructed by dropping or altering one of Euclid's axioms.

If CC is the daily bread of problem-solving in practical law, EC and TC are less common. Indeed, TC isn't really relevant here, because it's not up to practicing lawyers to change (transform) the law. One might say that judges change the law when they allow a new precedent. But they do this only when the new case 'fits' the old schema. In other words, they're exploring, perhaps even tweaking, the existing conceptual space--not transforming it.

Other examples of EC in daily practice include the search for loopholes in the tax-laws (or company law, or ...). Creative tax-lawyers explore the current system of allowances and restrictions to find previously unvisited, even unsuspected, places where their clients may legally perch. And, of course, creative tax-legislators seek to close them.

EC and, especially, TC really come into their own when we consider theoretical law. Here, more freedom of thought is possible. Current statutes and practices can be tested, tweaked, and at the limit transformed--in more or (probably) less fundamental ways--to give a vision of an alternative legal system. The newly-transformed system (or sub-system) would satisfy the *intelligibility* value, since it would be closely linked, to the previous one--for after all, it's the latter which provided the relevant conceptual space, one (or more) of whose dimensions has been altered. Whether it would also satisfy the other values on our list, perhaps even better than the prior system did, is another question. If it didn't, it would be less acceptable.

Changing the law by TC is very different from adopting a wholly different system. That could happen if a country decided--or was forced--to abandon its own legal system and take on the laws of some other culture. In that case, some of the *values* might have to be changed and/or reinterpreted. (For example, think of recent attempts to impose Islamic law on the Christian community in Nigeria; or consider the various defenses of Koranic values, as being worthy *in themselves* and/or worthy *as purer versions of 'Western' values*.)

AI in general can model both EC and TC. This may surprise some people, for it's commonly believed that a program can't transform itself. However, transformations of the given conceptual space are possible, using meta-level heuristics [Lenat 1983] or genetic algorithms (GAs).

The difficulty lies not in effecting the transformation, but in evaluating the result--or, what comes to much the same thing, in transforming it 'sensibly' in the first place. In short, we're face-to-face with values, again. In cut-and-dried domains, where the relevant values ('fitness criteria') can be clearly specified, evolutionary transformations can result in successful, or even optimal, solutions. But aesthetic values, for instance, are notoriously difficult to identify and/or spell out. That's why GA-programs in art are typically interactive, with some human being making the selection at each generation [Sims 1991; Todd & Latham 1992]. A GA-program in law would require such interactivity too, since legal values are only somewhat less problematic than aesthetic ones.

Exploring a given space by computer is less problematic. The very fact that the conceptual space is culturally accepted is evidence that it's *valued*. Hence any location within it, and (probably) any small-scale tweaking of it, will be acceptable too. However, to express a conceptual space clearly enough for a program to engage in EC on the base of it isn't the work of a moment. For automated EC in music, one needs a fine musician/musicologist as well as a competent programmer. *Pari passu*, for automated EC in law, one needs to define the relevant legal space clearly enough for it to be modeled in an AI system. In addition, one needs to provide procedures for moving through the space, for navigating within it, for locating its boundaries, and perhaps for tweaking it, too.

I said, above, that small-scale tweaking of an accepted space would "probably" generate acceptable results. But it's not always easy to say, in general terms, what counts

as "small-scale." I'm reminded, here, of an example given years ago by the computer scientist Lord Bowden, warning against over-reliance on programs for *arithmetic*, of all things. Consider a greengrocer, using his pocket calculator to determine the cost of hiring staff to pick seasonal fruit. He pays by the pound (avoirdupois), and knows that a 17-year-old girl can pick *x* pounds of blackberries in one hour, and a 23-year-old man, *y* pounds. Woe betide him, however, if--wanting a total of (3x + 3y) pounds--he sends them out blackberrying for three hours together. They'll either bring back much more than he can afford to pay for, because they've been showing off to each other, or very much less--for reasons I don't need to spell out. Blackbery-pickers of very different ages would have been a better bet.

Who'd have thought that the reliability of a greengrocer's arithmetic would depend on the ages of his fruit-pickers? Given that the law touches virtually every aspect of our lives, there must be many 'fruit-picker' examples in legal contexts, too. These include the unexpected losers (or, occasionally, beneficiaries) of "small-scale" changes in benefit-regulations, for example. Tweaking the current space, to include previously neglected deserving cases, may take protection away from others, "deserving" in different ways. And what of laws on inheritance, and the like? It's been suggested in Britain recently that the law should be changed to give long-term cohabiting homosexual couples the same inheritance-rights (and rights as "next of kin") as married couples. Some see this as a mere tweak, and long overdue at that. Others see it as a fundamental transformation, affecting--even undermining--the institution of marriage itself.

In these cases, at least, one has some specific laws/practices to act as signposts. With respect to the more abstract reaches of jurisprudence, I find it difficult to believe that AI-systems could help much here. (Maybe they could help with considerations of statute law, but not with the much less 'manageable' common law?) I don't think it plausible, for instance, to suggest that an AI-model could help someone aiming to design a 'Rawlsian' legal system. Again, this isn't a matter of *principle*. But jurisprudence, as an essentially philosophical enterprise, involves creative (EC/TC) thinking of a kind which will defeat AI-modellers for a very long time--perhaps even for ever.

References:

[1988] *Report of a Formal Investigation into St. George's Hospital Medical School*. London: Commission for Racial Equality.

Birnbaum, L. [1982] "Argument Molecules: A Functional Representation of argument structure," *Proceedings of the 2nd National Conference on AI (AAAI-82)*, Pittsburgh, 63-65.

Boden, M. A. *The Creative Mind: Myths and Mechanisms*. Routledge, 2003. (2nd edition, expanded.)

Gigerenzer, G., & P. M. Todd (eds.) [1999] *Simple Heuristics That Make Us Smart.* Oxford: Oxford University Press.

Hovland, C. J. (ed.) [1957] *The Order of Presentation in Persuasion*. New Haven, Conn.: Yale University Press.

Langley, P., H. A. Simon, G. L. Bradshaw, & J. M. Zytkow. [1987] *Scientific Discovery: Computational Explorations of the Creative* Process. Cambridge, Mass.: MIT Press. Pp. 243-306.

Leith, P. [1988] "The application of AI to Law," *AI & Society*, 2, 31-46.

Lenat, D. B. "The Role of Heuristics in Learning by Discovery: Three Case Studies". In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach*. .ne Palo Alto, Calif.: Tioga, 1983.

Rawls, J. [1971] *A Theory of Justice*. Cambridge, Mass.: Belknap Press.

Reed, C. [1997] "Representing and Applying Knowledge for Argumentation in a Social Context," *AI & Society*, 11, 138-154.

Reed, C., T. Norman, & D. Gabbay. [forthcoming] *Handbook of Argumentation and Computation.*

Sims, K. [1991] "Artificial Evolution for Computer Graphics, *Computer Graphics*, 25 (no. 4), 319-28.

Todd, S., & W. Latham. [1992]. *Evolutionary Art and Computers*. London: Academic Press.

Turner, S. R. [1994] *The Creative Process: A Computer Model of Storytelling and Creativity*. Hillsdale, NJ: Lawrence Erlbaum.

Whitby, B. [1996] *Reflections on Artificial Intelligence: The Legal, Moral, and Ethical Dimensions*. Exeter: Intellect Books.