In Praise of Evolution and Opacity

Notes for the Dagstuhl Workshop on Creativity

Inman Harvey
inmanh@gmail.com

1. The perspective I bring to questions of creativity in man and machine reflects my background in evolutionary robotics, which I consider an exercise in philosophy of mind, trying to understand the relationship between Behaviour and Mechanism. I have a suitcase full of philosophical premises and assumptions; some of these are common, some will be identifiable as associated with a Dynamical Systems approach to cognitive science, also influenced by Varela; and some will seem plain perverse.

2. This position paper it is in the form of numbered paragraphs, each making a single point. They may not make a continuous structured argument, but I hope nevertheless hang together.

3. The Darwinian revolution showed us that humans are just another species. We do have interesting differences -- above all language and culture -- but from an evolutionary perspective this is fairly insignificant.

4. There is a schism in cognitive science between those who define cognition in terms of what differentiates humans from other animals; and those who want to construct the concept of cognition as broadly as possible, to include how all living things relates to their worlds, whether they are humans, other animals, plants, microbes -- or robots and other artificial agents. I am firmly in the second camp.

5. Those in the first camp may consider cognitive properties, including reasoning, logic, and creativity, to exist in some Platonic world. They may well think in terms of some absolute, God-given measures of such properties. If they want to make a machine creative, they will see no need to relate this specifically to the behaviour of a named biological species -- e.g. *Homo sapiens*.

6. Beauty, emergence, and creativity are all in the eye of the beholder. This does not mean that any of these are arbitrary, random or irrational. It just means you need to be explicit in deciding which individual, or which group or culture, is going to count, for you, as the beholder.

7. It is not sufficient to just say that *Homo sapiens* is the relevant beholder. Do you mean 21st century Ashanti, or 8th century Western European, or what/whom?

8. Dan Dennett, in a talk at TED, compared human aesthetics with baboon aesthetics. In the context of sexual selection, there are sound reasons why the backside of a female baboon should look delightfully attractive to a male one. There are generic similarities with human aesthetics, but our tastes are somewhat different. Should a robot have human aesthetics, baboon aesthetics, or robot aesthetics?

9. People from the second camp, such as myself, will always want to relate machine creativity to human creativity (or bird or plant creativity). When using artificial evolution to design machines, you need a fitness function to evaluate each individual in each generation. This means there must be a human (or bird or plant) in the fitness evaluation loop.

10. When using artificial evolution, you may need or want to have millions of generations, evaluating every individual. With humans in the loop, this is not practical. But there is a trick,

Dagstuhl Seminar Proceedings 09291
Computational Creativity : An Interdisciplinary Approach
http://drops.dagstuhl.de/opus/volltexte/2009/2196

with an intermediate panel of automated critics. You can rapidly evaluate and evolve millions of individuals using a small population of automated critics. Then once a day, evaluate the taste of the automated critics against the judgement of one or more humans. The critics should slowly evolve to track human taste; and then they do the heavy duty work of assessing the fitness of the target machines, and rapidly evolving them

11. There is a speed limit, however. Roughly speaking (with lots of small print) it is one bit per generation absolute maximum. And with tiered evolution, it is the number of generations at the top-level that ultimately counts. So with e.g. 365 generations, one can only evolve a genepool with a relatively few bits of useful information.

12. In evolutionary computation, we talk of a trade-off between exploration and exploitation. Exploration is movement towards the new and unseen, away from the old and familiar. If you turn the exploration parameter up too high, this degenerates into random search. Exploitation means constraining one's search to areas that are familiar. If you turn the exploitation parameter up too high, the result is stagnation. Natural Darwinian evolution has found ways of tuning these parameters, so as to find a sweet spot of balance in this trade-off. One can do the same in artificial evolution.

13. There is a perceived paradox about machines and creativity, related to the controversy about free will and determinism. Surely a machine is either "just following rules ", or else (to a greater or lesser extent) "just functioning randomly". In neither case do we feel tempted to credit it with free will, and in neither case do awe want to give it credit for any creativity. If people trying to build rules into machines have not worked out their response to this apparent paradox, then they are certainly wasting their time.

14. Quotation attributed to von Neumann: "You insist that there is something that a machine can't do. If you will tell me precisely what it is that a machine cannot do, then I can always make a machine which will do just that."

15. There is a permanent tension, in e.g. government and education, between two poles: On the one hand an evaluation system for (e.g.) university applicants should be seen to be open, unbiased and objective; this leads to approved published criteria -- and leads further to attempts by savvy applicants to game-play the system, by writing down what they know officially meets the criteria. On the other hand, critics say this just rewards box-ticking and predictability, when they may well be seeking individuality and creativity; so they insist on relying on personal subjective judgment -- with dangers of conservatism and cronyism. The search for creative machines runs into the same kinds of problems.

16. One can analyse and describe a human being at various different levels of description. I focus on two in particular: the behavioural (how they act as a person), and the mechanism level (how the physics and chemistry is assembled together into a working system). As with the two perspectives on a Necker cube, each view is valid but you cannot have both simultaneously. One hopes that one's doctor is competent at handling both of these perspectives.

17. Intentions, needs, wants, learning and, aesthetics, creativity are all part of the language of the behavioural level of description. Circuits, activations, attractors, plasticity are all part of the language of the mechanical systems level of description. They should not be confused.

18. The big mistake of the GOFAI approach to cognitive science was to assume some degree of isomorphism between the mechanical structure of the brain and the behavioural structure of the person. This looks like phrenology.

19. There is an interesting range of circumstances in which a system can be deterministic at the mechanical level of description (I am deliberately ignoring noise and possibly random

quantum effects here) and yet nondeterministic (in the eye of the beholder) at the personal or behavioural level of description. Since we cannot hold both views simultaneously, this does not give rise to any contradiction. What is this range of circumstances?

20. The behavioural level of description should allow for high-level abstractions such as "he is angry", "she had an idea", which make sense, and are much more useful to the speaker than any possible mechanical level description of the same set of circumstances. Roughly, this probably implies a system that is complex at the mechanical level, yet capable of being described as a simpler system at the behavioural level.

21. One common aim of cognitive scientists and neuroscientists is the ***reduction*** of behavioural descriptions to physical or mechanical descriptions. There is an ambiguity here. We may have successful theories that use sub-personal concepts to give a simplistic, but useful, understanding of personal behaviour; these are pseudo-mechanisms. Folk psychology works, and can be improved. But there is no way that humans can understand fully how humans work right down at the molecular level.

22. The nematode worm *C. elegans* has 302 neurons. Humans struggle to partially understand how it works. It is clearly absurd to think that this nematode has sufficient brainpower to understand how a brain with 302 neurons might operate. It is equally absurd to expect human brainpower to appreciate how a brain with 100 billion neurons might operate. Sophisticated extensions to folk psychology are the best we can hope for.

23. If we want to build machines to which we can attribute free will, or creativity, but then at the mechanical level they must be complex and opaque. There is a basic contradiction between this need for opacity and the quest of many cognitive scientists for reduction.

24. How can you design machines when the workings may be, indeed need to be, opaque to the designer? Darwinian evolution did this job in the biological world; artificial evolution has the potential to do the same job for machines.

25. The speed limit for evolution is not well known, but needs to be acknowledged.