Maturana's Autopoietic Hermeneutics Versus Turing's Causal Methodology for Explaining Cognition

Stevan Harnad
Chaire de recherche du Canada
Institut des sciences cognitives
Université du Québec à Montréal
Montréal, Québec Canada H3C 3P8
http://www.crsc.uqam.ca/
and
Department of Electronics & Computer Science
University of Southampton
Highfield, Southampton, UK SO17 1BJ
http://www.ecs.soton.ac.uk/~harnad/

ABSTRACT: Kravchenko (2007) suggests replacing Turing's suggestion for explaining cognizers' cognitive capacity through autonomous robotic modelling by 'autopoiesis', Maturana's extremely vague metaphor for the relations and interactions among organisms, environments, and various subordinate and superordinate systems ('autopoietic systems') therein. I suggest that this would be an exercise in hermeneutics rather than causal explanation.

Keywords: cognition, computation, Turing test, distributed cognition, autonomy, autopoiesis, consciousness

Kravchenko (2007) has written a thoughtful and articulate critique of Harnad and Dror's (2006) argument (1) that cognition is the capacity of individual organisms to do what they can do, (2) that the way to explain cognition is to design a robot that can pass the Turing Test (i.e., able to do everything that organisms can do) and hence (3) that cognition is something that happens inside a cognizer (rather than being somehow 'distributed' across cognizers and the outside world).

First, Kravchenko's characterization of H&D's stance is largely accurate. The only point at which it is in error, however, is a crucial one: Not only did we not suggest that the only candidate process for what is going on inside the cognizer (whether organism or robot) is computation – i.e., symbol manipulation – but we are on record as being strongly critical of that view (Harnad 1990). The 'symbol grounding problem' arises because computation alone is not enough to generate or explain cognition. Compution consists of symbols and symbol manipulations that require an external interpreter, whereas cognition is autonomous and instrinsic to the cognizer. To make any symbols and symbol-manipulations going on inside a cognizer autonomous rather than parasitic on an external interpretation, the cognizer has to have full sensorimotor (robotic) capacity. That is why the Turing Test in question is the robotic one, and not just the linguistic (symbolic) one that Turing originally envisioned (Harnad 2006).

Sensorimotor input/output capacity includes the sensory surfaces of the cognizer, which are in turn causally connected to the distal objects, events, properties and traits that project onto its sensory surfaces; the motor output is likewise dynamic, taking the 'shape' of the distal objects that it manipulates. This is not symbol manipulation: Symbol shapes are arbitrary, and manipulated according to the syntactic algorithms of classical computation. In contrast, sensorimotor systems are dynamical physical systems.

But the sensorimotor systems still end with the proximal projections of the distal objects on their sensorimotor surfaces. There is a causal connection with the outside world, to be sure (natural cognizers are,

after all, Darwinian survival machines), but the cognizer itself ends at the sensorimotor surface. As philosopher's have put it, 'cognition is skin and in'.

Kravchenko, and some of the eloquent and influential (but not very clear or coherent) thinkers on whom he draws (notably, Maturana, and his remarkably nebulous notion of 'autopoiesis'; see Varela and Maturana 1980) prefer a smeared picture, in which cognizers are merely some sort of component – one is tempted to ask whether they are even autonomous components! – of an exceedingly vague 'relational' system, in which 'cognition' is seen as a distributed 'interaction' rather than as being whatever process is going on inside an autonomous cognizer that gives the cognizer the sensorimotor input/output capacities that it has (including, in the case of human cognizers, linguistic capacity).

The trouble with this extreme 'breadth' (shall we call it) in the scope of cognition is that it makes it very difficult to identify exactly what cognitive science is meant to do, and to explain. The Turing Test, which unapologetically takes the cognizer to be autonomous, takes its empirical task to be to explain the cognizer's sensorimotor capacity for interaction with its world. This at least has the merit of specifying an empirical problem and what would count as its solution.

What is the empirical problem that 'autopoiesis' is meant to address, and what would count as a solution? Perhaps the preferred object of study for autopoieticists is not the individual cognizer, but superordinate ecosystems in which cognition is going on in the organism/environment interactions. What then is cognition? All the kinds of relations and interactions taking place in an ecosystem? What is the problem, then, and what would count as a solution, for those who are interested in explaining cognition autopoietically?

My guess is that there is in fact no coherent empirical problem, hence no empirical solution, underlying the thinking of autopoieticists, because autopoieticists are not really doing empirical science at all, but are merely doing bionomic hermeneutics on the mind-body problem. The core intuition of autopoiesis comes from a particular stab at a solution to the (insoluble) 'mind/body' problem:

The mind/body problem (Harnad 2003) is a problem we all have with relating mental states (conscious states, which are also cognitive states) to bodily states (including brain states). We want to think they are the same thing, but we can't quite manage to see them that way, nor to explain how or why they are the same thing. Bodily states (including brain states) are clearly functional states – biologically adaptive functional states that make it possible for organisms to survive and reproduce, and that can be causally explained in the usual way, in terms of physical material and physical (including biological) processes. But mental states seem to be something different. It *feels like* something to be in a mental state. Physical states are just functional states: structure and function. It is not at all clear how (or why) functional states embody feeling (though some of them undoubtedly must). To put it another way, the mind/body problem is the problem of explaining how and why some functional states are felt, rather than merely being 'functed', like all other physical states.

Moreover, amongst our feelings are: what it feels like to see and hear, what it feels like to move (involuntarily as well as voluntarily), what it feels like to know and to understand and to mean; and what it feels like to know (or to feel you know) what you yourself – as well as other organisms – feel, know, understand and mean. All of this is what the old-school philosophers used to call 'being aware' and 'being aware of being aware' – i.e., consciousness. That is what it is to have a mind.

So my guess is that the way that autopoieticists have managed to soothe whatever unease they may have with 'solving' the mind/body problem by merely equating consciousness (hence cognition) with brain function is, instead, to smear it into a vague interactive relation in or with (depending on the still-incoherent question of the autonomy of cognizers) the world, involving subordinate and superordinate autopoietic systems.

Well, others don't find that autopoietic ointment all that soothing or helpful, and prefer something less vague and more amenable to the usual form of empirical explanation, which is to accept that cognizers are autonomous systems, and to explain their autonomous causal capacity through Turing robotic modelling.

A few other points could be made, in closing. Kravchenko is right to point out that 'in the brain' may be too restrictive a locus for cognition (if we accept that cognition is the mechanism underlying organisms' input/output capacity). Invertebrates and unicellular organisms are organisms too, and to the extent that they 'cognize', cognition is part of their *bodily* function, not their (nonexistent) cerebral function.

Kravchenko is also right to point out a limitation of (current) Turing robotics. Although he is wrong in claiming that robotics is limited to computation in its resources (sensorimotor systems include many other

potential transducer and effector materials and dynamics, plus analog processing as well as parallel and distributed neural networks, none of them consisting of computation – though all of course simulable computationally), he is right that at the present time it is not clear whether other biological materials and dynamics (e.g., biochemical ones) maynot prove to be necessary in order to generate and explain cognizers' capacities. But adding bioengineering functions to the Turing Test definitely does not make it into any less of an autonomous 'skin and in' matter, and into more of a distributed, relational one. The robot starts and stops with its transducer/effector interfaces with the world, regardless of what is going on inside its body.

It goes without saying that the behavioral capacity that the Turing Test has to generate includes social and linguistic interactions. So Kravchenko has not pointed out any omissions there. The rest of his scepticism about the Turing Test is just the flip-side of the mind/body problem, the 'other-minds' problem. The only one I can be sure is conscious is myself. With other people, other organisms, and indeed other robots, there is no way to be sure and no way to test 'directly'. There is, in fact, only the Turing Test. Kravchenko's relationism and autopoiesis, seen as some sort of smear across the entire ecosphere, may be personally soothing hermeneutically, but empirically it tells us next to nothing about cognition, nor about how to go about finding out more.

References

Harnad, S. 1990. "The symbol grounding problem". *Physica* D 42: 335-346.
Harnad, S. 2003. "Can a machine be conscious? " How? *Journal of Consciousness Studies* 10(4-5): 69-75.
Harnad, S. 2005a. "Distributed processes, distributed cognizers, and collaborative cognition". *Pragmatics & Cognition* 13(3): 501–514.
Harnad, S. 2005b. "To cognize is to categorize: Cognition is categorization". In H. Cohen and C. Lefebvre (eds), *Handbook of Categorization in Cognitive Science*. Amsterdam: Elsevier, 20-42.
Harnad, S. and Dror, I. 2006. "Distributed cognition: cognizing, autonomy and the Turing Test" *Pragmatics & Cognition* 14(2): 209-213.
Harnad, S. 2006. "The annotation game: On Turing (1950) on Computing, Machinery and Intelligence". In R. Epstein and G. Peters (eds), *The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Dordrecht : Kluwer, PAGES.
Kravchenko, A.V. 2007. "Whence the autonomy? A response to Harnad and Dror". Pragmatics & Cognition 15(3): PAGES.
Maturana, H. and Varela, F. 1980. *Autopoiesis and Cognition: The Realization of the Living*. Dordecht: Reidel.