Interview with Prof. Stevan Harnad

Silvia Coradeschi

Born in Budapest, Hungary, Stevan Harnad is professor at the Department of Psychology at Université du Québec á Montréal, holding the Canada Research Chair in Cognitive Sciences and is also Affiliate Professor in Electronics and Computer Science at University of Southampton, UK. He did his undergraduate work at McGill University and his graduate work at Princeton University. His research is on categorisation, communication and cognition and the open research web. KI interviewed him about Symbol Grounding on August 19, 2012.

KI: What inspired you to propose symbol grounding as a novel concept?

From 1978-2003 I was editor of the journal Behavioral and Brain Sciences (BBS), the journal in which computationalism - the thesis that cognition is computation - was beginning to become more important and influential. I too was impressed by the progress that artificial intelligence was making in generating and explaining our behavioral capacities computationally, especially compared to the empty decades of behaviorism, and in the face of neuroscience's failure to explain how our brains manage to generate our behavior. With what came to be called the "Turing Test" (TT), Alan Turing seemed to have defined the agenda for cognitive science - design a

Silvia Coradeschi
School of Science and Technology
Örebro University, AASS
E-Mail: silvia.coradeschi@oru.se

model that can do whatever we human cognizers can do, and you have explained how it can be done - and computation looked like the way to design the model.

But then in 1980 the philosopher John Searle published what was to become his influential critique of computationalism, the "Chinese Room Argument", a thought-experiment in which he pointed out that if there were a purely computational model that could pass the TT - i.e., communicate by email with humans, indistinguishably from a human, for a lifetime - then, if the computer program could pass the TT in Chinese, Searle himself, who does not understand Chinese, could memorize and execute all the computations too, while still not understanding Chinese. Hence computation alone cannot generate understanding. So cognition cannot be just computation.

What is missing? As Alan Turing showed in his epochal work, computation is formal symbol-manipulation - reading and writing symbols based on rules that operate only on the arbitrary shapes of the symbols (whether 0's and 1's or the words of a natural language) not on their meanings. So where does the meaning of symbols come from? My version of this impasse was the Chinese-Chinese dictionary-go-round: How can one learn the meanings of Chinese words from a Chinese-Chinese dictionary alone? Everything is defined there, but the definitions are meaningless unless you already know what some of them, at least, mean. How to "ground" the meanings of those symbols, so the rest of them can be defined and described through words alone? That is what I dubbed the "symbol grounding problem".

The natural candidate for grounding symbols is direct sensorimotor interaction with the things the symbols refer to, so that the grounded symbol system can recognize manipulate and name them,

autonomously. These are not just verbal email capacities: they are sensorimotor robotic capacities. And they are not just computational. Sensorimotor transduction is not computation; it is physical dynamics. And a good deal more dynamics are likely to be needed to pass the grounded robotic version of the TT, rather than just the symbolic email version.

To categorize is to do the right thing with the right kind of thing. The focus of my research has since been on how we acquire categories - first the ones that are grounded through direct sensorimotor induction, and then the ones we can learn via verbal instruction, through definition, description and explanation. Symbol grounding is also the necessary condition for the evolution of language, but language has to be grounded in the capacity to acquire some categories, at least, directly, through the senses, rather than indirectly, through language. (How many categories are needed to ground language - and which ones - is another question. So is the question of how learning to categorize things might change the way they look to us; this is called "categorical perception" and is a special case of the "Whorf Hypothesis" that experience and language can influence the way we perceive the world).

KI: How do you see that the field has progressed since your first formulation of symbol grounding?

It is now widely accepted that although computation is an essential tool in studying and modeling cognition, and it is also a potentially powerful and important component of cognition, cognition is not and cannot be just computation. Computation is hardware-independent symbol-manipulation; but there are many potentially important components of cognition that are not hardware-independent computation but dynamical, physical processes, such as sensorimotor transduction, analog structures and processes, even biochemical and molecular ones.

The widening of the Turing Test from the original purely symbolic version to the robotic Turing Test has opened the door to "situated" robotics, "embodied" and "embedded" cognition, and also parallel/distributed and other forms of analog processing. It has also deepened the study of the interface between nonverbal and verbal cognitive function.

A still-open question is how relevant and helpful neuroscientific data will be in explaining cognition. There is no question that the brain does it all. And also that much of brain activity is not computational (though like almost everything, it can be approximated and simulated computationally). But it is not clear which features of brain function are relevant to cognition, and how they cause cognition. It is still the robotic version of the TT that must be the arbiter not only of what is relevant, but of what is successful in causing cognition - which means: in generating the capacity to do all the things that human cognizers can do.

KI: How do you see the symbol grounding problem in the context of the recent trends in the semantic web and internet of things?

The semantic web is not really a semantic web! It is just a syntactic web, just as all computation is (meaningless symbols, manipulated on the basis of their shapes). What makes people want to call it (erroneously) "semantics" is the fact that human cognizers have tagged some of the symbols in the web with further symbols of their own, whose meanings we all understand. But that is not grounded meaning (semantics). It remains just as parasitic on the (grounded) meanings in the brains of its users (human cognizers) as Searle's Chinese pen-pal Turing Test does.

Another new web-inspired notion - alongside embodied cognition and situated cognition - is "distributed cognition": the notion that cognitive states may not just occur within the heads of cognizers, but might somehow be distributed across many heads, as well as across their external tools, including the web.

I think this is nonsense. What it forgets and leaves out is also what Turing left out: Cognitive states are not just action-capacity states, generating all of our doings; they are also felt states. It *feels like something* to cognize. Cognition would be trivially distributed if all there was in the world was doings. Then it would be arbitrary, in the Turing Test, what we designated as being the robot, and what we designated as being the rest of the world. But the reason it is not that arbitrary is because cognizers feel. And the notion that there is a further feeling entity, wider than the head of an individual human cognizer (or TT-passing robot), that likewise "cognizes", is as unlikely as the notion that there is a further feeling entity, wider than the head of an individual human cognizer, that likewise feels a headache - or that the web that told me that my grant was not funded is part of the headache that it induced in me.

The problem of explaining how and why cognizers can do what they can do (Turing's problem) has come to be called the "easy" problem of cognitive science (even though it is far from easy, and we are from solving it), in order to contrast it with what has been called the "hard" problem of cognitive science, which is to explain how and why cognizers feel. About the hard problem all I can say is that the Turing Test certainly cannot

test it, and symbol grounding certainly does not solve it.

The year 2012 is Alan Turing's year. In honor of the centenary of Turing's birth, over fifty speakers at an international, interdisciplinary Summer Institute on the Evolution and Function of Consciousness tried to address the hard problem that defies Turing testing and symbol grounding.

KI: Communication is an important aspect in symbol learning through sharing. Clark proposes that a common ground of shared beliefs and intentions can be achieved between two communication partners through a specific communication protocol, involving attention, signal decoding and understanding and, finally, the creation of joint projects. In this sense common ground is achieved when both interaction partners agree that they have reached mutual understanding. However, communication - at least in the real world - is noisy, as is the representation (or grounding) of concepts or symbols. How (much) does communication change the meaning of symbols? I.e. how do noisy transmission channels (which may possibly lead to misunderstandings) on the one hand and slightly different concepts on the other hand affect the meaning as it is constituted within an interaction and how does this affect the representation of a symbol?

I regret that this question is too vague to answer, except to say that, yes, language allows us to communicate, sharing not only symbols (words) but meanings (as long as enough of the words are grounded in the sensorimotor experience of sender and receiver). And, yes, noise and error can distort both symbol transmission and meaning. And of course if the grounding of the symbols is different for the sender and the receiver, the intended meaning may not be transmitted: If my meaning of "apple" was grounded only on round, red apples, and you tell me that you ate a fruit the color of grass, I will assume, wrongly, that it was not an apple.

But all categories are approximate (just as a picture is always worth not just more than 1000 words but more than an infinite number of words) and all categories (other than purely formal ones) are also provisional. In the apple example I just gave, I would have to learn - either from direct sensorimotor induction or from verbal (symbolic) instruction - that there are green apples too. That would resolve the communicative confusion.

Most of cognition is like that: an ongoing process of resolving the confusion among confusable alternatives, guided either by the direct error-correcting consequences of categorizing and miscategorizing (doing the right or wrong thing with the right or wrong

kind of thing) or - if you are an organism or robot that has language - guided by (grounded) verbal instruction.

KI: One research trend in robotics focusses on emotions as motivational foundation for action, perception and learning. How would you say do emotions affect symbol grounding processes?

Interpreting robotic processes or activities as emotional is as groundless as interpreting ungrounded symbols as meaningful: The problem of explaining emotions is just a particular case of the "hard" problem of explaining how and why we feel, rather than just do.

Motivation is much easier, because it can be treated as purely behavioral and functional - even a thermostat can have "goals" - as long as we leave out the (hard) fact that it feels like something to have a goal!

KI: What do you see as the challenges for the future regarding SG?

Scaling up today's toy robots to be able to pass the full-scale TT. Almost all the hard work still remains to be done, and no really revolutionary ideas are yet on the horizon. (We can keep on plodding with vastly over-interpreted toy robots that we call "grounded", but grounding only becomes meaningful and significant at full human scale, and that means the full power of human language - plus all that nonverbal capacity that grounds language.)

KI: Which of your current and past projects would you consider the most successful or the biggest milestone?

I don't think there have been any substantive milestones in cognitive science yet, other than Turing computation itself, and perhaps also Chomsky's Universal Grammar.

KI: What are the scientific issues that you are currently working on today?

My research group measures and models how humans learn new categories from direct experience and from words. We also analyze dictionaries for the properties of the smallest number of words from which all the rest can be defined by words alone. And I never miss an opportunity to point out how and why new (and old) attempts to solve the "hard" problem keep failing.

KI: We know you are a promoter of open access, what are your main motivations in this respect?

I waste an awful lot of time trying to persuade researchers to make their refereed research freely accessible on the web (and to persuade researchers' institutions and funders to mandate [require] that they make their refereed research freely accessible on the web), in order to maximize the uptake, usage, applications, impact and progress of research, to the benefit of researchers, their institutions, the R&D

industries, and the tax-paying public that fund the research and for whose benefit it is being conducted.

The web is not distributed cognition, but it has become our "Cognitive Commons" - the place where all the categories collectively acquired by generations of scholarly and scientific research are stored, used and shared.

I say time is being wasted because what I am preaching is trivial and ought to be self-evident. It's not rocket science but "raincoat science" as in: "Look kids, it's raining outside. Time to put on raincoats." Countless groundless worries have given rise to a syndrome that I've provisionally dubbed "Zeno's Paralysis" (but that historians will eventually have to explain): It is taking an unbelievably long time for most researchers to get round to doing the few keystrokes per paper it takes to make them all open access.

KI: Finally, which question should I have asked you and what would have been the answer to it?

You should have asked me about the relation between the symbol grounding problem and the "hard" problem of consciousness, but I have already managed to slip in both the question and the answer above!

KI: Stevan, thank you very much for your time and for this interview.