The Case for Radical Epigenetic Robotics

Alexander I. Kovács
Grad. Univ. Advanced Studies
Nat. Inst. of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
101-8430 Tokyo, Japan
aik@grad.nii.ac.jp

Haruki Ueno
Grad. Univ. Advanced Studies
Nat. Inst. of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
101-8430 Tokyo, Japan
ueno@nii.ac.jp

Abstract

We propose a way to study cognitive robotics whereby one completely departs from a task-centered methodology. The new approach, "radical epigenetic robotics", tries to chart new territory, likely not explored by task-centered approaches because there one starts with a certain cognitive task for a robot and then tries to come up with suitable "control" mechanisms in order to achieve this specific task. Arguably, this approach has a high risk of making many task-specific assumptions and of leaving unexplored possible mechanisms. To avoid these problems, radical epigenetic robotics starts not with an outside cognitive task, but from the inside, with the "control" mechanism, which we call "cognitive substrate"; the design of this "cognitive network" is the focus of research, resultant cognitive competencies of a robot are tested, ex post facto.

Keywords Epigenetic robotics, cognitive substrate, designer's dilemma

1 Introduction

The aim of this paper is to propose and justify a way of studying the fabrication of cognizers (in the form of robots) that jettisons one of the central assumptions in the field of cognitive robotics, viz., that one has to start from the cognitive competency that one wants to model. Only then, the assumptions goes, can one try to understand what mechanism to use to make progress. We turn things around: we start with the mechanism and then study what cognitive competencies result. We argue here, that this approach—if properly construed—is not at all foolish or blind (that is lacking a tangible goal).

Our approach, tentatively called "radical epigenetic robotics" (to be explained below), completely departs from standard task-centered methodologies. In a task-centered approach one starts with a certain cognitive task for a robot—say, wall-following, map-learning, gesture recognition, playing soccer, six-legged walking, learning of movement by imitation, object perception and learning, etcetera—and then tries to write developmental programs [1], evolve controllers [2], design behaviors [3], formulate knowledge to reason about[4] or somehow come up with suitable sensor-effector-mappings in order to achieve this specific task. While there is nothing wrong with this approach, one could argue that it (a) has a high risk of making many task-specific assumptions and (b) leaves unexplored possible "control" mechanisms.

Radical epigenetic robotics does not try to replace or compete with such task-centered approaches. It is complementary by trying to chart new territory; territory likely not explored by common approaches for reasons (a) and (b). In the next section we elaborate on why (a) and (b) are so common and why they can be problematic. In section 3 we take a quick look at the different flavors of robotics out there to position our approach and loosely define some important terms (such as "radical" and "epigenetic"). In section 4 we outline our proposal and in section 5 we elaborate on cognitive substrates. We conclude the paper with a discussion of some difficulties of our approach.

2 Problems with Task-Centered Approaches

[W]e must admit the likelihood that top-down reverse engineering will simply fail to encounter the right designs in its search of design space [5, p. 258]

Task-centered approaches—i.e. approaches that start with a certain cognitive task to be modelled—to repeat, can be said to (a) have a high risk of making many task-specific assumptions and (b) leave unexplored possible "control" mechanisms. Why is this so and what are the problems?

Regarding (a), there are of course good reason for simplifying assumptions. As Dennett [6, p. 308] puts it, "AI research, like all other varieties of research on this huge topic [understanding how the brain works - AIK], must make drastic oversimplifications in order to make even apparent progress."

However, such assumptions are often implicit and since they are made on a case-by-case basis this leads to many possibly incompatible models. An often voiced critique of artificial intelligence is that it works under the unjustified assumption that all the models in its different subfields will eventually be put together to result in truly intelligent systems that can, say, pass Harnad's Total Turing Test [7]. The critique is based on precisely the fact that many researchers work under different simplifying and implicit assumptions that can lead to utterly incompatible models.

As far as (b) is concerned, one could argue that by having a certain task in mind, one is likely to choose ways of achieving this task by utilizing techniques that will obviously or most likely work. In such a methodology, there are good reasons not to try out many (possibly much simpler) mechanisms for, at first sight, *they may just not look appropriate for the task at hand*; and so, as the epigraph states, one may "simply fail to encounter the right designs."

3 Oh No, Not Yet Another Robotics!

There are now a large number of different approaches to cognitive robotics. One finds (the original) cognitive [4], behavior-based [3], developmental [1], cognitive developmental [8], epigenetic [9], evolutionary [2], and adaptive neuro-robotics [10]. As it stands, the boundaries between them are not clear-cut, but what all these approaches, in one way or another, are addressing is how to get more out of the robots than was (and could have been) designed for; that is to have "control" mechanisms that do more than *just* control (whence the quotes).

The original cognitive robotics [4] was concerned with sense-plan-act style cognition where the planning component was based on logical programming languages, such as Golog. Cognition here was construed as *reasoning about X*, where X could be anything from goals, to perceptions, to actions, and so forth. Logicist cognitive robotics ignores learning and development setting it apart from the other approaches mentioned, which acknowledge that cognitive robots have to be "beneficiary of a longish period of infancy" [11, p. 157] in which certain cognitive structures develop and organize, rather than being fixed by the designer.

Our proposal shares this emphasis on epigenetic development, i.e., the development determined primarily by interaction rather than genes (that is a prior design), in the sense of Piaget [9]. Whatever the name, all approaches just mentioned are task-centered. The point of departure, and novelty, of our approach is that we believe that only by shifting the focus from specific tasks to appropriate, plausible "control" mechanisms (see below) can the the problems associated with task-centered approaches be avoided. This shift of focus to an inside-out approach is what we mean by "radical".

Dennett's quote above continues: "There are many strategies of simplification, of which [...] five, while ubiquitous in all areas of mind / brain research, are particularly popular in AI. [...] Many of the best-known achievements of AI have availed themselves of all five [...] strategies of simplifications [...]. Some critics hostile to any efforts in cognitive science enabled by these strategies, but *there is no point in attempting to 'refute' them a priori* [emphasis added - AIK]. Since they are strategies, not doctrines or laws or principles, their tribunal is 'handsome is as handsome does.'" So, rather than further argue against such simplifications we leave it at having justified our approach via (a) and (b) and take it (again) with Dennett [6, p. 309]: "one might just adopt some rival strategy or strategies, and let posterity decide which are the most fruitful."

4 Radical Epigenetic Robotics

To avoid problems (a) and (b), radical epigenetic robotics proposes to *not* start with a task-specification (outside), but focus on the inside mechanisms first. The five strategies of simplification mentioned above by Dennett are, in short, as follows:

- ignore both learning and development;
- ignore how isolated subcomponent under study might be attached to the larger system;
- hope that scaling-up from toy problem to real domain will be a straightforward extrapolation;
- bridge various gaps in one's model with unrealistic stopgaps; and
- avoid the complexities of real-time, real-world coordination.

Our research program tries to find new mechanisms that precisely don't make any of these assumptions, thats what we mean by *appropriate*. Since we aspires to avoid these simplifications, clearly, *something else has to give*. That something is that we will initially have to make do with the fabrication of rather simple proto-cognition. However, by definition, an appropriate mechanism will, e.g., scale well, make additions of new components easy, not introduce makeshifts whose fixing turns out problematic.

To be *plausible*, the mechanisms we investigate try to avoid as many known problematic assumptions as possible, such as e.g. regarding encodingist representation [12]. So, to be plausible, they must, e.g., be radical constructivist [13], interactivist [12], and there are other criteria. They must also offer a solution to what we call the Designer's Dilemma. Since we are talking about *fabricating* cognizers, qua artifact, there must be something that the engineer *designs*. We now know that neither behaviors nor knowledge are the right level we can design for cognition (this assertion is irrespective of whether we use machine learning or not) because the designer just cannot anticipate (and hence not accommodate) an infinite number of cases. The Designer's Dilemma forces us to focus on a level where we *can* (in principle) have complete knowledge, that is, not the cognitive task but the level of the cognitive substrate, as we call appropriate and plausible mechanisms for the fabrication of veritable cognizers.

5 Cognitive Substrates

We shall call an appropriate and plausible "control" mechanism *cognitive substrate*. By cognitive substrate we mean, roughly: the layer underlying or bringing about cognitive capabilities; so we could call the animal brain a "neurobiological cognitive substrate". The volume in design space of "control" mechanisms that we are specifically targeting is located where one finds what Varela et al. [14] have called "cognitive networks".

It would now be rather foolish to blindly try to find such mechanisms such that if implemented on a target robot architecture would lead to "interesting" behavior or what Beer calls minimally cognitive behavior [15]. We therefore urge to base the development of such cognitive substrates on a solid theoretical foundation, on a "general theory of cognitive systems". Apart from the animal brain, biologist have also for a while been speaking of the immune system as a "biological cognitive network" [16] with—though very much simpler [17]—similar capabilities to the brain (memory, perception, etc.). And if one looks at the literature on so-called complex adaptive systems one cannot help but notice the cognitively laden vocabulary used to talk about these systems. Though these parallels have been noticed and also lead to interesting applications, more general theoretical work is outstanding. We can, nevertheless, already guess at what our cognitive substrate might look like: a self-organizing, synergetic network of loosely-coupled anticipating self-sustained oscillatory elements.

The reader might now object, *How is this different from research into artificial neural networks (ANN); isn't that just such an approach where you start with a mechanism (the inside) and then see what you can do with it (the outside)?* The quip reply to this objection is: maybe, but nobody ever put ANN's in robots just to study "what happens". Besides, ANN's wouldn't be under such heavy fire from workers pointing out what ANN's cannot do to realize *whole* cognizers if they were even a suitable tool.

Whether ANN's (whatever kind you look at) have serious short-comings that rule them out as cognitive substrate (and they do) is besides the point. However, *how* to "hook up" the sensors and effectors of a robot with an ANN (whatever kind) to produce some form of proto-cognition is precisely the point of contention. Of course, you could just as well take ANN's and embark on the research program we propose. This hasn't been done yet and it is an open problem, whether any proto-cognition would be producible. That is, of course, if we start with a network which *hasn't been already trained for a certain task*, for in radical epigenetic robotics, there are no a priori tasks to train an ANN for.

6 Conclusion

> *Is this optimistic prospect an illusion? Is this bottom-up project as hopeless as trying to build a tower to the moon? You can't get there from here, say [...] the skeptics. Don't even try.* [18, p.196]

The radical epigenetic robotics approach proposed could perhaps be express by stating: "Let the robot think what it wants and do what it wants". As thus, the approach has a number of obvious disadvantages compared to other approaches and will likely face difficulties (and because of that, perhaps outright rejection). For example, since we start with a (theoretically motivated) cognitive substrate, and hopefully some predictions made by the theory as to the expected behavior of the system, the question remains how to evaluate resultant behavior of a robot. What counts as proto-cognition and how can it be verified to be veritable, i.e., attributable to the cognitive robot itself, and not some human designer. However, this problem of evaluation is an highly interesting problem by itself. Perhaps what is needed are new sciences of cognitive robot psychology and ethology.

A second problem (though not of the approach as such) is that the "general theory of cognitive systems" on which to base our proposed cognitive substrates doesn't yet exist. However, there are hints in the literature, such as a criterion differentiating cognitive from non-cognitive systems [16, 17]—we might want to call "Hershberg-Efroni-criterion"—which states, that "in a cognitive system the capabilities of the system are not preordained merely by the plan of the system but need interaction with their environment to define them exactly."

We are in an early phase of developing this new approach, but think it is timely to start debate.

Acknowledgements

AIK was supported by the Joint Studies Program of The Graduate University for Advanced Studies and an NII Fellowship Grant provided by Sumitomo Electric Industries, Ltd.

References

[1] J. Weng. Developmental robotics: Theory and experiments. *Int. J. of Humanoid Robotics*, 1(2):199–236, 2004.

[2] S. Nolfi and D. Floreano. *Evolutionary Robotics.* MIT Press, Cambridge, MA, 2000.

[3] R.A. Brooks. Intelligence without reason. A.I. Memo 1293, MIT, AI Lab, April 1991.

[4] R. Reiter. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems.* MIT Press, Cambridge, MA, 2001.

[5] D.C. Dennett. Cognitive science as reverse engineering: Several meanings of "top-down" and "bottom-up". In Dennett [18], chapter 16, pages 249–259.

[6] D.C. Dennett. Cognitive ethology: Hunting for bargains or a wild goose chase. In Dennett [18], chapter 21, pages 307–322.

[7] S. Harnad. Minds, machines and Searle. *Journal of Theoretical and Experimental Artificial Intelligence*, 1:5–25, 1989.

[8] M. Asada, K.F. MacDorman, H. Ishiguro, et al. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous System*, 37(1):185–193, 2001.

[9] C. Balkenius, J. Zlatev, C. Brezeal, et al., editors. *Proc. 1st Int. Workshop on Epigenetic Robotics.* Lund, Sweden, 2001.

[10] T. Ziemke. The construction of 'reality' in the robot: Constructivist perspectives on situated artificial intelligence and adaptive robotics. *Foundations of Science*, 6(1):163–233, 2001.

[11] D.C. Dennett. The practical requirements for making a concious robot. In Dennett [18], chapter 9, pages 153–170.

[12] M. H. Bickard and L. Terveen. *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution.* Elsevier Science, New York, NY, 1995.

[13] E. von Glaserfeld. *Radical Constructivism: A Way of Knowing and Learning.* Falmer Press, London, 1995.

[14] F. Varela, A. Coutinho, B. Dupire, and N.Vaz. Cognitive networks: Immune, neural, and otherwise. In A. Perelson, editor, *Theoretical Immunology, Part II*, SFI Series on the Science of Complexity, pages 359–375. Addison-Wesley, Redwood City, NJ, 1988.

[15] R. A. Beer. Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes, M. J. Mataric, J.-A. Meyer, et al., editors, *From Animals to Animats 4: Proc. Fourth Int. Conf. on Simulation of Adaptive Behavior*, pages 73–76, Cambridge, MA, 1996. MIT Press.

[16] U. Hershberg and S. Efoni. The immune system and other cognitive systems. *Complexity*, 6(5):14–21, 2001.

[17] U. Hershberg. Proposing a new focus for the study of natural and artificial cognitive systems. In C.G. Prince, Y. Demiris, Y. Marom, et al., editors, *Proc. 2nd Int. Workshop on Epigenetic Robotics*, Edinburgh, Scotland, 2002.

[18] D. C. Dennett, editor. *Brainchildren: Essays on Designing Minds.* Bradford Books, 1998.