A pre-neural goal for Artificial Intelligence

Micha Hersch

Abstract From its onset, the discipline of Artificial Intelligence aimed at understanding intelligence through a synthetic approach. Over time, progress has been made by considering lower and lower levels of intelligence. I argue that this trend should be completed by its next step by considering pre-neural forms of intelligence as models for AI. To justify the relevance of such primitive cognition to intelligence, I recall the works of Piaget, Jonas and Maturana and Varela. By considering how these authors relate to the question of teleology, I illustrate the kind of insights a pre-neural AI could provide, which pertain to fundamental aspects of natural cognition.

1 Introduction - the neural consensus

The numerous debates surrounding the discipline of Artificial Intelligence (AI) have failed to provide any commonly accepted definition of intelligence, be it natural or artificial. Yet, regarding natural intelligence, there seems to be an unspoken necessary condition accepted by the overwhelming majority of the AI community. This condition is that natural intelligence is implemented in neural circuits. This is the case for the proponents of symbolic AI which use human reasoning as their model, for connectionists, which are explicitly interested in neural networks, for researchers in low-level artificial intelligence who always consider neural sensori-motor coordination, and even research in swarm intelligence has taken neurally endowed insects as its main source of inspiration. So there seems to be an underlying assumption that intelligence emerged with the appearance of the neuron, whose capability for fast signal transduction and adaptive connectivity allowed information processing and eventually full-fledged intelligence. While this assumption is certainly true to a certain extent, its corollary is to exclude any non-neural phenomenon as model for

Micha Hersch
University of Lausanne e-mail: micha.hersch@unil.ch

artificial intelligence. In this paper, I argue for a pre-neural artificial intelligence, i. e., an artificial intelligence research program that takes pre-neural intelligence as a model. I believe that such a program is likely to provide valuable insights into the nature of intelligence.

In the following, I try to substantiate this claim by first providing a short retrospective on AI (section 2) and elaborate on a fundamental difference between artificial and natural intelligence, which pertains to the notion of the subject (Section 3). I will then briefly mention three theories that draw a continuum between life and cognition, claiming that cognition cannot be understood outside of its anchor, the living system (section 4), and thus justifying a pre-neural approach to AI. Doing so will provide an illustration of the kind of questions a pre-neural AI may try to contribute to, which will be discussed in section 6.

2 The evolution of AI

The relatively recent discipline of Artificial Intelligence (AI) emerged as an offspring of the older discipline of logic. As it appeared, it took up the task modern logic had initially set to itself, which was the study of human thinking. Indeed, for the founders of modern logic such as Boole, the aim of logic was to "to investigate the fundamental laws of these operations of the mind by which reasoning is performed, to give expression to them in the symbolical language of calculus" [3, p.3]. De Morgan expressed a similar view in the first sentence of his *Syllabus*, which states that "logic analyses the *forms*, or *laws of action*, of thought" [4, p.9] and Frege's *Begriffschrift* is an attempt to find the "formal language of pure thought" [9]. Beyond the formalism, logicians were interested in human thinking abilities, and more precisely in rational thinking, which was considered the "pure" thinking.

Likewise, the General Problem Solver, one of the first artificial intelligence systems is considered by its author to "simulate human thought"[21]. And indeed, the kinds of problems this approach set out to solve, were certainly human problems like proving theorems and playing chess. Although the basic assumptions underlying this "Good Old-Fashioned AI" [11] were questioned by philosophers such as Dreyfus [5] and Searle [29], its initial successes promoted the wide acceptance of this symbolic, logical approach to artificial intelligence within the engineering community.

However, as some of its overly optimistic promises failed to be fulfilled, in the eighties the connectionist approach [27] met a renewed interest with the work of Hopfield [14] and others. This approach, in which artificial neural networks occupied the center stage, was clearly inspired by the brain physiology. It emphasized the perceptual aspect of intelligence as well as the learning abilities, focusing on problems such as pattern recognition. As such this approach enlarged its scope to encompass not only human thinking but also mammalian thinking, for example by considering Pavlovian reflexes in rabbits [26] or the navigation abilities of rats [1].

One decade later, it was argued that intelligence had to be understood at the level of behavior. Coming from the robotics community, a claim was made that the goal of AI was not to "simulate" intelligence, but to actually implement it in a real environment [30], and more precisely in a robotic device. This led to the revival of the, by then, somewhat forgotten cybernetics tradition, which emphasized sensori-motor couplings as a way to produce intelligent behavior. It considered problems like obstacle avoidance and light following. Combined with influences from Varela's enactive theory of cognition [33], this led to the appearance of embodied cognition as a new framework for the study of artificial intelligence. According to this theory, intelligence cannot exist in a vacuum, but must be grounded in an environment through a body. Cognition emerged to enable to adequately guide the actions of the body in a given environment and can only be understood in this context. As biological models displaying this kind of sensor-motor coordination, animals such as turtles [13] were used.

The evolution described above, although slightly caricatural, is indicative of a general trend. The model of intelligence used by AI researchers has evolved from human intelligence, through mammalian intelligence to vertebrate intelligence. The interest of AI research has shifted from high to low level intelligence and has thus followed an evolution backward with respect to the evolution of natural intelligence. The main drive for this evolution is the observed gap between natural and artificial intelligence.

3 The ontological gap

At the onset of artificial intelligence, the existence of a fundamental gap between natural and artificial intelligence was not clear to most of the AI community, despite strong arguments put forth by philosophers [5]. However, over the years this has become more and more widely recognized. A few observations on the brief history of AI hint at this gap. One such observation is that what is most easily performed by artificial intelligence is most difficult to do for natural intelligence and vice-versa. Indeed, it turned out to be easier to beat Kasparov at chess than to beat a four year old kid at bedtime story understanding. This very strongly suggests that the modes of operation of artificial and natural intelligence greatly differ from one another, which is related to very different modes of being.
A related observation, also pointed out in [8], is that the explanatory power of traditional AI is very limited. Indeed recent successes such as a Jeopardy! player, do not provide any insight on how a human can play such a game. In fact it was not the intention of its developer to do so [7]. Thus, part of the AI community has departed from its initial goal of "understanding intelligence" [23]. Those who did not,

adopted the more recent approaches to AI such as embodied cognition. Not surprisingly, the explanatory power of artificial intelligence has increased with the evolution of AI to lower levels of intelligence. For example, it could be shown how simple optical flow computations could steer a flying device the same way a fly controls its flight [36], or how a subjective representation of the body can be acquired through sensori-motor contingencies [12], or how the salamander can control its amphibian locomotion [15].

However, if situating intelligence in a body in constant dynamical interaction with its environment has provided interesting insights into the intermingling of intelligence, the body and the environment, it has only filled a fraction of the gap between natural and artificial intelligence which still remains abyssal. As mentioned in [10], current artificial systems still lack any sense of *meaning* and of *agency*. These notions remain foreign to artificial intelligence and unexplained in natural systems.

In the rest of this paper and for illustration purposes, I will focus on one element of agency, namely the concept of teleology. This concept, which had vanished from our post-aristotelian scientific tradition was reintroduced by by the proponents of cybernetics such as Wiener [28]. In doing so, they stripped off its causal nature and explained it by a causal mechanism, the negative feedback loop. To clearly emphasize the non-causal aspect of this new teleology, it was then dubbed teleonomy. This concept, echoing Waddington's canalization processes in biological systems [35], has been extended into the study of attractor dynamical systems which have been widely used for understanding of animal behavior [18] and for controlling the behavior of artificial systems.

4 Cognition as a continuation of life

In order to understand the origin of the thinking subject in general and teleology in particular, it is worth considering simpler forms of intelligence, or minimal cognition [32]. Indeed some prominent thinkers have argued for a continuity between biological processes and intelligence, a view adopted in the Alife community [30]. According to this view, intelligence and in particular neural intelligence is an outgrowth of life and should thus be considered in this light. In the following, I will briefly mention the position of four influential figures, Piaget, Jonas and Maturana (and his student Varela), who, while all emphasizing this continuity, reach different conclusions on the notion of telelogy for the development of intelligence.

4.1 Piaget and the promise of cybernetical teleolonomy

Jean Piaget was a trained biologist turned psychologist and epistemologist. He is probably mostly known within the AI community to researchers focusing on developmental robotics for his work on sensori-motor loops and imitation in newborns and children [24], as this work has inspired many in the field [22, 2].
In a later book “Biology and knowledge”[25], Piaget studied “the relations between organic regulations and cognitive processes”. For him, “life is essentially self-regulation” (p.48) through processes such as assimilation and accomodation. And cognitive processes are “a result of organic self-regulation of which they reflect the essential mechanisms” (p.49). Cognition must then be understood in the broader framework of self-regulation. And here Piaget recognized the relevance of cybernetics in the theoretical understanding of self-regulation, and even counted the use of “mathematical and cybernetical models” as one of the four methods for his investigations (p.93). Indeed he states that “all concepts of cybernetics are of immediate signification for the cognitive domain” [1](p. 95). In this context he seems to fully adopts Wiener’s teleonomical explanation of behavior. For him, natural systems, like cognitive processes, are purposeful, and this purposefulness can be explained in term of regulatory mechanisms such as the feedback loop.

4.2 Jonas and the fallacy of cybernetical teleolonomy

Hans Jonas, a student of Heidegger, attempted to lay the foundation for a philosophical biology in his book “The phenomenon of life” [17]. There he also argues for a continuity between biological process and intelligence, and more generally between life and mind. According to him, life is a precursor of the mind and as such contains in essence the necessary ingredients of human intelligence. And the hallmark of life (or its simplest form) is metabolism. Intelligence as we know it in natural systems is an outgrowth of metabolism and has thus inherited its mode of operation. The continuum between simple cell metabolism and the human mind can be described along four axes.

1. The first axis is the notion of teleology. For Jonas, organisms are by essence teleological. Their behaviors are guided by a purpose, a goal, which originates in themselves. We could call this the teleological closure. The most basic purpose, which is present in the simple cell, is the preservation of its structure, as an organized self distinct from the environment. The behavior of the cell is usually organized around this goal.

2. The second axis is the notion of identity. Organisms develop a sense of identity, as a whole distinct from the environment, that need to be preserved through a

[1] Our translations

teleological behavior. In its most sophisticated form, the sense of identity develops into the human conscience.

3. The third axis is the notion of desire (or instinct, emotions). The desire comes from the difference between the goal and the present situation of the organism. As such it helps maintaing the goal and eventually reaching it. It is the drive to the goal.
4. The fourth axis is the notion of freedom. The most basic freedom experienced by the organisms is the freedom of form (or structure) with respect to matter. It is the ability to survive and transcend, the matter which constitutes it. Organisms tend to increase their freedom (for example through mobility), as it will provide them more opportunities to reach their goals.

According to Jonas, the set of explanatory categories needed to account for life and mind differs from those that were developped for Descartes' *res extensa*. As such, attempts such as those of Piaget, to explain life and mind as if their nature was the same as that of inanimate objects is bound to fail, as they contradict our own experience as living subjects.
Consistently with his theory, Jonas specifically criticizes the cybernetical teleology as a fallacy, in a dedicated essay of his book [16]. For Jonas, the feedback loop or any other regulatory mechanism, is a means to achieve a purpose but it will never originate the purpose itself. According to him, Wiener's teleonomical machines merely accomplish the purpose of their users, not their own. Cybernetical teleolonomy blurs the basic difference between the existence of a purpose and its realization. As such, while teleology needs to be explained, the cybernetical explanation is far from satisfactory.

4.3 Maturana and Varela and the irrelevance of teleolonomy

Francisco Varela and his mentor and colleague Maturana are probably the primary source of inspiration for the embodied cognition approach to artificial intelligence. They formulated the concept of autopoiesis and described its relationship to cognition in their book "Autopoiesis and cognition" [19]. According to their definition, an autopoietic system can be understood as a system that continously generates its own components and maintains itself as a unity in the space in which its components exist. Autopoeitic systems are autonomous, as they are their own producers and maintain their own organization and thus their own identity. The cell is the paradigmatic autopoeitic system and other examples include the immune system [34] or the human being. In this framework, cognition is defined as the phenomenological domain generated by autopoiesis, in other words the experienced reality resulting from autopoiesis. Now, since autopoietic systems generate their own domains and their own reality, any relevant description of such system has to use concepts that pertain to its phenomenological domain or to a universal logic that is valid for all phenomenological domains. Otherwise, the description only conveys knowledge about

the observer, as it is expressed in terms belonging to the world of the observer, which can be unrelated to the world of the observed. In particular, as clearly stated in the chapter entitled "Dispensability of teleonomy" [20] the use of teleonomy to explain living systems is irrelevant. The notion of purpose is within the observer and does apparently not belong to the universal logic of phenomenological domains, and in general, autopoietic systems are taken to be purposeless.
Thus, according to Maturana and Varela, the cybernetical explanation of purpose addresses a wrong problem, and its description in terms of inputs and outputs is misleading as autopoietic systems have neither inputs nor outputs, they are operationally closed.

5 Pre-neural artificial intelligence

We see that, while emphasizing the continuity between life and cognition, the three theories described above have very different position on the notion of purpose. For Piaget it is a result of regulatory mechanisms, for Jonas it is fundamental to any explanation of life and mind but remains to be reconciled with mechanistic causality, and for Maturana and Varela it belongs to the observer and is not an intrinsic feature of autopoietic systems, meaning that purposeless cognition is in their sense perfectly possible.

Artificial intelligence, with its synthetic approach, can attempt to probe these hypotheses. And thanks to the hypothesized continuity between life and cognition, it can do so by using pre-neural models of cognition. One candidate for such a model is plant intelligence [31]. There are a number of reasons for using plants as a model of intelligent system. First, plants display remarkable behaviors, the sophistication of which is often underestimated. Plants optimize their access to natural ressources such as light and nutrients, they can anticipate seasonal changes and adapt to very different environmental conditions. They can clearly be seen as displaying teleological behavior like phototropism or pathogen fighting. Moreover, the absence of a central nervous system makes the signal processing indistinguishable from the behavior. This results in a different view of intelligence and sensori-motor coordination, whithout a clear distinction between the sensory and the motor domains. They also interact with other plants and insects by sending and perceiving chemical signals so that their ecology can be seen as a primitive social context. Another interesting feature of plants is the hormonal regulation of their behavior, an aspect that is often neglected in AI models of neural intelligence. Thus plant cognition, unlike lower-level cellular cognition, is sufficiently complex to go beyond intracellular signalling cascade and transcriptional regulation, while being more amenable to investigation than animal cognition. Moreover, being sessile, plant bodies are radically different from animal bodies, which results in a different kind of cognition. This kind of cognition is often neglected from the discussions on the nature of intelligence, although it is likely to broaden our views on the topic. Indeed, by considering plant cognition, we are less likely to project our own cognitive categories, which will ease the difficult task of objectification of intelligence, a pre-requisite to any artificial intelligence.

The question whether plants are intrinsically purposeful has no easy answer. And this is revealed by the paradoxical behavior of many plant biologists, who formally design and describe their research on the assumption of mechanistic plant behavior, but informally ascribe intentional agency to their plants. Investigating this question will force us to better define and understand teleology in biological organisms, and in particular whether teleology can be assessed from a third person point of view.
Maybe plant behavior can be understood and modeled without a notion of teleology, which would show that very sophisticated and plastic behaviors, that appear to be oriented towards a goal can be implemented without it. This would be encouraging for AI, as it would push the limits of what can be expected from a purposeless artificial agent, in terms of both robustness and diversity of behavior.
But perhaps plants do have an intrinsic notion of teleology. The goal of AI would then be to investigate where it comes from and what it is made of. It could be that the sense of purpose can only develop as a result of an evolutionary history. Intelligence would thus not only require a body to be expressed, but also its grounding into an evolutionary process to acquire its “needful freedom”[17] required for agency.

For now, the study of pre-neural cognition has been mostly restricted to bacterial sensory motor-systems such as chemotaxis [6], or in the perspective of self-organization and synchronization. While plant cognition is very different from bacterial and neural cognition, its study is very relevant for the understanding of intelligent and coordinated behavior. The use of such a model will likely bring new perspectives on cognition, which may well prove fruitful.

6 Conclusion

The original endeavour of Artificial Intelligence, inherited from logic, was to understand and create intelligence. Due to the difficulty of this challenge, progress could only be made at the cost of lowering the bar for intelligence and considering low-level cognition such as sensori-motor coordination. If AI wants to remain true to this endeavour, it should continue in this direction and consider pre-neural intelligence, such as the one displayed by plants. This evolution is in line with a number of theories arguing for a continuity between life and cognition, such as those developped by Piaget, Jonas and Varela. As we have seen, fundamental questions regarding the nature of intelligence, such as the status of telelology in cognition remain relevant and are probably more amenable to investigation in lower forms of intelligence. By considering simpler organisms, it will be possible to better understand their mode of being and operation and thus their cognitive aspects. This is a path the field of artificial intelligence should resolutely engage in, lest it become one among many engineering fields, oriented to a given set applications but indifferent to the principles of natural intelligence.

Acknowledgements I thank Sven Bergmann for his support and comments on the manuscript, as well as Marion Haemmerli, Basilio Noris and Christophe Calame for helpful discussions. I also thank anonymous reviewers for their constructive feedback.

References

1. A. Arleo and W. Gerstner. Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. *Biological Cybernetics*, 83(3):287–299, 2000.
2. A. Billard. Play, Dreams and Imitation in Robota. In *Workshop on Interactive Robotics and Entertainment*, 2000.
3. G. Boole. *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*. Walton and Maberly, 1854.
4. A. De Morgan. *Syllabus of a proposed system of logic*. Walton and Maberly, 1860.
5. H. Dreyfus. *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Pres, 1992.
6. M.D. Egbert, X.E. Barandiaran, and E.A. Di Paolo. A minimal model of metabolism-based chemotaxis. *PLoS Computational Biology*, 6(12), 2010.
7. D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A.A. Kalyanpur, A. Lally, J.W. Murdock, E. Nyberg, J. Prager, et al. Building watson: An overview of the deepqa project. *AI Magazine*, 31(3):59–79, 2010.
8. S. Franchi and G. Güzeldere. Machinations of the mind: Cybernetics and artificial intelligence from automata to cyborgs. In S. Franchi and G. Güzeldere, editors, *Mechanical bodies, computational minds: artificial intelligence from automata to cyborgs*. MIT Press, 2005.
9. G. Frege. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. 1879.
10. T. Froese and T. Ziemke. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3-4):466–500, 2009.
11. J. Haugeland. *Artificial intelligence: The very idea*. The MIT Press, 1989.
12. M. Hersch, E. Sauser, and A. Billard. Online learning of the body schema. *International Journal of Humanoid Robotics*, 5(2):161–181, 2008.
13. O. Holland. Exploration and high adventure: the legacy of Grey Walter. *Royal Society of London Philosophical Transactions Series A*, 361:2085–2121, oct 2003.
14. J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554, 1982.
15. A. Ijspeert, A. Crespi, D. Ryczko, and J.-M. Cabelguen. From swimming to walking with a salamander robot driven by a spinal cord model. *Science*, 315(5817):1416–1420, 2007.
16. H. Jonas. Fith essay: Cybernetics and purpose: A critique. In *The phenomenon of life*, pages 108–126. Harper and Row, 1966.
17. H. Jonas. *The phenomenon of life*. Harper and Row, 1966.
18. J.A.S. Kelso. *Dynamic Patterns: The Self-Organization of Brain and Behavior*. MIT Press, 1995.
19. H.R. Maturana and F.J. Varela. *Autopoiesis and cognition: The realization of the living*, volume 42. Springer, 1980.
20. H.R. Maturana and F.J. Varela. *Autopoiesis and cognition: The realization of the living*, volume 42, chapter 2. Springer, 1980.
21. A. Newell and H.A. Simon. GPS, a program that simulates human thought. *Computers and thought*, pages 279–293, 1963.

22. M. Pardowitz and R. Dillmann. Towards life-long learning in household robots: The piagetian approach. In *Proc. 6th IEEE International Conference on Development and Learning, London, UK*, 2007.
23. R. Pfeifer and C. Scheier. *Understanding Intelligence*. The MIT Press, 2000.
24. J. Piaget. *La formation du symbole chez l'enfant*. Delachaux et Niestlé, 1947.
25. J. Piaget. *Biologie et connaissance: essai sur les relations entre les régulations organiques et les processus cognitifs*. Gallimard, 1967. Translated into English as: *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes.*, 1971, U. of Chicago Press.
26. R.A. Rescorla and A.R. Wagner. *A theory of Pavlovian conditionning: variations in the effectiveness of reinforcement and nonreinforcement*, pages 64–99. Appleton-Century-Crofts, 1972.
27. F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
28. A. Rosenblueth, N. Wiener, and J. Bigelow. Behavior, purpose and teleology. *Philosophy of science*, 10(1):18–24, 1943.
29. J.R. Searle. Minds, brains and programs. *Behavioral and Brain Sciences*, 3:417–457, 1980.
30. L. Steels. The artificial life roots of artificial intelligence. *Artificial Life*, 1(1_2):75–110, 1993.
31. A. Trewavas. Aspects of plant intelligence. *Annals of Botany*, 92(1):1–20, 2003.
32. M. Van Duijn, F. Keijzer, and D. Franken. Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior*, 14(2):157–170, 2006.
33. F. Varela, E. Thomson, and E. Rosch. *The embodied mind: Cognitive science and human experience*. MIT Press, 1991.
34. F.J. Varela, P. Bourgine, and P. Dumouchel. *Autonomie et connaissance*. Seuil, 1987.
35. C.H. Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811):563–565, 1942.
36. J.C. Zufferey and D. Floreano. Fly-inspired visual steering of an ultralight indoor aircraft. *IEEE Transactions on Robotics*, 22(1):137–146, 2006.