

Artificial Life, Artificial Unconscious

Mohammad Ali Rahebi

Abstract

The Unconscious is the Body

Michel Serres

Cybernetics and the Cartesian Problem

The way the body works without knowledge or consciousness has been a complex and certainly consequential problem since Descartes at the latest. The question of how we do something without knowing *how* we are in fact doing so was a scandal that drove Cartesians to Occasionalism (a concept that is perhaps more true now than ever, with cloud computing and distributed computation).

Since then, we have come very far indeed but it is in AI that we come to again feel the real force of this problem again. Minsky has famously stated that

Social cognition *qua* language based and mainly representational, is a highly inefficient mode of *signal processing* or computation in general. As Minsky (despite his own strict adherence to the representational models, strangely enough) said:

It's mainly when our other systems start to fail that we engage the special agencies involved with what we call "consciousness."

As I shall try to while discussing the Peircean notion of “community of believers” and the way the Cybernetic schema manages to overcome its necessity, the most efficient manifestations of artificial intelligence are not representational, as is the case with the recent success of neural networks and machine learning. In fact, if we are to investigate artificial *life* instead of artificial *consciousness* (a failed project; consciousness *qua* delay is simply obsolete in machinic terms), we have to look at the body, at the *habitual*, that is to say at the non-representational, non-knowledge-based mechanisms and automations that are not language-based in the least.

Beyond Turing

In a relatively recent book edited by Raffaella Giovagnoli, one of the organizers, *Computing Nature*, the issue of alternative computation models to Turing Machines was broached and discussed to some extent. Here I will argue that Cybernetics is one such alternative model that has been growing, from its origin as a strange interdisciplinary field in the late 1940s, to become the dominant model of computation. Although it is not named as such, the models that are not algorithmic and termination-oriented are all operating on the basis of the “cybernetic schema of intelligibility.” These are very familiar processes for all of us, being ubiquitous in “smart” machines and smart software. Giovagnoli et al describe Cybernetics as computational processes where

The main criterion of success of this computation is not its termination, but its behavior - response to changes, its speed, generality and flexibility, adaptability, and tolerance to noise, error, faults, and damage.

If consciousness is to be taken as linguistic and representational, and as such mediated by the social in its historicity and theoretical bias, Cybernetics would have to be seen as an artificial unconsciousness, and in fact, is much closer to artificial *life* than such AI trends as GOF AI or even most strands of embodied cognition.

In French philosophy, under the influence of Derrida and his reading of Plato, the problem of technology has been tied to that of writing, as some form of Ur-Technics that contains the essence of all future technics *qua* mnemotechnics. This idea has been also taken up by Bernard Stiegler who has taken it further by discussing all technologies as means of “retention” and “protection.” All of this starts with Plato, however, for in his famous *Phaedrus*, he comes to barrage writing as a dangerous method, a supplement to spoken language, which unbeknownst to its users, had deleterious effects on memory as well as on truth and on the *polis*. We are not here concerned with all that; what is of interest is something he mentions in passing, namely that if you were to write down something that is true at one moment (‘it is day’), it would become false the next (as night fell) because unlike the speaker, it did not have the ability to correct itself according to its surrounding reality, its environment. In the same manner, he says that when you write down the teachings of some philosopher, the written text will not be able to answer new questions or clarify obscure passage when asked to do so, in comparison to the person using living language.

This is in fact the same thing that distinguishes Cybernetics completely from all the technics that came before: cybernetic machines (think of your smartphone) *can* change their displayed content *and* their behavior in response to changes in their environment. Unlike previous forms of technology, which might have been mnemotechnics (or not), cybernetic technologies are adaptive, self-modifying, robust. In fact, this same Platonic problem occurs in Descartes (of whom we will not have time to speak) and later, Turing. In his famous “On Computation,” Turing states that although certain numbers might be

uncomputable for the Turing Machine, they might be very well computable for human mathematicians. The reason for this is that the human mathematician is capable of revising and changing “their strategy” completely from the ground up, while the Turing Machine, the representational computing machine, cannot change its own behavior when faced with an unsolvable problem. It does not have the ability to *change* its actions *spontaneously* and in response to its specific problem-environment. This is of course recognizable as the same Platonic problem that we encountered with writing and other technological artifacts. They are Leibnizian in principle, relying on some pre-established harmony, on the stability of the operations in the environment and thus are rendered absolutely inoperative once the smallest change occurs.

The cybernetic machine, moreover, does not start from a human-defined state or representation, but is an immanent, non-representational computing machine whose operations (e.g. in case of a neural network running Big Data analyses) cannot be even comprehended by a human observer. As long as there be no need for the cybernetic machine to interact with a human, there is no need for representational information; data is much more efficient.

Peirce, Habits, Neural Networks

Finally, I will present some of my current research on the relation between the American Pragmatist philosophies, the very important yet under-valued concept of *habits*, and the recently most successful manifestation of the cybernetic schema, namely Neural Networks.

The importance of the Pragmatists lies in their discovery of the importance of habits (for they too were early fans of feedback mechanisms, having become familiar with the pre-cybernetic feedback control called the flywheel governor). Yet, though their philosophy managed to move to pure empiricism, it could not solve the notion of intelligent encounter without recourse to social mediation and meaning-creation. This is now solved by the creation of neural networks that are in fact close analogs of Peirce’s community of believers if considered as a community of nodes holding specific weights and bias parameters.