

Proceedings

Progress in Machine Consciousness: Back to Basics

Igor Aleksander

Imperial College, London

Abstract

The basics go back as far as Aristotle, and philosophy carries these roots to the present day. The Aristotelian soul is strongly linked to anything that can be said to be alive (measured by acting to feed and reproduce), and within this are sub categories of plants, animals and people. Only the last of these have what was only in the 17th century called 'consciousness'. While more recently the possible consciousness of animals has been recognised, there remains something special about the human version. So it is not surprising that those who study consciousness through machines or even claim that some form of consciousness can arise in machines have a tough time as they are breaking one of the deepest tenets of all philosophy – that of the Aristotelian soul.

For the MC community this may or may not be of great concern – there is, however, a deeper danger. The danger is that a group of 'Machine Consciousness' experts might pat each other on their back for their splendid programs and computational theories, without really addressing the question of what it is for an object, living or otherwise, to be conscious, and hence shedding light on this problem as founded in ancient science.

The 2001 meeting at Cold Spring Harbour is often quoted as having given legitimacy to machine approaches within the framework of the Science of Consciousness through Christof Koch's optimism that

“...we know of no fundamental law or principle operating in this universe that forbids the existence of subjective feelings in artefacts designed or evolved by humans.”

To progress beyond this mere statement of belief presents the MC practitioner with several challenges amongst which there are the following:

1. What licence exists for any non-living organism to be called conscious?
2. Given a conscious artefact, how does it differ from a non conscious one?
3. What are the engineering advantages of using the language of consciousness?
4. What procedures are required to discover consciousness in a complex system?
5. Is phenomenology necessary to establish consciousness or is behaviour sufficient?
6. What specific computational theories address consciousness?
7. What specific neuro-scientific theories address consciousness?
8. What specific philosophical advances occur through considering MC?
9.

In this presentation I shall argue in the context of my own work that 1) no such license exists, but that keeping a clear distinction between living and artificial consciousness, mainly in terms of content and acquisition, is healthy for both. In the context of 2) and 5) I contend that a proven presence of a phenomenal representation is essential. Further, in 3), I suggest consciousness divested of ancient mystery is a good descriptor of a class of systems that operate on the basis of phenomenal representations of their environments, and in 4) well-trodden axiomatic ideas that map sensation into mechanism may have to be supplemented by theoretical notions such as Tononi's Φ measure of integration. For 6) I wish to stress the advantages of virtual machine phenomenology, leaving 7) and 8) to others.

It seems clear from the contributions at this conference that most workers in MC are searching for greater maturity in their own work. Particularly welcome are the introduction of externalist ideas in robotics, questions regarding the importance of embodiment, further studies into synthetic phenomenology, the functional meaning of qualia and affective issues. These points were mainly seen as question marks in 2001 and the Nokia conference is a timely test of whether some convincing, usable but previously unestablished notions are beginning to appear in the drive towards a sustainable paradigm of Machine Consciousness.