

Might a LIDA Controlled Robot be Phenomenally Conscious?

Uma Ramamurthy

Institute for Intelligent Systems, University of Memphis, Memphis, Tennessee, USA.

Abstract

The LIDA architecture is presented as a possible control system for a robot. LIDA is a complex software agent that models a broad swath of human cognition including “consciousness” in the sense of implementing Global Workspace Theory (Franklin & Graesser 1999). A robot based on such a control structure is said to be functionally conscious, without the subjective experience of qualia. This paper explores a possible way to make such a robot phenomenally conscious.

1. Introduction

The current leading cognitive theory of consciousness, Global Workspace Theory (Baars 1988, 2003), postulates that the primary functions of consciousness include a global broadcast serving to recruit internal resources to deal with the current situation and to modulate several types of learning. Be it human, animal or artificial, an agent is said to be *functionally consciousness* if its control structure (mind) implements Global Workspace Theory. We would consider humans, many animals (Seth, Baars, and Edelman 2005) and even some virtual or robotic agents (Franklin 2003, Shanahan 2006) to be functionally consciousness. We must carefully distinguish functional consciousness from the usual use of “consciousness,” which assumes phenomenal experience, the subjective experience of qualia. To keep this distinction clear we will refer to consciousness in this usual usage as *phenomenal consciousness*.

2. The LIDA Architecture

IDA (Intelligent Distribution Agent) is an intelligent software agent developed for the US Navy (Franklin et al. 1998). At the end of each sailor's tour of duty, he or she is assigned to a new billet. This complex assignment process is called distribution. The Navy employs about 300 trained people, called detailers, full time to effect these new assignments. IDA facilitates this process by completely automating the role of the human detailer. LIDA (Learning IDA) is a conceptual, and partially computational, cognitive architecture (Franklin et al. 2007), derived from IDA primarily by adding several modes of learning. LIDA exhibits both external and internal voluntary action selection, as well as consciously mediated action selection of both the internal and external variety. She uses her “consciousness” module to handle routine problems with novel content. LIDA makes a good case for the claim that she is functionally conscious. But, is LIDA phenomenally consciousness? It seems that LIDA implements part, but not all, of consciousness. What needs to be added to an LIDA based software agent to achieve phenomenal consciousness? We have no definitive answer to this question. However, we do have a conjecture as to at least part of the answer.

3. Merker's Evolutionary Pressure for Phenomenal Consciousness

Bjorn Merker suggested one selection pressure that may have served to increase the evolutionary fitness of phenomenal consciousness in us humans, and other conscious animals. He points out that phenomenal consciousness produces a stable, coherent perceptual world for each such animal by distinguishing real motion in the world from apparent motion produced by the movement of sensory receptors, and by suppressing awareness of the latter (Merker 2005). One can experience the loss of this stable, coherent sensory world by a simple experiment. Close one eye and press gently with an index finger on the lower eyelid of the open eye. The movement of the eyeball produces an apparent motion of whatever is present in the experimenter's perceptual field. Merker does not claim that phenomenal consciousness is the only process capable of producing such a stable, coherent perceptual world. Nor does he claim that this process of distinguishing and suppressing apparent motion provides the only evolutionary selection pressure. He simply

suggests that providing perceptual stability and coherency is one fitness benefit of phenomenal consciousness.

4. A Perceptually Stable and Coherent LIDA Controlled Robot

In a commentary on Merker's article, Franklin suggested that producing a robot provided with a stable, coherent perceptual world might be a step toward a phenomenally conscious machine (2005). A sense organ is said to be spatially sensitive if movement of the organ produces apparent motion at its surface independent of what is happening in the environment. Every autonomous, mobile robot will require spatially sensitive sensory mechanisms, for example vision, for moving appropriately in its world. Thus, the problem of distinguishing real motion from self-produced, apparent motion will be ubiquitous among such robots. One solution would be to build in mechanisms to shield the robot's action selection from apparent motion self-produced by the movement of its sense organs. Such shielding mechanisms might conceivably be based on any of several different principles. One such principle would have the robot construct its own individual, coherent and stable world, suppressing self-produced apparent motion, as Merker argues that consciousness does for some animals. We propose a LIDA controlled autonomous mobile robot with such a built-in shielding mechanism producing a coherent, stable, perceptual world view.

References

- Baars, Bernard J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baars, B J. (2003). How brain reveals mind: Neural studies support the fundamental role of conscious experience. *Journal of Consciousness Studies* 10: 100–114.
- Cotterill, R M J. 2003. Cyberchild: A simulation test-bed for consciousness studies. *Journal of Consciousness Studies* 10: 31–45.
- Franklin, S., Uma Ramamurthy, Sidney K. D'Mello, Lee McCauley, Aregahegn Negatu, Rodrigo Silva L., and Vivek Datla, (2007). LIDA: A Computational Model of Global Workspace Theory and Developmental Learning by AAAI 2007 Fall Symposium - AI and Consciousness: Theoretical Foundations and Current Approaches, Washington, D.C., November 2007.
- Franklin, Stan. (2005). Evolutionary pressures and a stable world for animals and robots: A commentary on Merker. *Consciousness and Cognition* 14: 115–118.
- Franklin, Stan. (2003). IDA: A conscious artifact? *Journal of Consciousness Studies* 10: 47–66.
- Franklin, S., and A. Graesser. (1999). A Software Agent Model of Consciousness. *Consciousness and Cognition* 8:285-305.
- Franklin, S., A. Kelemen, and L. McCauley. (1998). IDA: A Cognitive Agent Architecture. In *IEEE Conf on Systems, Man and Cybernetics*. : IEEE Press.
- Merker, Bjorn. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition* 14: 89–114.
- Seth, A K, B J Baars, and D B Edelman. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition* 14: 119–139.
- Shanahan, M P. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Consciousness and Cognition* 15: 433-449.