

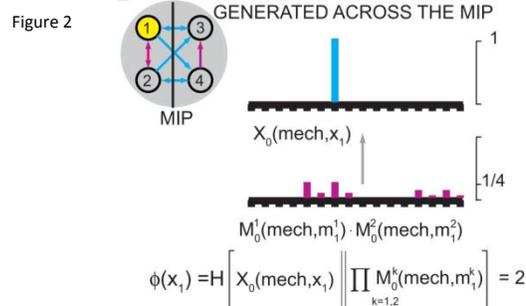
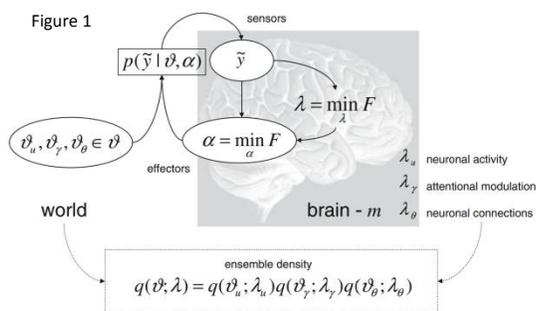
# What's it like to be a non-zombie machine?<sup>1</sup>

(Human misunderstanding of mechanisms and what it will take to transcend it)

Stephen Deiss, Engineer, UCSD CSE Dept.

In this presentation I will recast the 'hard problem' of consciousness in terms of sensations and their interpretation and describe how the ongoing hard problem conundrum results from historical biases about what a machine is. I explain how a broader conception of consciousness and mechanism aligns with modern theories of free energy and hierarchical predictive coding (Friston, Clark, Figure 1) and integrated information theory (Tononi, Figure 2). Understanding consciousness, whether in machines or brains, is largely a conceptual problem, and it requires a conceptual solution (provided here). Nevertheless, part of this deconstructive process will eventually find support in appearance of machines that can at least mimic what presumed conscious minds can do. Computational approaches continue to be domain limited and lack generalization with the prime reason being a lack of means to build scalable associative memories that enable rich perception and cognition. The usual approach is hybrid neuromorphic emulation or digital brain simulation because of the existence proof we have in our own heads and available tools. However, the well known wiring and message passing impasse of these approaches suggests the need for cognitive scientists and neuroengineers to focus on the problem of designing new and better kinds of physical associative memory, whether digital or analog, and whether neuromorphic or not. A machine that is incapable of 'filling in' its sensations with myriad others from a prediction oriented memory is cursed to operate with a limited conceptual horizon little more engaged with its environment than the proverbial thermostat.

Nonetheless, I will argue that the thermostat does have 'something going on' that only our biases prevent us from accepting. Same goes right on down for a simple hydrogen atom on a continuous scale of sensing and interpreting using whatever memory (inertial state) the system operates with. Looking for a specific phase transition somewhere along this continuum that allows consciousness to emerge has caused diversion and confusion from the real issue which is the associative architecture of our memory and how it is composed from simpler constraint-satisfying inertial systems. Once one accepts that consciousness can be fundamental in nature (Penrose, Tegmark, Tononi, Hoffman, Skrbina, Deiss), then the real structural substrates can be sought without misplaced anthropomorphic demands or illusions.



Friston and Stephan, Free-energy and the brain, *Synthese*, 159:417, 2007.

Clark, Whatever next? Predictive brains, situated agents, and the future of cognitive science, *Beh. & Brain Sci.*, 36:3, p1-73, 2013.

Balduzzi and Tononi, Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework, *PLOS: Comp. Biol.*, V4, N6, 2008.

Hameroff and Penrose, Consciousness in the universe: A review of 'Orch-OR' theory, *Phys. Of Life Rev.*, V11, N1, 2013.

Tegmark, Consciousness as a State of Matter, *arXiv:1401.1219*, 2014.

Hoffman, The interface theory of perception: Natural selection drives true perception to swift extinction, *Object categorization: Computer and human vision perspectives*, Dickenson et al eds, Cambridge Univ. Press, 2009.

Deiss, Universal Correlates of Consciousness (UNCC), Chapter 7 in *Mind that Abides: Panpsychism in the new millennium*, Skrbina, ed., John Benjamins, 2009.

Deiss, Time, Consciousness and the Foundations of Science, *Jnl. Cons. Expl. & Res.*, V1, N5, 2010.

<sup>1</sup> What...You don't already know?