What I Am
The Self as a Dynamic Data Structure Implemented Within a Cognitive Framework by a Functional System

Matthias Scheutz
School of Computer Science
The University of Birmingham
U.K.

m.scheutz@cs.bham.ac.uk

PSYCHE, 7(19), October 2001
http://psyche.cs.monash.edu.au/v7/psyche-7-19-scheutz.html

KEYWORDS: consciousness, functional organization, algorithmic information theory, mental states, qualia, implementation, supervenience.

REVIEW OF: Gregory R. Mulhauser (1998). *Mind out of Matter: Topics in the Physical Foundations of Consciousness and Cognition.* Dordrecht: Kluwer Academic Publishers. $147.50 hbk. 275 pp. ISBN 0792351037.

If you, like me, always had the suspicion that there should be a straightforward (but possibly very intricate) story about how mind arises out of matter, if you always viewed perspectival problems as not forcing one to abandon physicalism, if you always felt that first person phenomenal experiences convincingly do not hint at some strange sort of unbridgeable explanatory gap, in short, if you believe that mind has a place in the physical world and the mind-body problem is a relic of early attempts to understand the relation between mind and matter, Gregory R. Mulhauser's *Mind out of Matter: Topics in the Physical Foundations of Consciousness and Cognition* is for you. Free of purposefully sophistic arguments and artificially technical jargon still permeating some of the philosophical literature on the mind-body problem, Mulhauser's book touches on many rock-bottom issues in the philosophy of mind in a refreshingly novel way and does not leave a single stone unturned. Among the issues he addresses are "the difference between first person and third person perspective", "the (best) functional description of a given physical system", "the nature of phenomenal mental states", "the relation between consciousness and cognition", "the relation between representations and what they represent", "computational vs. dynamical systems descriptions of cognitive systems", "why consciousness need not play any role in quantum measurements", "the limits of

computability and physics", etc. Furthermore, the selection of zombie constructions in philosophy is enriched by another thought experiment as is the set of qualia attributes by the terms "chopped" and "frozen". Standard problems (e.g., the conceivability of zombies, the symbol grounding problem, or Jackson's neurophysiologist Mary, who has never seen a red object, but knows everything about the physics of seeing the color red) are revisited and receive an interesting new twist once seen through the glasses of algorithmic information theory.

In general, Chaitin's notion of "algorithmic information content of an object x" defined as "the length of the minimal self delimiting program(s) of a Universal Turing Machine which will generate a description of the that object for a given level of precision" takes center stage throughout the book. It figures crucially in many of Mulhauser's arguments, in particular, in his attempts to debunk difficulties with common approaches to the notion of "implementation of a functional system" and in his proposed notion of "representation", which is at the heart of what he calls "conscious data structures". In a sense, the whole book can be seen as an attempt to get a handle on the notion of "functional system", i.e., how a physical system can be described as having a particular kind of functional organization, and on the notion of "self-model", the seat of the "I", a dynamically changing representation of (parts of) the organism and the environment within an organism, and how it could be realized in a neural net.

Mulhauser's demonstration of what is wrong with naive "correspondence views of implementation" is on the right track, despite the fact that some of his arguments against Chalmers' view on implementation are too emotional and based on misunderstandings. It exposes the crucial problem of unrestricted translations from one domain to another (e.g., from a physical to a functional domain), namely that the complexity of the translation function could exceed that of either domain, hence that in the worst case all complexity is "hidden" in that function, while the mutual information of both domains is zero. Mulhauser's remedy is to construct a functional system implemented by a physical system from its physical description instead of defining a correspondence function. He introduces the notion of "functional logical depth" to measure the complexity of a process that takes certain inputs and produces certain outputs. A three-phase application of his notion of functional logical depth to a minimal description of a physical system given a set of inputs and outputs then attempts to "choose the simplest functional modules which can be interconnected in the simplest ways to give a system whose overall input/output behavior matches that which we're describing" (p. 87). Unfortunately, the details of this procedure are completely left out: neither is it clear what a minimal description of the system is (e.g. what would a minimal description of a human brain be?), nor is it clear what the parts of the system are that can be selected as modules (Mulhauser suggests particles, but without any argument). Yet, both are required to get the process off the ground. It would have been helpful to have an example of how this procedure is supposed to work, but it might be extremely hard to come up with even a simple example, thus rendering the whole strategy completely impracticable. On the other hand, Mulhauser himself admits that

> the point has never been to provide an actual recipe for producing a particular best description or functional decomposition of a given system. [...] It has been merely to show, in a way which is neither trivial nor vacuous, that there is one-and, moreover, that it is essentially unique. [...] But finding that best decomposition [...] remains an empirical goal. (p. 96)

Whether Mulhauser's strategy for extracting functional parts of a physical system is satisfactory as it stands, is open. It would have to be made much more precise (as almost all notions in any of the three steps, including the underlying notion of "functional logical depth", are underspecified) before a final judgement can be reached. Yet, it seems a worth-while endeavour to embark on this project.

Algorithmic information theory finds another application in Mulhauser's definition of representation: "Two objects represent each other to the extent that they are not algorithmically independent-equivalently, to the extent that they have substantial mutual information content" (p. 42). Variants of this notion of representation are at the center of the picture of self and consciousness, the "self model", which Mulhauser sets out to elaborate in an attempt to fill the explanatory role necessitated by the claim (which he argues for) that consciousness supervenes on cognition (which in turn supervenes on functional organization as well as the physical world). The self model is a "dynamic data structure implemented within a cognitive framework by a functional system. I propose the self model as the seat of conscious experience; I am a self model. Phenomenal experience is effected by change in the self model" (p. 129). According to Mulhauser the self model is similar to a "stack" used in computer systems in that it is also a data structure that does not actively engage in the system's functioning of which it is a part, but merely passively reflects changes in the functional system. Quite untypical for this kind of book, Mulhauser reviews some neural network theory and even suggests a type of neural network as a candidate for his self model, Grossberg's adaptive resonance model. While details, as in the case of functional systems, are missing, the direction itself seems very promising (as witnessed by various recent publications by leading neuroscientists, e.g., Damasio). Most likely, however, the breath of Mulhauser's discussion here will cause quite a few objections regarding the depth of some of his suggestions.

But even if one does not agree with Mulhauser's grand picture about the conscious self as a dynamic data structure, *Mind out of Matter* is worth reading for the bits and pieces of wisdom alone scattered throughout the whole book. Here are a few exemplary tastes:

> I suggest that any thought whatsoever, whether it 'refers' to something real or imaginary, is limited by its physical instantiation and that its information content [...] is bounded by the information content of its instantiating physical structures (p. 51)

> Human beings simply lack the capacity to dictate the precise states of their brains, and, as it happens, the states of seeing or smelling a rose straightforwardly differ from states of reading information or of working out logical implications of proposition [about roses] (p. 122)

> It is a body of information, physically instantiated by a functional system and changing dynamically with that functional system, which is conscious. On this view, phenomenal experience is an immediate feature of this change: it is 'what it is like to be' that body of changing, physically instantiated information (p. 130)
>
> Chaos is an interesting trait of some kinds of dynamical systems which may be relevant for explaining observed behavior of some cognizers, but it generally affords no opportunity to import noncomputability into that behavior" (p. 223), or "The basic message of functionalism as an explanatory approach survives essentially unscathed by super-Turing computation (p. 233).

There are many ways to read Mulhauser's book, and the author explicitly facilitates them by an abundance of cross-references, summaries of the state of the argument, excerpts and road maps of the book-at times this benevolent aid is more distracting than helpful. One quick way to check if one's interest could be sparked is to read "A Partial Picture in Soft Focus" (section 2 of chapter 10) first, which provides a succinct overview of the major themes and arguments on about four pages, followed by chapter one. By then, you know Mulhauser's story line and either you will be put off or -- and this is more likely -- you will be hooked and curious to discover the details of the plot.

Acknowledgements

This work is funded by a grant from the Leverhulme Foundation.