

Douglas Hofstadter: Going loopy over consciousness

Source: <http://sci.tech-archive.net/Archive/sci.physics/2007-07/msg00056.html>

- *From:* Sam Wormley <swormley1@xxxxxxxxx>
 - *Date:* Mon, 02 Jul 2007 02:48:46 GMT
-

Going loopy over consciousness
<http://physicsweb.org/articles/world/20/7/5>

Reviews: July 2007

I Am a Strange Loop
Douglas Hofstadter
2007 Basic Books
412pp £14.99/\$26.95 hb

Douglas Hofstadter's writing talent makes his love of paradox contagious. Reading *I Am a Strange Loop* inclines one to see whimsical connections, language games and self-reference everywhere. Part of Hofstadter invades one's brain and starts thinking there in its own right — a phenomenon that is itself a theme of the book. Hofstadter, therefore, is in effect co-writing this review, inclining it towards paradox. Which may be why my method of urging you to read the book will itself be paradoxical: I shall summarize why I find it ultimately unconvincing.

Hofstadter, a professor of cognitive science at Indiana University, expresses disappointment that his 1979 masterpiece *Gödel, Escher, Bach* (one of my favourite books) was not recognized as explaining the true nature of consciousness, or "I"-ness. I have to confess that it never occurred to me that it was intended to do so. I thought it merely explained the problem, highlighting stark flaws in common-sense ideas about minds. It also surveyed the infinite depth and meaning that can exist in "mere" computer programs. One could only emerge from the book (or so I thought) concluding that brains must in essence be computers, and consciousness an attribute of certain programs — and that discovering exactly what attribute is an urgent problem for philosophy and computer science. Hofstadter agrees with the first two conclusions but not the third; he considers that problem solved.

I Am a Strange Loop is supposed to restate and explain his solution: in short, that a mind is a near-infinitely extendable, self-referential loop of symbols that suffers — or rather, benefits

Douglas Hofstadter: Going loopy over consciousness

— from the hallucination of being an "I". Furthermore (Hofstadter says paradoxically), that hallucination is itself an "I".

Hofstadter's "strange loop" is a bit like an ordinary feedback loop, such as the images in a pair of parallel mirrors facing each other, but instead of merely depicting itself physically, it symbolically refers to itself. And unlike ordinary self-referential statements, like this one, the symbol inside a brain that refers to itself as "I" is not used by anyone else: it is someone.

Strangely, Hofstadter's half of this theory of consciousness (the loopy half), is quite convincing. The unconvincing half is essentially philosopher Daniel Dennett's theory from his book *Consciousness Explained* (which critics have justly renamed *Consciousness Denied*) — namely that our opinion that we are conscious is simply mistaken. Hofstadter calls it the "I myth". We can, of course, be mistaken about anything, so here Dennett laid down a valuable marker: the true explanation of consciousness will have to refute his position.

Hofstadter is a master of analogy and metaphor, which abound in this book. One of his metaphors is that of a soul (but devoid of religious connotations — these souls are unequivocally aspects of the brain) and, daringly, the idea of differently sized souls corresponding to degrees of consciousness. Children have smaller souls than adults, he says; animals have tiny (but non-zero) souls; the Franco-German philosopher and humanitarian Albert Schweitzer's soul was bigger than yours or mine. But Hofstadter's arguments for his analogies are, frustratingly, sometimes compelling but often absent.

The central analogy is between minds and other "strange loops": certain self-referential statements discovered by Kurt Gödel within formal mathematical systems. These statements assert their own unprovability within the system but are nevertheless provably true, akin to the paradoxical "this statement is false". And the way Gödel's proof works is by showing that certain very large numbers also have another meaning, as statements about numbers; and so a proof about numbers — which is itself just a number — turns into a proof about proofs, and in particular about itself.

The author flits between two somewhat conflicting strands of this analogy. He stresses that human consciousness depends on the universality of our thinking — the fact that we can extend our internal repertoire of symbols indefinitely, and eventually refer to anything at all. But he also draws the lesson that self-awareness is the heart of the matter. I do not see why. Most of my conscious thought is not about me. Gödelian statements refer meaningfully to themselves, but are not conscious. Universality implies the ability to contemplate oneself, but the converse is not true.

Correspondingly, Hofstadter does not seem to be able to decide whether animal minds are merely quantitatively inferior

Douglas Hofstadter: Going loopy over consciousness

("small-souled") or qualitatively. On the one hand he says that the "huge and fundamental breach between humans and...all other species...makes us unique, and...gives us what we call 'souls'". Yet, on the other : "to argue...that the word 'soul' does not even apply to animals...seems to me more like received dogma than like mature reflection". I think Hofstadter was right the first time: animals are not miniature people but are fundamentally different, and unmysterious, things. They cannot create new meanings at all because they lack the as-yet-unknown attribute of human brains that gives them universality.

The more Hofstadter invokes souls, feelings and animals — and the less he discusses computers, mathematics and meaning — the more, it seems to me, emotion replaces reason. For instance, what is his evidence for Schweitzer's oversized soul? Firstly, that Schweitzer empathized with insects. So if a soul is measured by its empathy with the small-souled, would an even greater soul empathize with cucumbers? Secondly, Schweitzer loved Bach's organ music, and musical taste is apparently a soul-size indicator. Hofstadter's argument for that? Absent. The animal theme culminates in a veritable celebration of sentimental anthropomorphism, describing Hofstadter's own "ability to mirror the interiorities" of grasshoppers and ants while listening to the music of Bach.

Hofstadter argues that emergent entities (such as people) and abstract concepts (such as numbers, and meanings) really do have causal effects on the microscopic constituents of events. He imagines a computer made of toppling dominoes that is designed to factorize integers. It is presented with the input "641" and set in motion to perform its computation. Why is one particular domino left standing? The most fundamental explanation does not refer to the sequence in which the other dominoes fell; rather it is "because 641 is prime".

I see no escape from this argument: regarding microphysical explanations as more fundamental than emergent ones is arbitrary and fallacious. Yet, from Hofstadter's point of view, I seem to have catastrophically missed the point, for he eventually disowns the argument. Consciousness (in its guise as free will), he says, cannot "push material stuff around" because "physical law alone would suffice to determine [its] behaviour". But physical laws can't push anything! They are just predictions and explanations — and by no means our only ones. Here I wondered what the point of the "641" argument was in the first place, and indeed of the whole book. Finally, Hofstadter embraces irrationality itself: "Our very nature is such as to prevent us from understanding our nature".

I judge claims to understand consciousness largely by this question: can you use that understanding to create an artificial-intelligence program? Judged by that criterion, Hofstadter does not have the answer. However, his claim that our nature prevents us from understanding our nature cannot be taken at face value. Like a

Douglas Hofstadter: Going loopy over consciousness

Gödelian claim to be unprovable, it applies only inside the system from which it is derived, namely Hofstadter's own philosophical framework. But, again like Gödel's construction, this simultaneously reveals that there is a truth to be discovered outside of that framework.

Something new is needed to discover that truth, and Hofstadter's loops are probably involved. "Strange loopiness" is a distinctive form of emergence, rooted not in complexity but in universality, the real substrate of "I"-ness. That is why, if you want to understand what an "I" is — what you yourself are — you should want to read this book. Unless your soul is too small.