BOOK REVIEW

William W. Davis

Douglas R. Hofstadter: Godel, Escher, Bach: an Eternal Golden Braid. Pp. xxi + 777. New York: Basic Books, Inc., Publishers, 1979. Hardcover, $18.50.

> This is, it seems to me, a general principle: You get bored with something not when you have exhausted its repertoire of behavior, but when you have mapped out the limit of the space that contains its behavior. (Hofstadter, p. 621)

Few, if any, readers can claim to have mapped out the behavior of Godel, Escher, Bach, simply because the themes of this multi-dimensional work range from Lewis Carroll-like dialogues and Zen koans to typogenetics and Godel's Incompleteness Theorem. While the book does contain a wealth of information, the reader will notice immediately that Professor Hofstadter offers more conjectures and suggestions, more illustrations and metaphors, than he does either argued conclusions or sound analogies. But the author does not claim to have crafted a tightly argued piece of analytic philosophy. On the contrary, in order to fully enjoy this book, the reader ought to carefully heed the book's subtitle: A Metaphorical Fugue on Minds and Machines in the Spirit of Lewis Carroll.

Nevertheless, the reader may feel somewhat confused or fatigued upon concluding Part I of Godel, Escher,

Bach. The focus of attention is the introduction of Godel's Theorem, which is presented as the culmination of a discussion concerning formal systems, interpretations and the development of the propositional and predicate calculus. Those who are unfamiliar with formal systems will find the discussion in Part I both interesting and helpful, while others might insist that the entire discussion is just too drawn out. But Part II of GEB promises something of interest to all. The discussion of formal systems is continued, and some of the most interesting developments in metamathematics are surveyed. The topics of special interest include the Church-Turing Thesis, Church's Theorem, Henkin sentences, and, of course, Godel's Incompleteness Theorem. These results are then applied to topics and problems in the field of computer systems and Artificial Intelligence.

The basic theme of the book centers upon the most interesting aspects of Godel's Theorem, namely, the notions of self-reference and levels of explanation and description which whereupon result. The author attempts to show how these notions are exemplified by Bach's Musical Offering and the drawings of M. C. Escher. In addition, Hofstadter extends his discussion of Godel's Theorem to an examination of computer systems and languages, and then to the domain of thought itself.

Of special interest to both the philosopher and the theorist in Artificial Intelligence alike is Hofstadter's discussion of the relation of thought processes to neurophysiological theory. According to Hofstadter, the fundamental question for the philosophically inclined theorist in Artificial Intelligence is: how does one reconcile the "software" of the mind with the "hardware" of the brain? This problem of reconciling mind and brain involves, as Hofstadter sees it, two questions: (1) How do neuron firings give rise to symbol activation? (2) Can symbol activation be described in terms which neither presuppose, nor refer to, neurophysiological entities and processes? (Hofstadter, pp. 302-9). An affirmative answer to (2) would mean that thought processes could, in principle, be realized in more than one kind of "hardware"; in, for example, a computer circuit (Hofstadter, p. 358). Question (1) is, of course, an old question in a new guise, and it is cleverly paraphrased by Hofstadter as the problem of

> I think, therefore I have no access to the level where I sum. (Hofstadter, p. 677).

Hofstadter's sensitivity to philosophical issues is best illustrated by his attempt to locate the problem presented by the paraphrase. First, he urges the reader to realize that, theoretically, human agents admit of descriptions on the micro-physical, the chemical, the biological, and the macro-behavioral level. The question here is whether or not these description levels converge, and, if so, at what point? (Hofstadter, p. 285). Specifically, Hofstadter asks us to suppose that thinking is in some sense a "flexible, intensional representation of the world" (Hofstadter, pp. 337-339). But then, Hofstadter notes, to ask how thoughts fit with neurophysiological theory is tantamount to forcing the element of intensional representation upon a micro-theoretical framework of the brain's activity. Instead, the question ought to be posed as follows: are neurophysiological processes analogous to conceptual representings, and, if so, in what respects? What corresponds to concepts pertaining to thoughts on the personal level of explanation, given a micro-theory of the brain? The answer must involve "a description which relates neural activity to 'signals' (intermediate-level phenomena)--and which relates signals, in turn, to 'symbols' and 'subsystems', including the presumed-to-exist 'self-symbol'" (Hofstadter, p. 709).

But how does one proceed with this description? What model is best suited for reconciling the software of thought with the hardware of the brain? Given Hofstadter's explicit recognition of the worth of analogy and his close connection with the computer field, one might expect him to offer a functionalist model for the problem of mind-brain. There are indeed many hints in this direction (see Hofstadter, pp. 49, 147). For one, in his discussion of what might be called the "other-brains" problem, Hofstadter suggests that an isomorphism or functional mapping could be established between brains which is based upon the similarity of their respective symbol repertoire and their symbol triggering mechanisms (Hofstadter, pp. 369, 375). Such a recommendation would provide a suitable basis for the functionalist approach.

However, Hofstadter's model for the mind-brain relationship is not that of the logical-structural states of a computing machine. Hofstadter is instead concerned with examining the implications of Godel's results upon the problem of mind-brain. His question is thus: do words and thoughts follow formal rules? (Hofstadter, p. 46). His answer is based upon what I will call the formal systems model, and his suggestion

is that brains can be likened to a formal system. The laws of neurophysiology are viewed as analogical extensions of the formation and transformation rules of a formal system (Hofstadter, p. 559; also "Typogenetics," Chapter XVI). Thought processes are thereby characterized as the "informal, overt, software level" which "floats" on a substrate (here, the brain). Thoughts are said to "depend" upon a "formal, hidden, hardware level" and are grounded, in this case, by neurophysiological entities and processes. Of course, the philosopher would like to see such metaphors unpacked. But the attempt to do so, Hofstadter argues, runs into the problem called the "Reductionist's Dilemma" (Hofstadter, p. 552). One horn of this dilemma has already been noted, namely, that if the characterization of thought is sought on the micro-level of the brain's activity, then the intensional, representational aspect of thought will be lost. Alternately, should one proceed from the other direction--from neuron firings to the description of, say problem solving--then one will lose the determinacy and the preciseness which is achieved at the micro-level.

While Hofstadter's "levels" account of micro- and macro-descriptions is an illuminating result of his formal systems model, it is not sufficient for the reconciliation of the software of thought and the hardware of the brain. Ultimately, the image of thoughts "floating" on the hardware of the brain remains at the metaphorical level. The reader is left with the impression that conceptual representation is somehow an emergent phenomenon, a function of some hitherto unknown, and yet to be discovered, description of a (non-conceptual) state of the brain.

My suggestion is that the formal systems model is too restrictive for an adequate solution to the mind-brain problem. This conjecture may seem surprising, especially given the author's enthusiasm for the method of analogy. Hofstadter writes: "One can think of the Bongard (Pattern Recognition) problem-world as a tiny place where 'science' is done--that is, where the purpose is to discern patterns in the world." "As patterns are sought, templates are made, unmade, and remade; slots are shifted from one level of generality to another; filtering and focusing are done, and so on" (Hofstadter, pp., 659-650). In light of these remarks, I suggest that a solution to the mind-brain reconciliation problem must involve in Hofstadter's words, the "remaking of a template." The formal systems model ought to be extended to at least include

an analogy between overt speech and thought as they apply to the personal level of explanation and description. For despite Hofstadter's extended discussion of symbols and information, his model ignores this important analogical dimension for thought. Indeed, if one of the basic questions which confronts the AI theorist is "what is thought?," why not begin with a paraphrase of the Sellarsian challenge--what are thoughts except something analogous to the way sentences are related to each other and to their contexts? A functional characterization of thoughts as analogues to overt speech would, in addition, leave open the question concerning the qualitative nature of thoughts. They could be neurophysiological entities, but there would be no reason, in principle, why thoughts could not be grounded by other forms of "hardware."

Hofstadter suggests, basing his discussion upon the formal systems model, that intensional phenomena (thoughts, ideas, images) can be described in terms of a Strange Loop, a process of level interaction which is an analogical extension of the Loop phenomenon in computer science. A Strange Loop is like a Henkin sentence, a sentence which asserts its provability and thereby becomes provable. Furthermore, as Hofstadter writes:

> This act of translation from low-level physical hardware to high-level psychological software is analogous to the translation of number theoretical statements into metamathematical statements. (709)

Likewise, Hofstadter's solution to the problem of self-awareness is based upon his formal systems model and, particularly, upon the self-referentiality of a Godel sentence in the predicate calculus (Hofstadter, pp. 406, 570). While the philosopher may suggest that the key to understanding self-awareness is to be found in the location of the similarities and the differences between self-awareness and other kinds of awareness (for example, perceptual knowledge), Hofstadter concludes that "the isomorphism which mirrors Typographical Number Theory inside the abstract realm of numbers can be likened to the quasi-isomorphism that mirrors the real world inside our brains, by means of symbols" (Hofstadter, p., 502).

These are but a few of Hofstadter's many analogies between formal systems and thought. And they are,

without a doubt, illuminating and interesting. I have questioned their comprehensive nature in relation to the mind-brain reconciliation problem. I have also emphasized the method of analogy as a valuable tool for philosophical explication and theory construction. It is quite unlikely that Hofstadter would share this enthusiasm for the analogical approach. In point of fact, Hofstadter is more concerned with capturing the phenomenon of analogical awareness and procedural knowledge with some suitable means of simulation (Hofstadter, pp. 361-363; 619). His very inquiry into analogical awareness (especially, the kind of awareness instantiated by his own mind) leads Hofstadter to the conclusion that Godel, Escher, Bach is "one, big self-referential loop." This is a very interesting result which, I believe, deserves much praise and further attention. But in what direction do we take our inquiry? Shall we look for ways out of the loop, or should we instead investigate the reasons which got us into it in the first place?