Retrospective Review of
   Godel, Escher, Bach: An Eternal Golden Braid
   by Douglas Hofstadter
   Basic Books, 1979.

Ernest Davis, New York University

Recently, while trying, like everyone else, to think deep thoughts about Deep Blue, I recalled a passage from *Gödel, Escher, Bach:* (p. 678)

> Question: Will there be chess programs that can beat anyone?
>
> Speculation: No. There may be programs that can beat anyone at chess, but they will not be exclusively chess programs. They will be programs of *general* intelligence, and they will be just as tempermental as people. "Do you want to play chess?" "No, I'm bored with chess. Let's talk about poetry."

If memory serves and I am not deceived by 20/20 hindsight, this prediction seemed far-fetched to me even in 1979, and it has, of course, been thoroughly disproved by the event. We now have a chess program that can beat almost anyone — soon, perhaps, with a little more speed, some more fine-tuning, one or two more clever search heuristics, a program that can beat absolutely anyone — built on the narrowest possible principles. Deep Blue not only has no idea of talking about poetry; it not only has no capacity for learning or abstraction; but indeed it knows astonishingly little about chess. The program has no idea of strategies, tactics, gambits, or forks; it cannot answer elementary questions about chess such as "If a bishop is now on a black square, will it ever be on a white square?" or "Is there a position in which white can mate black, if it is white's move, and black can mate white, if it is white's move?" or even "What is checkmate?" All that it can do is to generate a next move of extraordinary quality.

But in a deeper sense, the match between Deep Blue and Kasparov proves that Hofstadter was profoundly right. What Hofstadter is really talking about, the theme that his book pursued in many domains and many forms, is the relation between object levels and meta-levels and the need for a reasoner to be able to move between them. It is precisely this power of abstracting, of stepping outside the game to a higher-level view, that enabled Kasparov to find where Deep Blue was weak and to exploit this to win his match. And even if the next generation of chess program or the following defeats the ability of the human opponent to abstract within the confines of a chess match, the essential point remains: a reasoner who can abstract has creative and adaptive capacities immeasurably beyond those of a reasoner without it.

Having taken *Gödel, Escher, Bach* off the shelf, I indulged myself by rereading it for the first time in years. I enjoyed it and appreciated it more than I did sixteen years ago. At that time, I did not sufficiently appreciate the extraordinary patience and clarity and enjoyability of the presentation of the basic structure of Gödel's theorem and its underlying concepts — formal proof, logic and meta-logic, and self-reference. All this is presented in a form accessible to the lay reader with no cheating or inaccuracy, and with the omission only of the number-theoretic details of Gödel's proof. Having spent the last thirteen years trying to teach and write, I now marvel at Hofstadter's pedagogical skill at presenting this difficult material. The dialogues interspersed between the chapters are, indeed, often contrived, arch, heavy-handed, and excessively cute, but they have virtues that, pedagogically and presentationally if not aesthetically, far outweigh these drawbacks; they are uniformly clear, vivid, memorable, and thought-provoking. (My particular favorite, as a source of contemplation, is "The Magnificrab Indeed", which imagines a creature able to determine mathematical truth through aesthetic judgment.)

The only aspect of the book that really does not work for me is the attempt to integrate music.

The truth is that, since music does not denote, the issues of levels of representation and self-reference really do not arise. Similarly, the attempt to reflect musical structures in the dialogues mars them and is responsible for much of their contrived and arch feeling. All the same I don't actually object to it, because it is so clearly an expression of love on Hofstadter's part. It is like an unsuccessful attempt by a man to introduce old bachelor buddies to a new wife; it may not work, but one can hardly blame him for trying.

Rereading *Gödel, Escher, Bach*, I become nostalgic for the era when it was written. Part of this, of course, is nostalgia for the halcyon years of AI when the funding agencies had more money to throw around than there were researchers to take it, and the universities had more tenure-track slots than there were Ph.D's to fill them; the years when I was young and had the world before me, But more than that, *Gödel, Escher, Bach* reflects the pervasive feeling at the time that computer science generally and AI in particular could contribute to the general intellectual culture not just computational tools, user interface, and fast information retrieval but also ideas and insights of some profundity and some general interest. I encounter this feeling less and less these days.

I wish I could convey this spirit to my students; I would consider it a greater accomplishment than having them master Cormen, Leiserson, and Rivest, or Russell and Norvig or the internals of Windows-95. But I do not assign *Gödel, Escher, Bach* to them. How could I? College education, after all, is not concerned with conveying a spirit; it is concerned with problem sets, programming assignments, and exams. *Gödel, Escher, Bach* is too wild and wooly, too full of tangents, too unconcerned with technique, to serve as a textbook. So I point it out, and hope they read it on their own. I don't think any of them do, any more.

Still more, the effervescent, ebullient energy that bubbles over in *Gödel, Escher, Bach* into an endless, playful torrent of imaginings, inventions, jokes, dreadful puns and wordplay, digressions and diversions into everything under the sun, from DNA to Zen koans — this energy reflects the enthusiasm and creativity of AI in the late 70's. AI has "matured" as a field since 1979; like other types of maturing, this is a decidedly mixed blessing. I remember very well, and have no wish to revive, the methodological sins of the time so brilliantly excoriated by McDermott (1981): the programs, described in loving detail, that never existed or that ran on two examples; the extreme infrequency of any precise evaluation of anything, either empirical or theoretical; the tendency to pile together ill-fitting, inappropriate, and ill-understood computational mechanisms; the gaps of HACKER, the kludges of NOAH, the cheats of AM. But the best work of that time put forward creative approaches to large problems in a way that I rarely see today. We have gotten much more scrupulous about technique, but we have forgotten where we are supposed to be going.

One can see the change in the work of individual researchers. Compare Doug Lenat's Ph.D. thesis (1982) on AM — startlingly original, clear, witty, insightful, inspiring — which his book (1990) on CYC; one can hardly believe that the same person wrote them. Or, more fairly and more revealingly, compare Eugene Charniak's Ph.D. thesis[1] (1974) with his recent book *Statistical Language Learning* (1993). The later book has many advantages over the earlier: the material presented is solidly grounded in statistical theory; it has been implemented and extensively tested; and the book is a model of clear writing. By contrast, the earlier book (if I remember correctly) was unimplemented and often vague. The earlier Charniak did not know what he was doing, whereas the later Charniak knows what he is doing very precisely. But, on the other hand, the doctoral thesis dealt with a profound problem — the relation of knowledge and inference to natural language understanding — and pointed our understanding of this issue in a potentially revolutionary direction. Whereas *Statistical Language Learning* deals with a much narrower and more technical issue — the extraction and use of statistical patterns from vast textual corpora — and, though very valuable, is intellectually quite conservative. I am still romantic enough to believe that, in the long run, Charniak's earlier work will be more influential.

---

[1] I am working from memory here. Charniak's thesis was never published, and I have not seen a copy since 1981.

Gödel, Escher, Bach was a triumphantly successful presentation of quite difficult concepts for a popular audience. There has been nothing else like it in computer science before or since. Why not? AI is full of ideas and insights that could be appreciated by an intelligent general readership: knowledge representation, heuristic search, inductive learning ... Why is it that the only serious popular books discussing AI today are by opponents like Roger Penrose?

References

Eugene Charniak (1974). *Toward a Model of Children's Story Comprehension.* unpublished Ph.D. Thesis, M.I.T.

Eugene Charniak (1993). *Statistical Language Learning.* MIT Press.

Douglas Lenat (1982). "AM: Discovery in Mathematics as Heuristic Search." In R. Davis and D. Lenat (eds.) *Knowledge-Based Systems in Artificial Intelligence.* McGraw Hill

Douglas Lenat and R. V. Guha (1990). *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project.* Addison-Wesley.

Drew V. McDermott (1981). "Artificial Intelligence Meets Natural Stupidity." In J. Haugeland (ed.) *Mind Design: Philosophy, Psychology, Artificial Intelligence.* MIT Press.