BOOK REVIEWS

CHRISTOPHER CHERNIAK. ***Minimal Rationality.*** Cambridge, Mass.: MIT Press, 1986, x + 161 pp. $19.95 (cloth).

It is widely held that rationality is an essential property of agents with beliefs and desires. Indeed, for Cherniak this is "the most basic law of psychology . . .: No rationality, no agent" (p. 1). When it comes to saying what rationality amounts to, however, most philosophers offer accounts that are so idealized that they run head on into what Cherniak takes to be "the next most fundamental psychological principle: Agents are finite objects" (p. 1). Since they are finite objects, they could not possibly instantiate some idealized philosophical accounts of rationality. They could not, for example, believe *all* the logical consequences of their beliefs, because no finite object could deduce them. Nor could they eliminate all inconsistent beliefs, since detecting some inconsistencies would take more time than we have—lots more. So if having beliefs and desires requires idealized rationality, creatures in the "finitary predicament" (and that's all of us) cannot have beliefs and desires.

Cherniak suspects that this paradox encourages some philosophers to view beliefs and desires as less than fully real. Dennett, for example, takes them to be instrumentalistic posits, on an ontological par with the lines in a parallelogram of forces. While for Davidson, beliefs and desires are not the proper subjects of "strict" or "serious" laws that can lead to "accurate" predictions of behavior (p. 4). Cherniak's reaction to the paradox is quite different—and for the cognitive scientist who hopes to incorporate the propositional attitudes of folk psychology into a serious science, it is obviously the right one. He urges that we reject over-idealized philosophical accounts of rationality, and develop in their stead a less demanding model of rationality that can be instantiated by finite creatures.

The central project of this important and intriguing book is to explore just how much, or how little, of the idealized model of rationality we can plausibly take to be a necessary condition for agenthood. "How stupid can you be?" (p. 26) Cherniak asks, and still be the sort of thing that has beliefs and desires? In attempting to locate the necessary-rationality condition in the space of less-than-ideal abilities, a lower limit is provided by the requirement that we be able to use attributions of beliefs and desires, along with various ancillary theories, to predict behavior. Without some minimal deductive and inductive ability, Cherniak argues, an account of what a creature believes and desires is useless in predicting its behavior (p. 6). I am a bit skeptical here, but I don't propose to push the point. For the real focus of Cherniak's interest is not to determine how high the necessary-rationality condition must be set, but to determine how low it must be set. The upper limit, Cherniak argues, is imposed by the finitary predicament along with various facts about the workings of our own cognitive apparatus. It had better not turn out that finitary creatures with cognitive systems like ours have too little deductive ability to count as agents, for if it did it would follow that *we* aren't agents. From this plausible beginning Cherniak draws a sequence of fascinating and unexpected consequences. The upper limit of required deductive ability is going to have to be much lower than almost anyone prior to Cherniak had supposed.

Perhaps the most surprising constraints on how much deductive ability can be expected of an agent emerge in Cherniak's discussion of complexity theory, a branch of mathematics that has hitherto been largely neglected by philosophers. The well-known classic undecidability results of Gödel, Church and others demonstrate the "absolute" impossibility of decision procedures for theoremhood or consistency in various domains. Complexity theory demonstrates that decision procedures in a very much wider class of domains are computationally intractable, and thus practically, if not absolutely, impossible. Consider just a single example. How large a belief set could an "ideal" computer check for consistency by the truth table method?

> Suppose that each line of the truth table for the conjunction of [a set of] beliefs could be checked [by an ideal computer] in the time a light ray takes to traverse the diameter of a proton . . . and suppose that the computer was permitted to run for twenty billion years, the estimated time from the "big bang" dawn of the universe to the present. A belief system containing only 138 logically independent propositions would overwhelm the time resources of this supermachine. (p. 93)

These sorts of results throw an entirely new light on the "surprisingly ubiquitous use of prima facie suboptimal 'heuristic strategies'" (p. 81) that has been uncovered in many recent experiments by psychologists who take themselves to be studying human irrationality. Since people can't possibly use algorithms that guarantee a correct answer, the use of "formally incorrect heuristics need not in fact be irrational at all. They are not just unintelligible or inadvisable sloppiness, because they are a means of avoiding computational paralysis while still doing better than guessing" (p. 82).

Similar conclusions emerge from Cherniak's discussion of the structure of human memory. Idealized accounts of rationality take inconsistency to be a cardinal sin. But, if recent accounts of human memory are on the right track, then, for two quite different reasons, inconsistencies will be abundant and unavoidable. First, it is widely held that human memory consists of two very different subsystems: a large, and largely passive, long term memory, and a much smaller, much more computationally active short term memory. If short term memory is where the computational action is, people are only likely to draw out the implications of their beliefs, and notice contradictions among them, when the beliefs have been retrieved from long term memory and activated in short term memory. But since the capacity of short term memory is limited, the detection and elimination of contradictions is bound to be incomplete. Second, in many models of memory, long term memory is itself compartmentalized:

> Logical relations between beliefs in different "compartments" are less likely to be recognized than relations among beliefs within one compartment, because in the former case the relevant beliefs are less likely to be contemporaneously activated, and . . . it is only when they have been activated together that such relations can be determined. (p. 67)

Moreover, the compartmentalization of memory is not plausibly viewed as a quirk or a flaw, for there would be intractable problems in searching a memory as large as ours, unless the search can be restricted to a few compartments:

> Such memory organization is advisable overall, in the long run, despite its costs. Correspondingly, a person's action may seem irrational when considered in isolation, but it may be rational when it is more globally considered as part of the price of good memory management. (p. 67)

Cherniak's last chapter is in some ways the most fascinating. In it he explores the idea that the program of naturalized epistemology, which has been so elegantly embellished in the first five chapters, contains the seeds of a new kind of skepticism about the limits of knowledge. A common theme in the naturalized epistemology literature is that natural selection has built us with a host of perceptual and cognitive biases without which our cognitive machine would be unable to function. But our "cognitive tool kit" was designed "for a specific terrestrial hunter-gatherer environment of middle sized objects . . . (p. 125). It would be an "odd accident" if a system "pretuned" specifically to this environment turned out to be adequate to fathom the truths of astro- and microphysics. A second reason to suspect that human knowledge may be limited derives from the findings of complexity theory. The range of formal procedures known to be computationally intractable leads one to wonder whether knowledge systems might not become so computationally unwieldy that they will shatter. Indeed, "the universe may be not merely inhumanly complex, but 'transcendentally' unmanageable for any physically realizable entity, for example, an ideal computer occupying the twenty billion light year radius and twenty billion year age of the universe" (p. 129). A disquieting idea, this, and one among many in this unique and

original book which is likely to provoke debate for some time to come. *Stephen P. Stich, University of California, San Diego.*

K. OKRUHLIK AND J. BROWN (Editors). ***The Natural Philosophy of Leibniz.*** Dordrecht: D. Reidel, (1985), viii + 342 pp. $49.50 (cloth).

The ten essays collected here are from two conferences given in 1982, one at the University of Western Ontario, the other at the University of Toronto. The topics covered range from Leibniz's conception of the methodological role of angels and the place in science of miracles (in papers by Robert Butts and Robert McRae, respectively) to his views on the status of scientific laws, the nature of monadic relations, and the metaphysics of mechanics (in papers by Kathleen Okruhlik, Graeme Hunter, and François Duchesneau, respectively). All told, a pleasing farrago.

The centerpiece of the collection—and a full 103 pages of it—is Daniel Garber's "Leibniz and the Foundations of Physics: the Middle Years". Leibniz is well known for having believed something like this: the real world is composed of simple, indivisible, incorporeal substances, each isolated from all the others yet together in perfect harmony, while what we usually take for the real world, everyday physical objects, are merely well-founded phenomena. One upshot of this (in)famous view is that physics, which provides a mechanistic explanation of the behavior of the physical objects of common experience, is cut loose from the real world and reality becomes the private domain of metaphysics. However close this is to Leibniz's position in the *Monadology,* it is fundamentally wrong, Garber argues, when it comes to giving an account of Leibniz's ontology from about 1680 through approximately 1700—the "Middle Years". Relying heavily, although not solely, on the *Correspondence with Arnauld,* Garber attributes to Leibniz the view that the truly real is not limited to soul-like monads but includes quasi-Aristotelian corporeal substance—form and matter united. (Judging from recent work on Leibniz, Garber's reading is already having an impact.) For Leibniz, embracing corporeal substance as real has two advantages; it makes room for the reality of bodies of common experience (which are aggregates of corporeal substance) and it acknowledges science's access to the real by construing physics as giving "an account of the laws that govern a *real world of quasi-Aristotelian substances*" (Garber, p. 28). For us, this interpretation of Leibniz has the advantage of making more comprehensible Leibniz's view (as expressed in the *Discourse on Metaphysics*) that metaphysics serves as the foundation for the true physics and his conviction that philosophy must "do justice to physics" (*Critical Thoughts on the General Part of the Principles of Descartes,* Loemker, 1956, p. 675). The more familiar interpretation of Leibniz's ontology, the one that captures his post-1700 position fairly accurately, makes it hard to see why physics and metaphysics should be so closely linked. In contrast, Garber's interpretation makes clear the intimacy of the link. Befitting its length, Garber's paper is incredibly rich; it certainly deserves the space it is given. As rich as the paper is, it contains an unconvincing account of Leibniz's devotion to the primacy of mechanistic explanation and it stops tantalizingly short of saying why Leibniz abandoned his belief in corporeal substance. Even so, it stops short only after having traveled an extremely long distance. Garber's paper alone justifies this collection.

Happily, the book's justification need not rest with a single excellent essay. All the papers are, at the very least, intriguing, and most are much more than that. Especially provocative are two papers relating to Leibniz's account of space and time: Ian Hacking's "Why Motion Is Only Well-Founded" and Richard Arthur's "Leibniz and Time".

Hacking begins his paper claiming that Leibniz thought of space and time as merely well-founded phenomena because he thought motion was merely phenomenal, and not (as some may have thought) vice versa. Because Leibniz made scant reference to motion in his correspondence with Clarke, there is reason to doubt that Leibniz saw his argument as going in the direction Hacking says it does. Yet Hacking's purpose in this paper is not so much to defend this interpretation of the order of dependence. Instead, he concentrates on explaining why Leibniz thought of motion as only phenomenal. And to do this Hacking sets himself the task of articulating for Leibniz a clear criterion for distinguishing real from