Brian Cantwell Smith

On the Origin of Objects

On the Origin of Objects

On the Origin of Objects

Brian Cantwell Smith

A Bradford Book

The MIT Press
Cambridge, Massachusetts
London, England

Library of Congress Cataloging-in-Publication Data

Smith, Brian Cantwell.
On the origin of objects / Brian Cantwell Smith.
p. cm.
Includes bibliographical references.
ISBN 0-262-19363-9 (hc: alk. paper)
1. Metaphysics. 2. Object (Philosophy). I. Title.
BD111.S573 1996
111—dc20 95-37637
CIP

Contents

Preface

It is difficult to find in life any event which so effectually condenses intense nervous sensation into the shortest possible space of time as does the work of shooting or running an immense rapid. There is no toil, no heart breaking labour about it, but as much coolness, dexterity, and skill as [a person] can throw into the work of hand, eye, and head; knowledge of when to strike and how to do it; knowledge of water and rock, and of the one hundred combinations which rock and water can assume—for these two things, rock and water, taken in the abstract, fail as completely to convey any idea of their fierce embracings in the throes of a rapid as the fire burning quietly in a drawing-room fireplace fails to convey the idea of a house wrapped and sheeted in flames.

—Sir William Francis Butler (1982)

A few years ago, at a workshop on the perception of objects, I raised a question about one of the philosophers' ontological assumptions, which seemed to me altogether too neat. I was concerned about basic issues of object identity—about what makes a newspaper into a single entity as opposed to a large number of separate pieces of paper, what distinguishes the headache that I had this morning from the one I had last night, what it is for Microsoft Word on PCs to be "the same program" as Microsoft Word on the Macintosh. I was particularly troubled by how ready most of the other workshop participants were to base their perceptual theories on the presumption that such ontological questions had clear, determinate answers.

Over lunch, several people pushed the worry aside, as if it were a matter of internal theoretical hygiene. "Don't worry," someone assured me; "most philosophers would insist on precise individuation criteria." Sure enough, most do—I knew that. My question was *why*—whether such presumptions were warranted. And I had a special reason for asking. Having spent more than twenty-five years working in the trenches of practicing computer science, in a long-term effort to develop an empirically responsible theory of computation, I had never met such a logically pure entity, never met such a lapidary individual thing. And computation, alleged to be one of science's most dazzling achievements, is presumably as promising a site as any for ontological clarity and distinctness. Moreover, those few objects that are possessed of anything like crystalline ontological coherence, such as numbers, and perhaps the underlying bits and bytes, are manifest exceptions, no matter how highly prized. By and large, or so at least my experience suggests, the world is an unruly place—much messier than reigning ontological and scientific myths would lead one to suspect.

The workshop where this all took place was held at King's College, Cambridge, in 250-year-old hand-cut stone buildings, surrounded by immaculate English gardens. The setting was ironically apposite. Looking out the window, I felt like an eighteenth-century explorer, newly returned from the New World, faced with the task of explaining, over tea and crumpets, the stark and searing beauty of the arctic, with its thousands of miles of desolate rock and wild rivers, unchecked ranges of lapped, craggy mountains, vast outcroppings of exfoliating granite, and millions upon millions of acres of profusely variegated and often impenetrable lichen, brambles, and muskeg. It is hard to imagine a more vivid contrast. And for better or worse—but mostly, I believe, for worse—the conception of "object" that has been enshrined in present-day science and analytic philosophy, with its presumptive underlying precision and clarity, is more reminiscent of fastidiously cropped hedge rows, carefully weeded rose gardens, and individually labeled decorative trees, than it is of the endless and rough arctic plain, or of a million-ton iceberg midwifed with a deafening crack and splintering spray from a grimy 10,000-year-old ice flow.

It is not wildness, per se, that differentiates the cases. No one travels the arctic without tools and equipment, endless preparation, and companions—or escapes an enmeshing web of social and political influences. But even the artifacts, on an expedition, are to a degree unstable, contingent, in flux. It is not just the dirt and grease ground into endlessly mended paddles—paddles grasped and worked by salty hands of people with sunburnt skin and densely matted hair. Think also of late-night conversations around the campfire, about possible routes and upcoming dangers (are we following "the same route" as Sir John Franklin?), about how to make one's way across fifty-seven miles of partially frozen swamp. Landscapes, plans, even material ordnance, are fecund masses of jumbled identity.

A far cry from a meticulous bed of yellow jonquils. Or so at least I thought. But then, while I watched, the Cambridge garden began to yield up its secrets, and the world fell back into its usual restless semblance of place. It was obvious, first, that the perfection with which I felt so at odds was in large part an idealization, if not illusory—that crumbly earth and erupting masses of new shoots are as essential to the life of a garden as cut stems or kempt grounds. Moreover, as I was mulling on the role of tamed wildness in English horticulture, a gardener came into view, and patiently bent down to work. Of course! That is what is profound about gardens: *they have to be maintained*. It takes continuous, sometimes violent, and often grueling, work, work that is in general nowhere near as neat as its intended result, to maintain gardens as the kind of thing that they are. Neither discovered, nor in any simple sense merely constructed, gardens, in order to be gardens, must be cared for, tended—even loved. What more could one ask for, by way of ontological moral?

This is a book about metaphysics—one that attempts to do justice to the tundra, to gardening, to politics, to rock. As indicated, my path into these subjects has come through computer science, but that is mostly by the way. Although some technical material is reviewed in chapter 1, computational considerations are largely set aside, in order to tell a tale about the territory into which that long exploration has led. The result is something of a metaphysical romp—occasionally riding rough-shod over turf already well explored (and well tilled) by generations of writers: from philosophy, feminism, theology, science studies, physics, poetry. Notwithstanding the germ of truth in the remark that "progress is made by stepping on the toes of giants," links with these literatures need to, and at some later point will, be forged. Nevertheless, my aim for the present text is simple: by presenting the story stripped of its computational heritage, to open up a conversation about perspectives, requirements, insights, and struggles—a conversation with others who have been led, via different routes, to this same metaphysical terrain.

To those inspired to take the trip—whether from explicit professional wrestling with such issues, or as the result of late-night reservations about how to participate authentically in academic life—I hope to say two things. First: *yes*, it is possible to base uncompromising theoretical inquiry on alternative foundations: messier foundations, contested foundations, foundations that run closer to the wry and weathered texture of ordinary life. No one, least of all God, has decreed that intellectual rigor must (or even can) be founded on a pristine foundational atomism. Second, though, I also want to make evident just how much such a transformation costs. Politics, creativity, ambiguity, irreverence—none of these can be grafted, at a later stage, onto a silent steel core, or even poured, like life-giving water, over inherently desiccated foundations. The whole story has to be turned upside down.

As is inevitable for a project under way for decades, I owe more debts to more people than possibly can, or should, be said here. I trust they know who they are, and why.

With respect to this particular manuscript, the initial suggestion to write it, and much subsequent encouragement, came from long-time friend and colleague Kathleen Akins, who orga-

nized a conference in Vancouver in January 1992 on "The Perception of Objects," as part of the Simon Fraser University Cognitive Science Series, where the material was first presented. Since then I have benefited greatly from the workshop participants at King's, as well as from audiences at the University of Illinois, Indiana University, Stanford University, and the Xerox Palo Alto Research Center. All these groups, it must be said, were subjected to pell-mell presentations of a view that cannot, it seems, be conveyed in an hour. I hope this book goes some way towards explaining what happened.

I am also extremely grateful to Adrian Cussins, and to the singular group of fellow travelers he assembled—Adam Lowe, Bruno Latour, and Charis Cussins—each of whom in his or her own way is telling a similar story, though Adam tends to do so in acrylic and epoxy. The first full draft of the manuscript was prepared for a meeting of this small group at ChatelPerron, an eleventh-century castle south of Paris—an idyllic setting for naturalized theology.

Closer to home, I have worked for many years in collaboration with colleagues at the Xerox Palo Alto Research Center: Mike Dixon, Gregor Kiczales, John Lamping, Jim des Rivières, and other members (past and present) of the Embedded Computation Area—and many other members of the wider PARC community, for whom this in part was written: Annette Adler, Bob Anderson, Bob Bauer, Danny Bobrow, Austin Henderson, Dan Huttenlocher, Gitte Jordan, Geoff Nunberg, Estee Solomon-Gray, Lucy Suchman, and Mark Weiser, to name just a few.

Special thanks also to: John Haugeland, co-presenter at the Vancouver conference and long-time metaphysical companion, and Adrian Cussins, again, for trenchant, detailed reviews of an earlier draft, without which the present version would be far more obscure; Teri Mendelsohn and Betty and Harry Bradford Stanton of the MIT Press, for faith in the project and stewardship

of the manuscript; Shelley Evenson and John Rheinfrank, for layout, typography, and the world of design, as well as for their help in designing this manuscript; Helga Wild, for all the artwork; and Sharon Olson, for editing. And to John Batali, Niklas Damiris, Güven Güzeldere, Stefan Heck, Mimi Ito, Natalie Jeremijenko, Joanna Rączaszek, Arnold Smith, Kären Weickert, and (again) Helga Wild, fellow travelers, sustaining friends, and stalwart coconspirators. Also to Charles Smith, for early faith; Barbara Grosz, for the original CSLI; Susan Newman, for help with the social; Walter Fontana, for chemistry and biology; Meg Graham, for reliance; Jon Barwise, for surety; Peter Finkelstein, for balance; Tytti Solantaus, for identity; Leigh Star, for bravery; Stefano Franchi, for play; and Jim Mahoney, who understands improvisation.

I also want to give this to David Levy, Marcia Lind, Mitch Marcus, Susan Stucky, and Henry and Catharine Thompson, separately and together, because of what it is. It is for my parents, too, Wilfred and Muriel Smith (to whom I have added little)—and above all, again, for Susan. Finally, a very warm thank you to John Seely Brown, for fifteen years of sustaining friendship—to say nothing of running the best place in the world to do research on the edge.

On the Origin of Objects

Introduction

This book introduces a new metaphysics—a philosophy of presence—that aims to steer a path between the Scylla of naive realism and the Charybdis of pure constructivism. The goal is to develop an integral account that retains the essential humility underlying each tradition: a form of epistemic deference to the world that underlies realism, and a respect for the constitutive human involvement in the world that underwrites constructivism.[1] The proposal is also intended to be neutral with respect to (a late twentieth-century update of) C P Snow's famous two cultures—serving equally as a foundation for the academic, intellectual, and technological, on the one hand, and for the curious, the moral, the erotic, the political, the artistic, and the sheerly obstreperous, on the other.

Fundamental to the view is a claim that objects, properties, practice, and politics—indeed everything ontological—live in what is called the "middle distance": an intermediate realm between a proximal though ultimately ineffable *connection*, reminiscent of the familiar physical bumping and shoving of the world, and a more remote *disconnection*, a form of unbridgeable separation that lies at the root of abstraction and of the partial (and painful) subject-object divide. No sense attends to the idea of complete connection or complete disconnection; limit idealizations are outmoded. Yet an essential interplay of patterns of partial connection and partial disconnection—restless figures

The footnotes in this book, positioned here in the bottom left corner of the page, are intended to be contrapuntal to the main body of the text. As well as supplying the usual references and other supporting material, they are used to fill in certain details, particularly philosophical details, that only some readers will view as essential to the argument.

[1] "Constructivism" is intended to signify not only what is known as *social constructionism*, but a more widespread constructive metaphysical stance especially prevalent in artificial intelligence and cognitive science.

of separation and engagement—is shown to underlie a single notion taken to unify representation and ontology: that of a subject's *registration* of the world.

The notion of metaphysics is itself overhauled in the process. For the project requires finding something whose existence (even possibility) is not obvious in advance: a way to feed our undiminished yearning for foundations and grounding, while at the same time avoiding the reductionism and ideological fundamentalism that have so bedeviled prior foundationalist approaches. Doing so (this is the same fact) answers the fundamental open challenge to pluralism, by making it evident how an integral metaphysical reality can simultaneously underwrite our unutterable differences and uniqueness, and at the same time sustain basic notions of worth and value, standards and significance, truth and even beauty. Thus the proposal shows what any successful metaphysics must show: how an irrevocable commitment to pluralism is compatible with the recognition that not all stories are equally good.[2]

Perhaps surprisingly, the end result is a picture of rather stark simplicity. Simplicity, though, is more a retrospective than a prospective quality—and anyway it is not a picture of a simple world, but a simple picture of a world of surpassing richness.

1 The foundations of computation

It may help, first, to explain how a computer scientist has come to tell such a bluntly metaphysical tale.

The interest in metaphysics and ontology has grown rather

[2]Thus Haraway: "So I think my problem and 'our' problem is how to have *simultaneously* an account of radical historical contingency for all knowledge claims and knowing subjects, a critical practice for recognizing our own 'semiotic technologies' for making meanings, *and* a no-nonsense commitment to faithful accounts of a 'real' world, one that can be partially shared and friendly to earth-wide projects of finite freedom, adequate material abundance, modest meaning in suffering, and limited happiness." Haraway (1991), p. 187, italics in the original.

straightforwardly out of a long-term (and ongoing) concern with the foundations of computer science and artificial intelligence. For more than twenty-five years I have been striving to develop an adequate and comprehensive theory of computation, one able to meet two essential criteria:

1. **Empirical:** It must do justice to computational practice (e.g., be capable of explaining Microsoft Word—including, for reasons that will emerge, the program itself, its construction, maintenance, and use);[3] and
2. **Conceptual:** It must provide a tenable foundation for the computational theory of mind—the thesis, sometimes known as "cognitivism,"[4] that underlies artificial intelligence and cognitive science.

The first, empirical, requirement, of doing justice to practice, helps to keep the analysis grounded in real-world examples, as a guard against excessive abstraction. It is humbling, too, since if anything is reliable about the computer revolution, it is its unerring ability to adapt, expand, dodge expectations, and generally outstrip our theoretical grasp. The criterion's primary advantage, however, is to provide a vantage point from which to question the legitimacy of all theoretical perspectives. For I take it as a tenet that what Silicon Valley *treats* as computational, in practice, *is* computational, in fact; to deny that would be considered sufficient grounds for rejection. But no such a priori com-

[3]To require that an account of computation *do justice* to a program's design, use, construction, repair, etc., is not to insist that a theory of computation *be* a theory of use (design, repair, etc.). For any technology or tool or equipment, there is almost certainly no limit to, and therefore no theory of, the uses to which it can be put. The idea, rather, is to make it a criterion on a successful theory of a technology that it explain or make intelligible how it is that the technology can subserve or play the roles that it does, while recognizing that they will be essentially open-ended. Thus if the materiality of computation is essential to a non-trivial fraction of the roles to which it is put, then an adequate theory, by these lights, will have to treat of its materiality.

[4]Cf. for example Haugeland (1978).

mitment is given to any story about computation, including the widely held Turing-theoretic conception of computability that currently goes by the name "the theory of computation." By the same token, I reject all proposals that assume that computation can be defined. By my lights, an adequate theory must make a substantive empirical claim about what I call *computation in the wild*:[5] that eruptive body of practices, techniques, networks, machines, and behavior that has so palpably revolutionized late-twentieth-century life.

The second condition, that an adequate theory of computation provide a tenable foundation for a theory of mind, is of a rather different character. Here labeled *conceptual*, it is more a metatheoretic requirement on the form or status of the theory than a constraint on its substantive content. In committing myself to honor the criterion, in particular, I make no advance commitment to cognitivism's being true (or false); I just want to know what it says.

That is not to say that the content of cognitivism is left open. Cognitivism's fundamental thesis—that the mind is computational—is given substance by the first, empirical criterion. The point is only that cognitivism, as I read it, is not a theory-laden claim, in the sense of framing specific hypotheses about what computers are. Rather, it has a more of an ostensive character: that people (i.e., us) are computers in whatever way that computers (i.e., those things over there) are computers, or at least in whatever way some of those things are computers.[6] In my view,

[5]With a nod to Hutchins' *Cognition in the Wild* (1995).

[6]To say that people, or minds, or intelligence, are computational is an asymmetrical claim: it is not, conversely, to say that to be computational is to be a person, or mind, or intelligent. Patently, some computers are impersonal, non-mental, and stupid. That raises a question about the status of the identifying characteristic of the subset. The strongest version of the computational claim, to which I believe most artificial intelligence advocates would subscribe, should be reconstructed along something like the following lines: (i) that people, minds, etc., are a subclass of the set of computers—implying that the property of being a computer is necessarily ex-

that is, cognitivism holds that people manifest, or exemplify, or are, or can be explained by, or can be illuminatingly understood in terms of, *whatever properties it is that characterize some identifiable species of the genus exemplified by computation-in-the-wild.* It would not be very interesting, or at least I would not consider it very interesting, for someone to argue that people are (or are not) "computers" according to a notion of computation that does not apply to present-day practice—does not apply to Unix, say, or to the PowerPC, or to the embedded controllers in a computerized lathe. Such a claim would be uninteresting because the hunch that has motivated artificial intelligence and propelled computational psychology into the mainstream of intellectual life—the hunch that has excited several generations of graduate students—is ultimately grounded in practice. Two weeks in any artificial intelligence laboratory is enough to demonstrate this. The cognitive revolution is fueled, both directly and indirectly, by an embodied and enthusiastically endorsed, but as-yet largely tacit, intuition based on many years of practical computational experience.

It follows that any theoretical formulation of cognitivism is doubly contingent. Thus consider Newell and Simon's popular "physical symbol system hypothesis," or Fodor's claim that thinking consists of formal symbol manipulation, or Dreyfus' assertion that cognitivism (as opposed to connectionism) re-

...

hibited by every mind; and (ii) that the property that distinguishes that subset—i.e., its characteristic function—is itself also a computational property (i.e., is computationally defined, occurs as part of a theory of computation, involves a computational regularity, whatever). Clearly, other variants are possible, such as that being a mind involves being computational *plus something else*, where that "something else" is noncomputational: having transducers of a certain form (on the assumption that 'transducer' is not a computational property), being conscious, or whatever. However this goes, a more finely textured characterization is needed than is captured for example in Searle's binary distinction between "strong" vs. "weak" AI (Searle 1980, 1984). See also Jackendoff (1987).

quires the explicit manipulation of explicit symbols.[7] Not only do these writers make a hypothetical statement about *people*, that they are physical, formal, or explicit symbol manipulators, respectively; they do so by making a hypothetical statement about *computers*, that they are in some essential or illuminating way characterizable in the same way. Because I take the latter claim to be as subservient to empirical adequacy as the former, there are two ways in which these writers could be wrong. In claiming that people are formal symbol manipulators, for example, Fodor would naturally be wrong if computers were formal symbol manipulators and people were not. But he would also be wrong, even though cognitivism itself might still be true, if computers were not formal symbol manipulators, either.

That, then, constitutes what I will call the *computational project*: to formulate a true and satisfying theory of computation that honors these two criteria. Needless to say, neither criterion is easy to meet. Elsewhere, I report on a study of half a dozen reigning construals of computation, with reference to both criteria—formal symbol manipulation, automata theory, information processing, digital state machines, recursion theory, Turing machines, the theory of effective computability, complexity theory, the assumptions underlying programming language semantics, and the like—and argue, in brief, that *each fails*, on *both counts*.[8] These are non-standard conclusions, and I do not press the present reader even to suspend reasonable disbelief, let alone to accept them. To understand the present book, however, and especially to understand why I use the examples that I do, it helps to know that they are, in fact, conclusions that I hold. To begin with, that is, I believe that each of the enumer-

[7]Newell & Simon (1976), Newell (1980), Fodor (1975, 1980), Dreyfus & Dreyfus (1988), and "Introduction to the MIT Press Edition" in Dreyfus (1992).

[8] *The Middle Distance* (forthcoming); henceforth TMD. The present volume can be seen as a travel brochure for a new land to which that study takes a full-fledged trip.

ated construals is inadequate to explain practice, being incapable of making sense of current systems, much less of unleashing a new generation. As a result, each fails the empirical criterion. For various theoretical reasons, moreover, even if they were empirically adequate, no one of them alone, nor any group in combination, would provide a tenable foundation for a theory of mind. So none of them meets the conceptual criterion, either.

What is the problem? Why do these theories all fail?

The most celebrated difficulties have to do with semantics. It is widely recognized that computation is in one way or another a symbolic or representational or information-based or semantical—i.e., as philosophers would say, an *intentional*—phenomenon.[9] Somehow or other, though in ways we do not yet understand, the states of a computer can model or simulate or represent or stand for or carry information about or signify other states in the world (or at least can be taken by people to do so—see the sidebar on original vs. derivative semantics, page 10). This semantical or intentional character of computation is betrayed by such phrases as *symbol* manipulation, *information* processing, programming *languages*, *knowledge representation*, *data* bases, etc. Furthermore, it is the intentionality of the computational that motivates the cognitivist thesis. The only compelling reason to suppose that we (or minds or intelligence)

[9]Although the term 'intentional' is philosophical, there are many philosophers, to say nothing of some computer and cognitive scientists, who would deny that computation is an intentional phenomenon. Reasons vary, but the most common go something like this: (i) that computation is both *syntactic* and *formal*, where 'formal' means "independent of semantics"; and (ii) that intentionality has fundamentally to do with semantics; and therefore (iii) that computation is thereby not intentional. I believe this is wrong, both empirically (that computation is purely syntactic) and conceptually (that being syntactic is a way of not being intentional); I also disagree that being intentional has *only* to do with semantics, which the denial requires. See the "Independent of semantics" sidebar on page 15, as well as TMD·II.

Original vs. derivative semantics

Many people have argued that the semantics of computational systems is intrinsically *derivative* or *attributed*—i.e., of the sort that books and signs have, in the sense of being ascribed by outside observers or users—as opposed to that of human thought and language, which in contrast is assumed to be *original* or *authentic.** I am dubious about the ultimate utility (and sharpness) of this distinction, and also about its applicability to computers—for a cluster of reasons, having to do partly with the fact that computers are increasingly being embedded in real systems, and can thus directly sense and affect what their symbols denote, and partly with arguments about the division of linguistic labor, or the division of labor more generally, arguments implying that some of our referential capacities are inherited from the communities of which we are a part.

Even if the distinction were useful, however, and even if the claim based on it were true—even if computational intentionality were in point of fact derivative—that would take nothing away from the point made in the text, that a theory of computation needs a theory of intentionality. First, it would be absurd to take derivativeness to be a *defining* characteristic of computation; if, in fact, derivativeness is a property of computers, that fact ought to be demonstrable—derivable from an account of what computers are like, to say nothing of an account of what originality and derivativeness are like. Such an argument would require a prior theory of intentionality. Second, derivative semantics is still real semantics—and of a quite complex kind. *Derivative*, after all, does not mean *fake* or *unreal.*

*Searle and Haugeland are the primary proponents of this distinction; see Searle (1980, 1982, 1984) and Haugeland (1981 p. 32–33; and 1985 p. 25–28, 87, & 119–123). For another very clear application to computers see Dretske (1985). An opposing (nondifferentiated) view is argued by Dennett; e.g., in chapter 8 of (1987).

might be computers stems from the fact that we, too, deal with representations, symbols, meaning, information, and the like.

For someone with cognitivist leanings, therefore—as opposed, say, to an eliminative materialist,[10] or to some types of connectionist—it is natural to expect that a comprehensive theory of computation will have to focus on its semantical aspects. This raises problems enough. Consider just the issue of representation. In order to meet the first criterion, of empirical adequacy, a successful candidate will have to make sense of the myriad kinds of representation that saturate practical systems—from bit maps and images to knowledge representations and data bases; from caches to backup tapes; from low-level finite-element models used in simulation to high-level analytic descriptions supporting reasoning and inference; from text to graphics to audio to video to virtual reality. As well as being vast in scope, it will also have to combine decisive theoretical bite with exquisite resolution, in order to distinguish models from implementations, analyses from simulations, and virtual machines at one level of abstraction from virtual machines at another level of abstraction in terms of which the former may be implemented.

[10]Strictly speaking, an eliminative materialist—or rather, in the vocabulary of chapter 5, an eliminative physicalist—would be someone who thinks that in due course (e.g., in millennial science) all non-physical levels of explanation, presumably including intentional, computational, and psychological varieties, will be "eliminated" in favor of physicalist or materialist alternatives. This strong form of eliminativism is further from common sense than reductionism: whereas the reductionist expects (or at least allows) higher-level notions, such as belief and truth, to survive but to be reducible to lower-level material accounts, the eliminativist in contrast expects the higher-level account to be replaced or discarded.

Not all self-proclaimed eliminativists are this thorough-going, though. It all depends on what is in line for elimination. Thus the Churchlands, well-known defenders of an eliminativist position, are more committed to the elimination of propositional attitude psychology than they are to the elimination of all intentional (semantic) vocabulary and/or ontology. See, e.g., Patricia Churchland (1986, 1992) and Paul Churchland (1989).

In order to meet the second, conceptual, criterion, moreover, any account of this profusion of representational practice must be grounded on, or at least defined in terms of, a theory of semantics or content. This is necessary in part in order that the concomitant psychological theory avoid vacuity or circularity, and in part in order that the main theory meet a minimal kind of naturalistic criterion: that we understand how computation is part of the natural world.[11] This is made all the more difficult by the fact that the word 'semantics' is used in an incredible variety of different senses across the range of the intentional sciences. Indeed, in my experience it is virtually impossible from any one location within that range to understand the full significance of the term, so disparate is that practice in toto.[12]

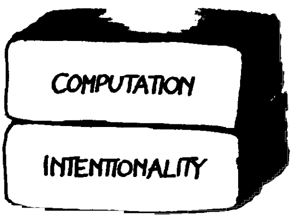


FIGURE 0·1 FIRST CUT

Genuine theories of content, moreover—of what it is that makes a given symbol or structure or patch of the world be *about*

[11] My use of 'naturalism' is non-standard: although it is a (metatheoretic) criterion I embrace, I do not thereby mean to lend support to physicalist explanation, or indeed to give any special status to what are known as the natural sciences. As explained in §1.a of chapter 3, all I am really committed to is that the account show how computation is not, as it were, *supernatural*: mysterious, spooky, or otherwise intellectually disreputable.

[12] In computer science, to take just one salient example, the term "the semantics of α," where α is an expression or construct in a programming language, means approximately the following: the topological (not geometrical) temporal profile of the behavior to which execution of this program fragment gives rise. By "topological" I mean that the overall temporal order of events is typically dictated, but that their absolute time-structure (e.g., exactly how fast the program runs) is not. As a result, a program can typically be sped up, either by adjusting the code, or by running it on a faster processor, without, as it is said, "changing the semantics." This reading is discussed further in chapter 1, §2. For an historical analysis of how the word 'semantics' has come to be used for such an effective kind of phenomenon—something that, at least to a traditional logician, sounds more like proof theory than model theory—see TMD·III.

or *oriented towards* some other entity or structure or patch—are notoriously hard to come by.[13] What is needed, to put a Husserlian spin on it, is an account of why, when we look out the window, we see a tree—i.e., have a (potentially) conscious experience of or about a tree, not of a two-dimensional leafy and barked surface, let alone of a pattern of incident electromagnetic radiation, or even have no experience at all, but instead just participate in an electromagnetically-mediated causal loop. Some putatively foundational construals of computation are implicitly defined in terms of just such a background theory of semantics, but do not explain what semantics is, and thus fail the second conceptual criterion. This group includes the "formal symbol manipulation" construal so favored in the cognitive sciences, in spite of its superficial formulation as being "independent of semantics" (see the sidebar on page 15). Other construals, such as those that view computation as the behavior of discrete automata—and also, I would argue, even if this is not immediately evident, the recursion-theoretic one that describes such behavior as the calculation of effective functions—fail to deal with computation's semantical aspect at all, in spite of sometimes using semantical vocabulary, and so fail the first empirical criterion. In the end, I find myself driven to the conclusion represented in figure 0·1, which for discussion I will call a "first cut" on the computational project: that, in spite of the advance press, especially from cognitivist quarters, computer science, far from supplying answers to the fundamental intentional mysteries, must, like cognitive science, await the development of a satisfying theory of intentionality.[14]

[13] Best known are Dretske's semantic theory of information (1981), which has more generally given rise to what is known as "indicator semantics," Fodor's "asymmetrical-dependence" theory (1987), and Millikan's "teleosemantics" or "biosemantics" (1984, 1989). For comparison among these alternatives see, e.g., Fodor (1984) and Millikan (1990).

[14] As suggested in footnote 9 (p. 9) and in the "Independent of semantics" sidebar on page 15, philosophers are less likely than computer scientists to expect a theory of computation to be, or to supply, a theory of

2 The ontological wall

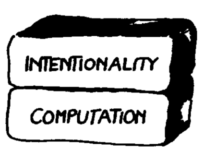
This, then, was my position for almost twenty years: (i) I was in awe of the depth, texture, scope, pluck, and impact of practice; (ii) I was critical of the inadequate state of current theory; and (iii) I was sure in my belief that what was needed, above all else, was a (situated, embodied, embedded, indexical, critical, reflexive, and all sorts of other things) theory of representation and semantics. In line with this metatheoretic attitude, I kept semantical and representational issues in primary theoretical focus.[15] Since, as already indicated, what currently goes by the name "the theory of computation"—essentially a derivative of recursion and complexity theory—pays very little attention to such intentional problems, to strike even this much of a semantical stance was to part company with the center of gravity of the received theoretical tradition.

And yet, in spite of the importance and magnitude of these intentional difficulties, I have gradually come to believe something even more sobering: that the most serious problems standing in the way of developing an adequate theory of computation are as much *ontological* as they are semantical. It is not that the semantic problems go away; they remain as challenging as ever. It is just that they are joined—on center stage, as it were—by even more demanding problems of ontology.

Except that to say "joined" is misleading, as if it were a matter of simple addition—i.e., as if now there were two problems on

...

intentionality. Thus they would not expect the metatheoretic structure to be as expected by most computer scientists and artificial intelligence researchers—namely, as indicated in the diagram to the right, with a theory of intentionality resting on a theory of computation. But that does not mean they would necessarily agree with figure 0·1. As discussed in TMD·II, many philosophers seem to think that a theory of computation can be *independent* of a theory of intentionality. Clearly, I do not believe this is correct.



[15]As is typical of those who work in artificial intelligence, I also paid more attention to representation than to algorithms.

Independent of semantics

Because formal symbol manipulation is usually defined as "manipulation of symbols independent of their interpretation," some people believe that the formal symbol manipulation construal of computation does not rest on a theory of semantics. But that is simply an elementary, though apparently very common, conceptual mistake.

The "independence of semantics" postulated as essential to the formal symbol construal is independence at the level of the phenomenon; it is a claim about how symbol manipulation *works*. Or so at least I have come to believe, based on more than twenty years of investigating what version of formality practitioners are actually committed to. (Whether the condition holds of computation-in-the-wild is a separate issue.) The intuition is simple enough: that semantic properties, such as referring to the Sphinx, or being true, are not of the right sort to do effective work. So it cannot be by exemplifying them that computers run. At issue in the present discussion, in contrast, is independence *at the level of the theory*—or, perhaps, to put it less epistemically and more ontologically, independence at the level of the *types*. And here the formal symbol manipulation construal is as dependent on semantics as it is possible to be: *it is defined in terms of it*. And defining yourself in terms of something is not a very good way to be independent of it, as the parent of any teenager knows.

Symbols must have a semantics—i.e., have an actual interpretation, be interpretable, whatever—in order for there to be something substantive for their formal manipulation to proceed independently of. Without a semantic character to be kept crucially in the wings, the formal symbol manipulation construal would collapse in vacuity. It would degenerate into something like "the manipulation of structure," or, as I put it in TMD·II, *stuff manipulation*—i.e., materialism.

the table, where before there had been just one. No such luck. The two issues are inextricably entangled—a fact of obstinate theoretical and metatheoretical consequence. This book can be viewed as an attempt to follow out, as simply but as rigorously as possible, the consequences of this entanglement.

A methodological consequence will illustrate the problem. Within the analytic tradition (by which I mean to include not just analytic philosophy, e.g. of language and mind, but most of modern science as well, complete with its formal/mathematical methods), it is traditional to analyze semantical or intentional systems, such as computers or people, under the following presupposition: (i) that one can parse or register the relevant theoretical situation in advance into a set of objects, properties, types, relations, equivalence classes, and the like (e.g., into people, heads, sentences, real-world referents, etc.), as if this were theoretically innocuous; and then (ii), with that ontological parse in hand, go on to proclaim this or that or the other thing as an empirically justified result. Thus for example one might describe a mail-delivering robot by first describing an environment of offices, hallways, people, staircases, litter, and the like, through which it is supposed to navigate; and then, taking this characterization of its context as given, ask how or whether the creature represents routes, say, or offices, or the location of mail delivery stations.

If one adopts a reflexively critical point of view, however, as I will try to do in chapter 1, one is inexorably led to the following conclusion: that, in that allegedly innocent pretheoretical "set-up" stage, one is liable, even if unwittingly, to project so many presuppositions, biases, blindnesses, and advance clues about the "answer," and in general to so thoroughly prefigure the target situation, without either apparent or genuine justification, that one cannot, or at least should not, take any of the subsequent "analysis" seriously. And that is problematic, in turn, not just because it rejects standard analyses, but because it seems to shut all inquiry down. What else can one do, after all? How can one not parse the situation in advance (since it will hardly do to merely whistle and walk away)? And if, undaunted, one were to go ahead and parse it anyway, what kind of story could possibly serve as a justification? It seems that any conceivable form of defense would devolve into another instance of the same problem.

In sum, the experience is less one of facing an ontological challenge than of running up against an ontological wall. Perhaps not quite of slamming into it, at least in my own case, since recognition dawned slowly. But neither is the encounter exactly gentle. It is difficult to exaggerate the sense of frustration that can come, once the conceptual fog begins to clear, from seeing one's theoretical progress blocked by what seems to be an insurmountable metaphysical obstacle.

Like the prior claim that all extant theories of computation are inadequate to reconstruct practice, this last claim, that theoretical progress is stymied for lack of an adequate theory of ontology, is a strong statement, in need of correspondingly strong defense. Again, this book does not present rigorous arguments for it, though I hope chapter 1 goes some way towards suggesting why it is true. I mention it only in order to motivate the present project. In my judgment, that is, despite the progress that has been made so far, we are not going to get to the heart of computation, representation, cognition, information, semantics, or intentionality, until the ontological wall is scaled, penetrated, dismantled, or in some other way defused.

3 A convergence of fields

One reaction to the wall might be depression. Fortunately, the prospects are not so bleak. For starters, there is some solace in company. It is perfectly evident, once one raises one's head from the specifically computational situation and looks around, that computer scientists, cognitive scientists, and artificial intelligence researchers are not the only ones running up against the demands of ontology. Similar conclusions are being reported from many other quarters. The words are different, and the perspectives complementary, but the underlying phenomena are the same.

The concerns are perhaps most pressing for literary critics, anthropologists, and other social theorists, vexed by what analytic categories to use in understanding people or cultures that, by the theorists' own admission, comprehend and constitute the world using concepts alien to the theorists' own. What makes the problem particularly obvious, in these cases, is the potential for *conceptual clash* between theorist's and subject's worldview[16]—a clash that can easily seem paralyzing. One's own categories are hard to justify, and reek of imperialism; it is at best presumptuous, and at worst impossible, to try to adopt the categories of one's subjects; and it is manifestly impossible to work with no concepts at all. So it is unclear how, or even whether, to proceed.

But conceptual clash, at least outright conceptual clash, is

[16]It is hardly surprising that among the first writers to recognize the potential of undischarged bias in their theoretical equipment were those, including many comparative humanists, social scientists, theologians, etc., whose task it was to characterize other people or cultures that lived by different conceptual schemes. But the fact that ontological problems are obvious in such cases does not mean that such difficulties are unique to them. For example, it is traditional within the humanities to suppose that univocal, objectivist, monist, or "scientific" approaches are fine for rocks, but inapplicable in human affairs. It is true, of course, that rocks' lack of a conceptual system obviates the possibility of conceptual clash—and rocks will not complain, either, being mute. It would take a lot more work than this, however, to show that methodological monism can do a rock justice.

Computers are not cultures, at least not yet, but neither are they rocks. As chapter 1 tries to demonstrate, problems analogous to conceptual clash arise for any intentional subject matter, no matter how modest.

not the only form in which the ontological problem presents. Consider the burgeoning interest in "complex systems" coalescing in a somewhat renegade subdiscipline at the intersection of dynamics, theoretical biology, and artificial life. This community debates the "emergence of organization," the units that selection operates on, the structure of self-organizing systems, the smoothness or roughness of what are called "fitness landscapes," and the like.[17] In spite of being virtually constitutive of the discipline, these discussions are conducted in the absence of adequate theories of what organization comes to, of what a "unit" consists in, of how "entities" arise (as opposed to how they survive), of how it is determined what predicates should figure in characterizing a fitness landscape as rough or smooth, etc. The ontological lack is to some extent recognized in increasingly vocal calls for "theories of organization" (a theory of organization is nothing but a metaphysics with a business plan). But the calls have not yet been answered.[18]

Ontological problems have also plagued physics for years, at least since foundational issues of interpretation were thrown into relief by the developments of relativity and quantum me-

[17] See for example Kaufmann (1993), Gell-Mann (1994), and Gleick (1984).

[18] A dramatic historical example involves the development of graphics displays. It is relatively easy to program cellular automata with various kinds of rule, and to see—palpably, in front of one's own eyes—little "organisms" and other organized or patterned entities emerge—dynamically, like worms or clusters or hives of activity. But when one "sees" such a creature emerge, one is relying on one's perceptual apparatus. No one yet has a theory that, given a field of cellular activity, can reliably identify such "objects" or "emergent entities." And *identification* is not really the problem, anyway. If the underlying theory—of selection, say, or organization, or behavioral emergence, or evolution—is to be worth its salt, it should be defined in terms of a theory of such things. (It has even been speculated that the entire field of non-linear dynamics, popularly called "chaos theory," could not have happened without the development of such displays. No one would have "seen" the patterns in textual lists of numbers.)

chanics (including the perplexing wave-particle duality, and the distinction between "classical" and "quantum" world views[19]). They face connectionist psychologists, who, proud of having developed architectures that do not rely on the manipulation of formal symbol structures encoding high-level concepts, and thus of having thereby rejected propositional content, are nevertheless at a loss as to say what their architectures *do* represent.[20] And then there are communities that tackle ontological questions directly—not just philosophy, but poetry and art, where attempts to get in or around or under objects and ontology have been pursued for centuries.

So there are fellow travelers (perhaps even fellow readers). In this book, however, at least for one more chapter, I will remain focused on the case with which I started: computation in the wild. For the plan for the book is to respond to the encounter with the ontological wall by developing an actual, concrete, metaphysical proposal—not simply to press for properties or features that an adequate metaphysics must have. Unlike at least some arguments for realism or irrealism, and also unlike some treatises pro or con this or that philosophy of science, the present project is not intended to be *meta*-metaphysical. But answering the concrete demand requires detailed, representative examples—examples, as one says, that "go the distance." And for this purpose the computational realm has unparalleled advantage. Midway between matter and mind, computation stands in excellent stead as a supply of concrete cases of middling complexity—what in computer science is called a "validation suite"—against which to test specific metaphysical hypotheses. "Middling" in the sense of being neither so simple as to invite caricature, nor so complex as to defy comprehension. Crucially, too, they are cases with which we are as much practically as theoreti-

[19] See the sidebar on page 146.

[20] See e.g. Smolensky (1988) and Cussins (1990). Though not, strictly speaking, a connectionist, Brooks' (1991) proposal to set aside representation entirely can be viewed as an extreme version of this view.

cally familiar (we build systems better than we understand them[21]).

The end result is intended to be general, however—i.e., to cope as well with other fields' ontological problems as with computation's—so I have tried to present the examples in chapter 1 in a straightforward enough way that readers from other disciplines can understand what is going on. As far as possible, technical details have been restricted to a series of sidebars. As a result, the text should be accessible to anyone even moderately familiar with conceptual issues in computation, artificial intelligence, or cognitive science. Still, in places chapter 1 is rough going. If computers are not your cup of tea, do not fret the details. Better to identify situations in your own experience that exhibit analogous (or disanalogous) structure than to try to figure these out, which will anyway not work if they are not "in the bones" familiar. Computation is *an* example, throughout; it is never *the* example.

At the beginning of chapter 2, with motivation and examples in place, I set aside any further consideration of computational issues per se, and for the remainder of the book go straight for the metaphysical jugular. The project from that point forward remains as described at the outset: to develop a new metaphysical/ontological account—an account adequate, at a minimum, though that will turn out to be much less of a minimum than anyone would have had any right to expect, to the computational task.

4 Rhetoric

One final comment, on writing. It turns out that the normative demands on writing in computer science and philosophy are

[21] In the spirit of Vico's "verum et factum convertuntur" (1744/1984), Dretske (1994) argues that to understand a mechanism is to know how to build one. I am sympathetic to a necessity reading of the claim, expressed for example in Dretske's title: "If You Can't Make One, You Don't Know How It Works." But present computational practice, in my view, defies a sufficiency reading.

rather different. In computer science, a text is treated as something like a map or travel guide: it gives you a route to the author's real contribution, but once that contribution is reached (understood, implemented), the text has served its purpose, and might as well be thrown away. It is assumed that the destination can be evaluated on autonomous criteria—of elegance, utility, practicality, etc. (how fast it runs, how easy it is to use, how many people will buy it). Part of what it is to be a professional computer scientist is to know how to perform this kind of on-location property assessment. It would seem as bizarre for serious normative criteria to be applied to the guiding text as for the capabilities of a sports car to be confused with the merits of its owner's manual, or the value of a buried treasure to be confused with the merits of the path through the jungle one was forced to take to reach it.

The situation in philosophy is a little different. Here the text, or perhaps the argument that the text expresses, is subject to more serious normative evaluation. At its worst, this can sometimes make it seem as if conclusions are held to no higher (e.g., ethical) standard than the quality of the technical arguments that have been offered in their favor. But even at its best there remains a substantial difference between the two fields—a difference reflected in the fact that texts, at least arguably, are the primary product in philosophy, but of secondary importance in computer science.

This book may look like philosophy, but do not be fooled. Although in many ways it takes the form of an argument, if you scratch its surface it will betray its computational roots. It is only as much of an argument as I thought was necessary in order to show the reader the direction I was headed. I was less interested, this time around, in developing watertight arguments than in introducing a new territory—a territory that I believe is worth exploring on its own merits.[22] The particular path I follow is by no means the only one to that territory; nor do I have any reason to suppose that it is necessarily the best. It is one of the routes I have traveled, however; I hope it is one that others find worthwhile as well.

[22]The TMD series presents a more rigorous and detailed argument.

References

Page numbers in brackets at the end of each reference indicate places in the text where the given work is cited. Footnotes are explicitly indicated (e.g., 284 n) only in cases where the author is not cited by name in the main text. 245 s refers to the sidebar on page 245.

Agre, Philip E. (1996): *Computation and Human Experience*, Cambridge: Cambridge University Press · [140 n]

Akins, Kathleen (1993): "What It Is Like to Be Boring and Myopic," in Bo Dahlbom, ed., *Dennett and His Critics*, Oxford: Blackwell, pp. 124–160 · [68 n]

Aydede, Murat (1993): *Syntax, Functionalism, Connectionism, and the Language of Thought*, doctoral dissertation, College Park: Univ. of Maryland · [35 n. 13]

Barrow, John (1991): *Theories of Everything*, Oxford: Clarendon · [144 n. 6]

Barwise, Jon and John Perry (1983): *Situations and Attitudes*, Cambridge, Mass.: MIT Press / A Bradford Book · [31 n. 8, 284 n]

Bawden, Alan (1993): *Linear Graph Reduction: Confronting the Cost of Naming*, doctoral dissertation in computer science, Cambridge, Mass.: Massachusetts Institute of Technology · [41 n]

Bechtel, William and Adele A. Abrahamsen (1991): *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*, Oxford: Basil Blackwell · [64 n. 37]

Beeson, Michael (1985): *Foundations of Constructive Mathematics*, Berlin: Springer-Verlag · [355 n. 5]

Bellah, Robert (1985): *Habits of the Heart*, Berkeley: University of California Press · [91 n]

——— (1991): *The Good Society*, New York: Knopf (distributed by Random House) · [91 n]

Berger, Peter L. and Thomas Luckmann (1966/1980): *The Social Construction of Reality: a Treatise in the Sociology of Knowledge*, Garden City, New York: Doubleday (1966); New York: Irvington Publishers (1980) · [86 n. 1]

Bloor, David (1976/1991): *Knowledge and Social Imagery*, Chicago: University of Chicago Press · [86 n. 1]

Boyd, Richard (1979): "Metaphor and Theory Change: What Is 'Metaphor' a Metaphor For?" in Anthony Ortoni, ed., *Metaphor and Thought*, Cambridge: Cambridge University Press, pp. 356–419 · [33 n. 9]
Brentano, Franz Clemens (1874/1973): *Psychology from an Empirical Standpoint*, New York: Humanities Press · [111, 209]
Brooks, Rodney A. (1986): "A Robust Layered Control System for a Mobile Robot," IEEE *Journal of Robotics and Automation*, Vol. 2, pp. 14–23 · [50]
——— (1991): "Intelligence Without Representation," *Artificial Intelligence*, Vol. 47, Issues 1–3, pp. 139–159 · [20 n. 20, 351 n.]
Butler, Sir William Francis (1982), as quoted in the *Wild Rivers Survey* series of pamphlets, Parks Canada, Ottawa, Canada · [vii]
Callon, Michel (1986): "Some Elements of a Sociology of Translation: Domestication of the Scallops and Fishermen of St. Brieuc Bay," in J. Law, ed., *Power, Action, and Belief: A New Sociology of Knowledge?* London: Routledge and Kegan Paul · [86 n. 1]
Castañeda, Hector-Neri (1975): *Thinking and Doing: The Philosophical Foundations of Institutions*, Dordrecht, Holland: Reidel · [248 n]
Chapman, David and Philip Agre (1987): "Abstract Reasoning as Emergent from Concrete Activity," in Michael P. Georgeff and Amy L. Lansky, eds., *Reasoning about Action and Plans: Proceedings of the 1986 Workshop*, Los Altos, California: Morgan Kaufmann, pp. 411–424 · [248 n]
——— (1990): "What Are Plans For?" in Pattie Maes, ed., *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, Cambridge, Mass.: MIT / Elsevier, pp. 17–34 · [351 n]
Chisholm, Roderick M. (1957): *Perceiving: A Philosophical Study*. Ithaca, New York: Cornell University Press · [111 n]
Churchland, Patricia Smith (1986): *Neurophilosophy*, Cambridge, Mass.: MIT Press / A Bradford Book · [11 n, 281]
——— (1992): *The Computational Brain*, Cambridge, Mass.: MIT Press / A Bradford Book · [11 n, 281]
Churchland, Paul M. (1989): *A Neurocomputational Perspective*, Cambridge, Mass.: MIT Press / A Bradford Book · [11 n, 281]
——— (1995): *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge, Mass.: MIT Press / A Bradford Book · [46 n. 21]
Constable, Robert L., et al. (1986): *Implementing Mathematics with the Nuprl Development System*. New Jersey: Prentice-Hall · [30 n. 5]
Crick, Francis and Christof Koch (1992): "Towards a Neurobiological Theory of Consciousness," *Seminars in Neuroscience*, Vol. 2 · [145 n]
Cussins, Adrian (1990): "The Connectionist Construction of Concepts,"

in Margaret Boden, ed., *The Philosophy of Artificial Intelligence*, Oxford: Oxford University Press, pp. 368–440 · [20 n. 20, 46 n. 21, 65 n. 39, 281]

——— (1992): "Content, Embodiment, and Objectivity: The Theory of Cognitive Trails," *Mind*, Vol. **101**, No. 404, Oct. 1992, pp. 651–688 · [281]

——— (forthcoming): *Constructions of Thought* · [42 n, 50 n. 25, 140 n. 4, 281]

Davidson, Donald (1974): "On the Very Idea of a Conceptual Scheme," *Proceedings and Addresses of the American Philosophical Association, Vol.* 47, reprinted in Donald Davidson, *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press (1984), pp. 183–198 · [197 n]

Dennett, Daniel C. (1982): "Beyond Belief," in Andrew Woodfield, ed., *Thought and Object: Essays on Intentionality*, Oxford: Clarendon Press · [192 n]

——— (1987): *The Intentional Stance*, Cambridge, Mass.: MIT Press / A Bradford Book · [10 s, 73, 258 n]

——— (1991): *Consciousness Explained*, Boston: Little, Brown and Co. · [231 s]

Derrida, Jacques (1967): *Writing and Difference* (*L'écriture et la différence*), Paris: Editions du Seuil.

Dixon, Michael (1991): *Embedded Computation and the Semantics of Programs*, Xerox Palo Alto Research Center Report SSL–91–1, 182 pp. · [35 n. 11, 77 n. 1, 245 s, 246 n, 256 n. 11]

Dretske, Fred I. (1981): *Knowledge and the Flow of Information*, Cambridge, Mass.: MIT Press / A Bradford Book · [13 n. 13, 31 n. 8]

——— (1985): "Machines and the Mental," Presidential Address to the 83rd Annual Meeting of the Western Division of the American Philosophical Association, Chicago: *APA Proceedings* 59, pp. 23–33 · [10 s]

——— (1994): "If You Can't Make One, You Don't Know How It Works," in Peter French, Theodore Uehling, et al., eds., *Midwest Studies in Philosophy*, Notre Dame, Indiana: University of Notre Dame Press, pp. 468–482 · [21 n, 147]

Dreyfus, Hubert L. (1992): *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, Mass.: MIT Press · [7–8, 281]

Dreyfus, Hubert L. and Stuart E. Dreyfus (1988): "Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint," *Dædelus* special issue on artificial intelligence, Winter 1988; issued as Vol. **117** No. 1 of the *Proceedings of the American Academy of Arts and Sciences*, pp. 15–43 · [7–8]

Dumas, Alexandre (1894): *The Count of Monte Cristo*, Boston: Little, Brown & Co. · [196 n. 6]

Dupré, John (1993): *The Disorder of Things*, Cambridge, Mass.: Harvard University Press · [326 n. 1, 334 n]

Dupuy, Jean-Pierre (1994): *Aux origines des sciences cognitives*, Paris: La Découverte · [111 n]

Evans, Gareth (1982): *The Varieties of Reference*, Oxford: Oxford University Press · [65, 125 n, 279, 281 n]

——— (1985): *Collected Papers*, Oxford: Oxford University Press.

Fausto-Sterling, Anne (1993): "The Five Sexes: Why Male and Female Are Not Enough," *The Sciences*, Vol. 33 No. 2 (March-April), pp. 20–25 · [79]

Fauvel, John, Raymond Flood, Michael Shortland, and Robin Wilson, eds. (1988): *Let Newton Be!* Oxford: Oxford University Press · [361 n]

Fodor, Jerry A. (1975): *The Language of Thought*, Cambridge, Mass.: Harvard University Press · [7–8, 350 n]

——— (1980): "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," *Behavioral and Brain Sciences*, Vol. 3 No. 1; March 1980, pp. 63–73. Reprinted in Fodor, Jerry: *Representations*, Cambridge, Mass.: MIT Press, 1981 · [7–8, 192 n, 350 n]

——— (1984): "Semantics, Wisconsin Style," *Synthese*, Vol. 59, pp. 231–250 · [13 n. 13]

——— (1987): *Psychosemantics*, Cambridge, Mass.: MIT Press · [13 n. 13, 31 n. 6]

Fodor, Jerry A. and Brian McLaughlin (1990): "Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work," *Cognition*, Vol. 35, pp. 183–196 · [35 n. 13, 279 n. 2]

Fodor, Jerry A. and Zenon Pylyshyn (1988): "Connectionism and Cognitive Architecture: A Critical Analysis," in Steve Pinker and Jacques Mehler, eds., *Connections and Symbols*, Cambridge, Mass.: MIT Press (a *Cognition* special issue), pp. 3–71 · [35 n. 13, 279 n. 2]

Franchi, Stefano (forthcoming · a): "Intentionally True," unpublished ms. · [111 n]

——— (forthcoming · b): *Endgames: Game and Play at the End of Philosophy* (working title), doctoral dissertation in philosophy, Stanford: Stanford University · [208 n]

Frege, Gottlob (1892/1993): "On Sense and Reference," reprinted in A. W. Moore, ed., *Meaning and Reference*, Oxford: Oxford University Press, pp. 23–42 · [43 n, 243 n. 1, 256]

——— (1893/1980): *The Foundations of Arithmetic : A Logico-Mathematical Enquiry into the Concept of Number*, second revised edition, Evanston, Illinois: Northwestern University Press · [131 n]

Gell-Mann, Murray (1994): *The Quark and the Jaguar*, New York: W. H. Freeman and Co. · [19 n. 17]

Gibson, James J. (1979): *The Ecological Approach to Visual Perception*, Boston: Houghton Mifflin · [217, 225 n, 351]

Gleick, James (1987): *Chaos*, New York: Viking · [19 n. 17]

Goel, Vinod (1995): *Sketches of Thought*, Cambridge, Mass: MIT Press / A Bradford Book · [46 n. 22, 254 n. 9, 281 n]

Goodman, Nelson (1976): *Languages of Art: An Approach to a Theory of Symbols*, second edition (first published 1968), Indianapolis: Hackett Publishing Co. · [357]

——— (1983): *Fact, Fiction, and Forecast*, fourth edition (first published 1954), Cambridge, Mass.: Harvard University Press · [31 n. 6, 284 n]

Grice, H. Paul (1989): *Studies in the Way of Words*, Cambridge, Mass.: Harvard University Press.

Halpern, Joe (1987): "Using Reasoning about Knowledge to Analyze Distributed Systems," *Annual Review of Computer Science*, Vol. 2, pp. 37–68 · [31 n. 8]

Haraway, Donna (1991): *Simians, Cyborgs, and Women: The Reinvention of Nature*, New York: Routledge—especially the article "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective" · [4 n, 86 n. 1, 208 n, 254 n. 8, 306 n, 340 n]

Harding, Sandra (1986): *The Science Question in Feminism*, Ithaca, New York: Cornell University Press · [86 n. 1, 93 n. 11]

——— (1991): *Whose Science? Whose Knowledge: Thinking from Women's Lives*, London: Open University Press · [86 n. 1]

Hardy, Godfrey H. (1929): "Mathematical Proof," *Mind*, January 1929, pp. 1–25 · [357]

Haugeland, John (1978): "The Nature and Plausibility of Cognitivism," *Behavioral and Brain Sciences*, Vol. 2, pp. 215–260 · [5 n. 4]

——— (1981): "Semantic Engines," in John Haugeland, ed., *Mind Design: Philosophy, Psychology, Artificial Intelligence*, Cambridge, Mass.: MIT Press / A Bradford Book (1981), pp. 1–34 · [10 s]

——— (1982): "Analog and Analog," *Philosophical Topics* (Spring 1981); reprinted in J. I. Biro & Robert W. Shahan, eds., *Mind, Brain, and Function: Essays in the Philosophy of Mind*, Norman, Oklahoma: University of Oklahoma Press (1982), pp. 213–225 · [177 n, 235]

——— (1985): *Artificial Intelligence: The Very Idea*, Cambridge, Mass.: MIT Press / A Bradford Book · [10 s]

——— (1991): "Representational Genera," in William Ramsey, Stephen P. Stich, and David E. Rumelhart, eds., *Philosophy and Connectionist Theory*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 61–90 · [46 n. 21]

Hutchins, Edwin (1995): *Cognition in the Wild*, Cambridge, Mass.: MIT Press · [6 n. 5]

Jackendoff, Ray S. (1987): *Consciousness and the Computational Mind*, Cambridge, Mass.: MIT Press · [7 n]

Kafka, Franz (1925/1992): *The Trial*, New York: Knopf · [326 n. 2]

Kauffman, Stuart A. (1993): *The Origins of Order*, New York: Oxford University Press · [19 n. 17]

Keller, Evelyn Fox (1985): *Reflections on Gender and Science*, New Haven: Yale University Press · [86 n. 1]

Keynes, John Maynard (1947): "Newton, the Man," *Newton Tercentenary Celebrations*, Cambridge: Cambridge University Press · [361 n]

Kirsh, David (1987): "Putting a Price on Cognition," in *Southern Journal of Philosophy*, Vol. XXVI, supplement.

——— (1991): "Today the Earwig, Tomorrow Man?" in *Artificial Intelligence*, Vol. 47 No. 1–3, pp. 161–184. Entire issue reprinted as David Kirsh, ed., *Foundations of Artificial Intelligence*, Cambridge, Mass.: MIT Press / A Bradford Book · [46 n. 21, 65]

Kitcher, Philip (1992): "The Naturalist's Return," *Philosophical Review*, Vol. 101, pp. 53–114 · [138 n. 1, 140 n. 4]

Knorr-Cetina, Karen (1981): *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford and New York: Pergamon Press · [86 n. 1]

Kripke, Saul (1976): "Is There a Problem about Substitutional Quantification?" in Gareth Evans and John McDowell, eds., *Truth and Meaning: Essays in Semantics*, Oxford: Clarendon Press, pp. 325–419 · [45 n]

Kuhn, Thomas (1962/1970): *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press (1 st edition 1967, 2 nd edition 1970) · [86 n. 1]

Kundera, Milan (1984): *The Unbearable Lightness of Being*, New York: Harper & Row · [208]

Latour, Bruno (1987): *Science in Action: How to Follow Scientists and Engineers Through Society*, Cambridge, Mass.: Harvard University Press · [79 n. 4, 86 n. 1, 335 n. 9]

——— (1988): *The Pasteurization of France*, Cambridge, Mass.: Harvard University Press · [77, 86 n. 1, 293]

——— (1993): *We Have Never Been Modern*, Wheatsheaf, New York: Harvester Press · [87 n. 3]

Latour, Bruno and Steven Woolgar (1979/1986): *Laboratory Life: The Social Construction of Scientific Fact*, London: Sage (1979). Reprinted (without the word 'social' in the title) Princeton: Princeton University Press (1986) · [86 n. 1]

Lave, Jean and Etienne Wenger (1991): *Situated Learning: Legitimate Peripheral Participation*, Cambridge and New York: Cambridge University Press · [195 n. 4, 289 n]

Leith, Philip (1990): *Formalism in AI and Computer Science*, New York: Ellis Horwood · [357 n. 9]

Lettvin, Jerome Y., Humberto R. Maturana, Warren S. McCulloch, and Walter H. Pitts (1959): "What the Frog's Eye Tells the Frog's Brain," *Proceedings of the IRE*, Vol. 47 No. 11 (November), pp. 1940–1959. Reprinted in Warren S. McCulloch, *Embodiments of Mind*, Cambridge, Mass.: MIT Press (1965/1988), pp. 230–255 · [149 n]

Levy, David M., Daniel C. Brotsky and Kenneth R. Olson (1988): "Formalizing the Figural: Aspects of a Foundation for Document Manipulation," *Proceedings of the ACM Conference on Document Processing Systems*, New York: Association for Computing Machinery · [48 n]

Livingston, Eric (1986): *The Ethnomethodological Foundations of Mathematics*, London: Routledge & Kegan Paul · [357 n. 9]

Locke, John (1959): *An Essay Concerning Human Understanding*, Dover Publications, unabridged, 2 volumes.

Longino, Helen (1990): *Science as Social Knowledge*, Princeton: Princeton University Press · [86 n. 1]

McAlister, Linda L. (1974): "Chisholm and Brentano on Intentionality." *Review of Metaphysics*, Vol. XXVIII No. 2, pp. 328–338 · [111 n]

McGinn, Colin (1982): "The Structure of Content," in Andrew Woodfield, ed., *Thought and Object: Essays on Intentionality*, Oxford: Clarendon Press · [192 n]

Marcus, Mitchell P. (1980): *A Theory of Syntactic Recognition for Natural Language*, Cambridge, Mass.: MIT Press · [61 n]

Marcus, Mitchell, Don Hindle, and Margaret Fleck (1983): "D-Theory: Talking about Talking about Trees," *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics* · [61 n]

Marcus, Mitchell and Don Hindle (1990): "Description Theory and Intonation Boundaries," in G. Altmann, ed., *Computational and Cognitive Models of Speech*, Cambridge, Mass.: MIT Press · [61 n]

Millikan, Ruth (1984): *Language, Thought, and Other Biological Categories*, Cambridge, Mass.: MIT Press · [13 n. 13]

——— (1989): "Biosemantics," *Journal of Philosophy*, Vol. LXXXVI No. 6, June 1989, pp. 281–297 · [13 n. 13]

——— (1990): "Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics," *Philosophical Topics*, Vol. **18** No. 2, Fall 1990, pp. 151–161 · [13 n. 13]

Monk, Ray (1990): *Ludwig Wittgenstein: The Duty of Genius*, New York: Penguin Books · [357 n. 8]

Montague, Richard (1963): "Syntactical Treatments of Modality, with Corollaries on Reflexion Principles and Finite Axiomatizability," *Acta Philosophica Fennica*, Vol. **16**, pp. 153–167 · [45 n]

Nagel, Thomas (1974): "What Is It Like to Be a Bat?" *Philosophical Review*, Vol. LXXXIII No. 4 (October), pp. 435–50 · [67 n. 40]

——— (1986): *The View From Nowhere*, Oxford: Oxford University Press · [305 n. 17]

Neisser, Ulric (1989): "Direct Perception and Recognition as Distinct Perceptual Systems," in Meeting of the Cognitive Science Society, Ann Arbor, Michigan: August 1989 · [351 n]

——— (1994): "Multiple Systems: A New Approach to Cognitive Theory," *European Journal of Cognitive Psychology*, Vol. **6**, pp. 225–241.

Newell, Allen (1980): "Physical Symbol Systems," *Cognitive Science*, Vol. **4**, pp. 135–183 · [7–8, 74]

Newell, Allen and Herbert A. Simon (1976): "Computer Science as Empirical Inquiry: Symbols and Search," *Communications of the Association for Computing Machinery*, Vol. **19** (March 1976), No. 3, pp. 113–126. Reprinted in John Haugeland, ed., *Mind Design*, Cambridge, Mass.: MIT Press / A Bradford Book (1981), pp. 35–66 · [7–8, 74]

Nunberg, Geoffrey (1993): "Indexicality and Deixis," *Linguistics and Philosophy*, Vol. **156**, pp. 1–43 · [169 n. 22]

Perry, John (1979): "The Problem of the Essential Indexical," *Nous*, Vol. **13** · [170, 248 n]

Pettit, Philip and John McDowell, eds. (1986): *Subject, Thought and Context*, Oxford: Oxford University Press · [192 n]

Popper, Karl R. and Eccles, John C. (1977): *The Self and Its Brain*, Berlin: Springer-Verlag · [158–59 s]

Putnam, Hilary (1981): *Reason, Truth, and History*, Cambridge: Cambridge University Press · [86]

——— (1987): *The Many Faces of Realism*, La Salle, Illinois: Open Court · [86, 112–13 s]

———, edited by James Conant (1990): *Realism with a Human Face*, Cambridge, Mass.: Harvard University Press · [86, 96, 112–13 s, 142]

Quine, Willard von Orman (1953): "Three Grades of Modal Involvement," *Proceedings of the XIth International Congress of Philosophy*, Vol. **14**, pp. 65–81 · [45 n]

——— (1964): *Word and Object*, Cambridge, Mass.: MIT Press · [138 n. 1]

——— (1971): "The Inscrutability of Reference," in D. D. Steinberg and L. A. Jacobovits, eds., *Semantics*, Cambridge: Cambridge University Press · [169 n. 22]

Reichenbach, Hans (1947): *Elements of Symbolic Logic*, New York: Free Press · [157, 167]

Rosenschein, Stanley (1985): "Formal Theories of Knowledge in AI and Robotics," *New Generation Computing*, Vol. **3** · [31 n. 8]

Rudwick, Martin J. S. (1985): *The Great Devonian Controversy: The Shap-*

ing of Scientific Knowledge among Gentlemanly Specialists, Chicago: University of Chicago Press · [86 n. 1]
Searle, John (1980): "Minds, Brains, and Programs," *Behavioral and Brain Sciences*, Vol. 3, No. 3, Sept. 1980, pp. 417–458 · [7 n, 10 s, 28, 30 n. 4]
——— (1982): "The Myth of the Computer: An Exchange," *New York Review of Books* · [10 s]
——— (1983): *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press · [252 n]
——— (1984): *Minds, Brains, and Science*, Cambridge, Mass.: Harvard University Press · [7 n, 10 s, 28, 145 n]
——— (1992): *The Rediscovery of the Mind*, Cambridge, Mass.: MIT Press / A Bradford Book · [29 n. 2, 31 n. 6, 43, 145 n, 196 n. 8, 292]
Shapin, Steven and Steven Schaffer (1985): *Leviathan and the Air Pump: Hobbs, Boyle, and the Experimental Life*, Princeton: Princeton University Press · [86 n. 1]
Smith, Brian Cantwell (1984): "Reflection and Semantics in Lisp," Salt Lake City: *Proceedings of the 11th Annual Principles of Programming Languages Conference (POPL)*, pp. 23–35 · [37 n. 16, 77 n. 1, 102 n]
——— (1986): "Varieties of Self-Reference," in Joseph Y. Halpern, ed., *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference*, Monterey, California, March 19–22, 1986, Los Altos, California: Morgan Kaufmann, pp. 19–43. Also available as Center for the Study of Language and Information Technical Report No. CSLI–87–76, Stanford: Stanford University · [77 n. 1, 248 n]
——— (1987): "The Correspondence Continuum," Center for the Study of Language and Information Technical Report No. CSLI–87–71, Stanford: Stanford University, 59 pp. · [34 n, 58 n. 31, 77 n. 1, 251 n, 262 n]
——— (1988): "The Semantics of Clocks," in James H. Fetzer, ed., *Aspects of Artificial Intelligence*, The Netherlands: Kluwer Academic Publishers, 1988, pp. 3–31. Also available as Center for the Study of Language and Information Technical Report No. CSLI–87–75, Stanford: Stanford University, March 1987 · [74 n, 261 n]
——— (1994): "Discreteness Run Amok," in Joan Fujimura and Lucy Suchman, eds., *Vital Signs: Cultural Perspectives on Coding Life and Vitalizing Code*, Department of Anthropology, Program in History and Philosophy of Science, Stanford University, June 2–4, 1994 · [203 n]
——— (forthcoming): *The Middle Distance: An Essay on the Foundations of Computation and Intentionality*, a series of books in the philosophy of computation · [8 n. 8, 23 n] Tentative titles:
Volume I: Introduction · [72 n, 186 n, 290 n. 10]
Volume II: Formal Symbol Manipulation · [9 n, 14 n. 14, 15 s, 365 s]
Volume III: Effective Computability · [12 n. 12, 210 s]

Volume IV: The Correspondence Continuum · [34 n, 42 n, 58, 77 n. 1, 186 n, 251 n]
Volume V: Digital State Machines · [235 n. 9]
Smolensky, Paul (1988): "On the Proper Treatment of Connectionism," *Behavioral and Brain Sciences*, Vol. 11, pp. 1–23 · [20 n. 20, 46 n. 21, 64 n. 37, 65 n. 39, 281]
——— (1991): "Connectionism, Constituency, and the Language of Thought," in Barry Loewer and Georges Rey, eds., *Meaning in Mind: Fodor and His Critics*, Oxford: Basil Blackwell, pp. 210–237 · [35 n. 13]
Snow, Charles Percy (1964): *The Two Cultures and the Scientific Revolution* and *A Second Look: An Expanded Version of 'The two cultures and the scientific revolution,'* Cambridge: Cambridge University Press · [3, 94]
Star, Leigh (1989): *Regions of the Mind*, Stanford: Stanford University Press · [305 n. 16]
Strawson, Peter (1959): *Individuals*, London: Methuen · [65 n. 39, 124, 175, 232, 279 n. 2, 281 n]
——— (1992): *Analysis and Metaphysics*, Oxford: Oxford University Press.
Struik, Dirk (1962): *A Concise History of Mathematics*, London: G. Bell · [355 n. 5]
Suchman, Lucy (1987): *Plans and Situated Actions*, Cambridge: Cambridge University Press · [310]
——— (1994): "Working Relations of Technology Production and Use," *Computer Supported Cooperative Work*, Vol. 2, pp. 21–39 · [306 n, 340 n]
——— (in press): "Centers of Coordination: A Case and Some Themes," in L. Resnik, R. Saljo, and C. Pontecorvo, eds., *Discourse, Tools, and Reasoning*, Berlin: Springer-Verlag · [335 n. 7]
Taylor, Charles (1989): *Sources of the Self: The Making of the Modern Identity*, Cambridge, Mass.: Harvard University Press.
Vico, Giambattista (1744/1984): *The New Science of Giambattista Vico*, Ithaca, New York: Cornell University Press · [21 n]
Weinberg, Steven (1993): *Dreams of a Final Theory*, New York: Pantheon Books · [144 n. 6]
Wenger, Etienne (forthcoming): *Communities of Practice: Learning, Meanings, and Identity* · [36 n. 14]
Wild, Helga and Niklas Damiris (in prep.): *The Writing Condition: Bodies, Language, and the Emergence of Social Praxis*, Stanford: Stanford University Press · [363 n]

Designed by Shelley Evenson and Brian Cantwell Smith.
Typeset in Adobe Garamond text with Frutiger Ultra Black display
by Melissa L. Ehn at Wilsted & Taylor Publishing Services.