PRAISE FOR *DARWIN'S DANGEROUS IDEA*

"Passionate. . . . Mr. Dennett likens Darwin's theory to a 'universal acid' that eats through our most cherished assumptions."

—David Papineau, *The New York Times Book Review*

"It is . . . a pleasure to meet a philosopher who understands what Darwinism is about, and approves of it."

—John Maynard Smith, *The New York Review of Books*

"What you get, in short, is a jam-session view of man's fate, which in Dennett's hands turns out to seem no less moving than the well-scored religious ideas it has inexorably replaced."

—Adam Gopnik, *The New Yorker*

"Perhaps the most delicious meditation ever upon natural selection."

—Stewart Brand, author of *How Buildings Learn*

"Dennett's book brings together science and philosophy with wit, complex clarity, and an infectious sense that these ideas matter to us and the way we live now."

—A. S. Byatt, *The Sunday Times,* London

"If we had to choose one man to represent Earth in a debating tournament with extraterrestrials, the American philosopher Daniel Dennett would be a prime candidate."

—Matt Ridley, *The Times,* London

"A discussion of Darwin and contemporary evolutionary biological theory that is exemplary in its presentation of ideas crucial to our understanding of ourselves."

—Joyce Carol Oates, *Times Literary Supplement,* London

"The best nonfiction book I read [in 1995]...an ambitiously conceived attempt to chart the continuing intellectual revolution embodied in the theory of natural selection."

—Pat Barker, *The Sunday Times,* London

"A challenging, provocative read, especially the sections on ethics, language, and meaning.... Above all, the final chapters are a great rallying call in favor of diversity, of tolerance and accommodation, of honesty. The mood is one almost of spirituality, but there is an underlying tone of realism not all of which is easy medicine."

—Mary Mulvihill, *The Irish Times*

"An intellectual challenge full of devastating criticism and original insight, and it is also fun to read."

—*The Economist*

"A grand, provocative, gripping synthesis."

—*Publishers Weekly*

"Whether you conclude by the time you reach the end of this massive volume that Dennett makes his case or not, you have to admit that he is up to the challenge. Few mortals would be."

—Roger Lewin, *Los Angeles Times*

"Dennett's dangerous idea: to use his gift for lucid explanation and his twinkling wit to cure the strange allergy to Darwin in modern intellectual life. It is essential—and pleasurable—reading for any thinking person."

—Steven Pinker, author of *The Language Instinct*

"With characteristic insight, verve, and wit, Dan Dennett ranges over central problems in philosophy, evolutionary biology, and cognitive science. Readers of *Darwin's Dangerous Idea* will not always be convinced, but they will be informed, entertained, and provoked by its arguments. Dennett's big themes are Life and Mind, and he is a wonderful advertisement for the life of the mind."

—Philip Kitcher, author of *The Lives to Come*

"If you want an exciting, wide-ranging romp through great ideas, read this book. Daniel Dennett proves fully worthy of Charles Darwin."

—Jared M. Diamond, author of *The Third Chimpanzee*

"In this clear and rigorous testing of Darwinian theory across modern science, Dennett persuades us that evolution by natural selection is vital to the future of philosophy."

—Edward O. Wilson, author of *On Human Nature*

Also by Daniel C. Dennett

Consciousness Explained
The Intentional Stance
Elbow Room
The Mind's I (with Douglas R. Hofstadter)
Brainstorms
Content and Consciousness

DARWIN'S DANGEROUS IDEA

EVOLUTION AND THE MEANINGS OF LIFE

Daniel C. Dennett

SIMON & SCHUSTER PAPERBACKS
New York Toronto London Sydney

TO VAN QUINE
teacher and friend

SIMON & SCHUSTER PAPERBACKS
Rockefeller Center
1230 Avenue of the Americas
New York, NY 10020

Designed by Irving Perkins Associates
Manufactured in the United States of America
23 25 27 29 30 28 26 24

The Library of Congress has cataloged the hardcover edition as follows:
Dennett, Daniel Clement.
Darwin's dangerous idea : evolution and the meanings of life /
Daniel C. Dennett.
p. cm.
Includes bibliographical references and index.
1. Natural selection. 2. Evolution (Biology)—Philosophy.
3. Human evolution—Philosophy. I. Title.
QH375.D45 1995
146'.7—dc20 94-49158
CIP

ISBN-13: 978-0-684-80290-9
ISBN-10: 0-684-80290-2
ISBN-13: 978-0-684-82471-0 (Pbk)
ISBN-10: 0-684-82471-X (Pbk)

The author is grateful for permission to reprint the following copyrighted material:

Figure 7.1: © 1991 Sidney Harris. Reprinted with permission. All rights reserved.

(*Continued at back of book*)

Contents

Preface

Darwin's theory of evolution by natural selection has always fascinated me, but over the years I have found a surprising variety of thinkers who cannot conceal their discomfort with his great idea, ranging from nagging skepticism to outright hostility. I have found not just lay people and religious thinkers, but secular philosophers, psychologists, physicists, and even biologists who would prefer, it seems, that Darwin were wrong. This book is about why Darwin's idea is so powerful, and why it promises—not threatens—to put our most cherished visions of life on a new foundation.

A few words about method. This book is largely about science but is not itself a work of science. Science is not done by quoting authorities, however eloquent and eminent, and then evaluating their arguments. Scientists do, however, quite properly persist in holding forth, in popular and not-so-popular books and essays, putting forward their interpretations of the work in the lab and the field, and trying to influence their fellow scientists. When I quote them, rhetoric and all, I am doing what they are doing: engaging in persuasion. There is no such thing as a sound Argument from Authority, but authorities can be persuasive, sometimes rightly and sometimes wrongly. I try to sort this all out, and I myself do not understand all the science that is relevant to the theories I discuss, but, then, neither do the scientists (with perhaps a few polymath exceptions). Interdisciplinary work has its risks. I have gone into the details of the various scientific issues far enough, I hope, to let the uninformed reader see just what the issues are, and why I put the interpretation on them that I do, and I have provided plenty of references.

Names with dates refer to full references given in the bibliography at the back of the book. Instead of providing a glossary of the technical terms used, I define them briefly when I first use them, and then often clarify their meaning in later discussion, so there is a very extensive index, which will let you survey all occurrences of any term or idea in the book. Footnotes are for digressions that some but not all readers will appreciate or require.

One thing I have tried to do in this book is to make it possible for you to read the scientific literature I cite, by providing a unified vision of the field, along with suggestions about the importance or non-importance of the controversies that rage. Some of the disputes I boldly adjudicate, and others I leave wide open but place in a framework so that you can see what the issues are, and whether it matters—to you—how they come out. I hope you will read this literature, for it is packed with wonderful ideas. Some of the books I cite are among the most difficult books I have ever read. I think of the books by Stuart Kauffman and Roger Penrose, for instance, but they are pedagogical *tours de force* of highly advanced materials, and they can and should be read by anyone who wants to have an informed opinion about the important issues they raise. Others are less demanding—clear, informative, well worth some serious effort—and still others are not just easy to read but a great delight—superb examples of Art in the service of Science. Since you are reading this book, you have probably already read several of them, so my grouping them together here will be recommendation enough: the books by Graham Cairns-Smith, Bill Calvin, Richard Dawkins, Jared Diamond, Manfred Eigen, Steve Gould, John Maynard Smith, Steve Pinker, Mark Ridley, and Matt Ridley. No area of science has been better served by its writers than evolutionary theory.

Highly technical philosophical arguments of the sort many philosophers favor are absent here. That is because I have a prior problem to deal with. I have learned that arguments, no matter how watertight, often fall on deaf ears. I am myself the author of arguments that I consider rigorous and unanswerable but that are often not so much rebutted or even dismissed as simply ignored. I am not complaining about injustice—we all must ignore arguments, and no doubt we all ignore arguments that history will tell us we should have taken seriously. Rather, I want to play a more direct role in changing what is ignorable by whom. I want to get thinkers in other disciplines to take evolutionary thinking seriously, to show them how they have been underestimating it, and to show them why they have been listening to the wrong sirens. For this, I have to use more artful methods. I have to tell a story. You don't want to be swayed by a story? Well, I *know* you won't be swayed by a formal argument; you won't even *listen* to a formal argument for my conclusion, so I start where I have to start.

The story I tell is mostly new, but it also pulls together bits and pieces from a wide assortment of analyses I've written over the last twenty-five years, directed at various controversies and quandaries. Some of these pieces are incorporated into the book almost whole, with improvements, and others are only alluded to. What I have made visible here is enough of the tip of the iceberg, I hope, to inform and even persuade the newcomer and at least challenge my opponents fairly and crisply. I have tried to navigate between the Scylla of glib dismissal and the Charybdis of grindingly detailed infighting, and whenever I glide swiftly by a controversy, I warn that I am doing so, and give the reader references to the opposition. The bibliography could easily have been doubled, but I have chosen on the principle that any serious reader needs only one or two entry points into the literature and can find the rest from there.

In the front of his marvelous new book, *Metaphysical Myths, Mathematical Practices: The Ontology and Epistemology of the Exact Sciences* (Cambridge: Cambridge University Press, 1994), my colleague Jody Azzouni thanks "the philosophy department at Tufts University for providing a near-perfect environment in which to *do* philosophy." I want to second both the thanks and the evaluation. At many universities, philosophy is studied but not done—"philosophy appreciation," one might call it—and at many other universities, philosophical research is an arcane activity conducted out of sight of the undergraduates and all but the most advanced postgraduates. At Tufts, we *do* philosophy, in the classroom and among our colleagues, and the results, I think, show that Azzouni's assessment is correct. Tufts has provided me with excellent students and colleagues, and an ideal setting in which to work with them. In recent years I have taught an undergraduate seminar on Darwin and philosophy, in which most of the ideas in this book were hammered out. The penultimate draft was probed, criticized, and polished by a particularly strong seminar of graduate and undergraduate students, for whose help I am grateful: Karen Bailey, Pascal Buckley, John Cabral, Brian Cavoto, Tim Chambers, Shiraz Cupala, Jennifer Fox, Angela Giles, Patrick Hawley, Dien Ho, Matthew Kessler, Chris Lerner, Kristin McGuire, Michael Ridge, John Roberts, Lee Rosenberg, Stacey Schmidt, Rhett Smith, Laura Spiliatakou, and Scott Tanona. The seminar was also enriched by frequent visitors: Marcel Kinsbourne, Bo Dahlbom, David Haig, Cynthia Schossberger, Jeff McConnell, David Stipp. I also want to thank my colleagues, especially Hugo Bedau, George Smith, and Stephen White, for a variety of valuable suggestions. And I must especially thank Alicia Smith, the secretary at the Center for Cognitive Studies, whose virtuoso performance as a reference-finder, fact-checker, permission-seeker, draft-updater/printer/mailer, and general coordinator of the whole project put wings on my heels.

I have also benefited from detailed comments from those who read most or all the penultimate-draft chapters: Bo Dahlbom, Richard Dawkins, David Haig, Doug Hofstadter, Nick Humphrey, Ray Jackendoff, Philip Kitcher, Justin Leiber, Ernst Mayr, Jeff McConnell, Steve Pinker, Sue Stafford, and Kim Sterelny. As usual, they are not responsible for any errors they failed to dissuade me from. (And if you can't write a good book about evolution with the help of this sterling group of editors, you should give up!)

Many others answered crucial questions, and clarified my thinking in dozens of conversations: Ron Amundsen, Robert Axelrod, Jonathan Bennett, Robert Brandon, Madeline Caviness, Tim Clutton-Brock, Leda Cosmides, Helena Cronin, Arthur Danto, Mark De Voto, Marc Feldman, Murray Gell-Mann, Peter Godfrey-Smith, Steve Gould, Danny Hillis, John Holland, Alastair Houston, David Hoy, Bredo Johnsen, Stu Kauffman, Chris Langton, Dick Lewontin, John Maynard Smith, Jim Moore, Roger Penrose, Joanne Phillips, Robert Richards, Mark and Matt (the Ridley conspecifics), Dick Schacht, Jeff Schank, Elliot Sober, John Tooby, Robert Trivers, Peter Van Inwagen, George Williams, David Sloan Wilson, Edward O. Wilson, and Bill Wimsatt.

I want to thank my agent, John Brockman, for steering this big project past many shoals, and helping me see ways of making it a better book. Thanks also go to Terry Zaroff, whose expert copyediting caught many slips and inconsistencies, and clarified and unified the expression of many points. And Ilavenil Subbiah, who drew the figures, except for Figures 10.3 and 10.4, which were created by Mark McConnell on a Hewlett-Packard Apollo workstation, using I-dea.

Last and most important: thanks and love to my wife, Susan, for her advice, love, and support.

DANIEL DENNETT
September 1994

PART I

STARTING IN THE MIDDLE

Neurath has likened science to a boat which, if we are to rebuild it, we must rebuild plank by plank while staying afloat in it. The philosopher and the scientist are in the same boat. . . .

Analyze theory-building how we will, we all must start in the middle. Our conceptual firsts are middle-sized, middle-distanced objects, and our introduction to them and to everything comes midway in the cultural evolution of the race. In assimilating this cultural fare we are little more aware of a distinction between report and invention, substance and style, cues and conceptualization, than we are of a distinction between the proteins and the carbohydrates of our material intake. Retrospectively we may distinguish the components of theory-building, as we distinguish the proteins and carbohydrates while subsisting on them.

—WILLARD VAN ORMAN QUINE 1960, pp. 4–6

CHAPTER ONE

Tell Me Why

1. Is Nothing Sacred?

We used to sing a lot when I was a child, around the campfire at summer camp, at school and Sunday school, or gathered around the piano at home. One of my favorite songs was "Tell Me Why." (For those whose personal memories don't already embrace this little treasure, the music is provided in the appendix. The simple melody and easy harmony line are surprisingly beautiful.)

Tell me why the stars do shine,
Tell me why the ivy twines,
Tell me why the sky's so blue.
Then I will tell you just why I love you.

Because God made the stars to shine,
Because God made the ivy twine,
Because God made the sky so blue.
Because God made you, that's why I love you.

This straightforward, sentimental declaration still brings a lump to my throat—so sweet, so innocent, so reassuring a vision of life!

And then along comes Darwin and spoils the picnic. Or does he? That is the topic of this book. From the moment of the publication of *Origin of Species* in 1859, Charles Darwin's fundamental idea has inspired intense reactions ranging from ferocious condemnation to ecstatic allegiance, sometimes tantamount to religious zeal. Darwin's theory has been abused and misrepresented by friend and foe alike. It has been misappropriated to lend scientific respectability to appalling political and social doctrines. It has been pilloried in caricature by opponents, some of whom would have it compete in our children's schools with "creation science," a pathetic hodge-podge of pious pseudo-science.[1]

Almost no one is indifferent to Darwin, and no one should be. The Darwinian theory is a scientific theory, and a great one, but that is not all it is. The creationists who oppose it so bitterly are right about one thing: Darwin's dangerous idea cuts much deeper into the fabric of our most fundamental beliefs than many of its sophisticated apologists have yet admitted, even to themselves.

The sweet, simple vision of the song, taken literally, is one that most of us have outgrown, however fondly we may recall it. The kindly God who lovingly fashioned each and every one of us (all creatures great and small) and sprinkled the sky with shining stars for our delight—*that* God is, like Santa Claus, a myth of childhood, not anything a sane, undeluded adult could literally believe in. *That* God must either be turned into a symbol for something less concrete or abandoned altogether.

Not all scientists and philosophers are atheists, and many who are believers declare that their idea of God can live in peaceful coexistence with, or even find support from, the Darwinian framework of ideas. Theirs is not an anthropomorphic Handicrafter God, but still a God worthy of worship in their eyes, capable of giving consolation and meaning to their lives. Others ground their highest concerns in entirely secular philosophies, views of the meaning of life that stave off despair without the aid of any concept of a Supreme Being—other than the Universe itself. Something *is* sacred to these thinkers, but they do not call it God; they call it, perhaps, Life, or Love, or Goodness, or Intelligence, or Beauty, or Humanity. What both groups share, in spite of the differences in their deepest creeds, is a conviction that life does have meaning, that goodness matters.

But can *any* version of this attitude of wonder and purpose be sustained in the face of Darwinism? From the outset, there have been those who thought they saw Darwin letting the worst possible cat out of the bag: nihilism. They thought that if Darwin was right, the implication would be that nothing could be sacred. To put it bluntly, nothing could have any point. Is this just an overreaction? What exactly are the implications of Darwin's idea—and, in any case, has it been scientifically proven or is it still "just a theory"?

Perhaps, you may think, we could make a useful division: there are the parts of Darwin's idea that really are established beyond any reasonable doubt, and then there are the speculative extensions of the scientifically

1. I will not devote any space in this book to cataloguing the deep flaws in creationism, or supporting my peremptory condemnation of it. I take that job to have been admirably done by Kitcher 1982, Futuyma 1983, Gilkey 1985, and others.

irresistible parts. Then—if we were lucky—perhaps the rock-solid scientific facts would have no stunning implications about religion, or human nature, or the meaning of life, while the parts of Darwin's idea that get people all upset could be put into quarantine as highly controversial extensions of, or mere interpretations of, the scientifically irresistible parts. That would be reassuring.

But alas, that is just about backwards. There are vigorous controversies swirling around in evolutionary theory, but those who feel threatened by Darwinism should not take heart from this fact. Most—if not quite all—of the controversies concern issues that are "just science"; no matter which side wins, the outcome will not undo the basic Darwinian idea. That idea, which is about as secure as any in science, really does have far-reaching implications for our vision of what the meaning of life is or could be.

In 1543, Copernicus proposed that the Earth was not the center of the universe but in fact revolved around the Sun. It took over a century for the idea to sink in, a gradual and actually rather painless transformation. (The religious reformer Philipp Melanchthon, a collaborator of Martin Luther, opined that "some Christian prince" should suppress this madman, but aside from a few such salvos, the world was not particularly shaken by Copernicus himself.) The Copernican Revolution did eventually have its own "shot heard round the world": Galileo's *Dialogue Concerning the Two Chief World Systems*, but it was not published until 1632, when the issue was no longer controversial among scientists. Galileo's projectile provoked an infamous response by the Roman Catholic Church, setting up a shock wave whose reverberations are only now dying out. But in spite of the drama of that epic confrontation, the idea that our planet is not the center of creation has sat rather lightly in people's minds. Every schoolchild today accepts this as the matter of fact it is, without tears or terror.

In due course, the Darwinian Revolution will come to occupy a similarly secure and untroubled place in the minds—and hearts—of every educated person on the globe, but today, more than a century after Darwin's death, we still have not come to terms with its mind-boggling implications. Unlike the Copernican Revolution, which did not engage widespread public attention until the scientific details had been largely sorted out, the Darwinian Revolution has had anxious lay spectators and cheerleaders taking sides from the outset, tugging at the sleeves of the participants and encouraging grandstanding. The scientists themselves have been moved by the same hopes and fears, so it is not surprising that the relatively narrow conflicts among theorists have often been not just blown up out of proportion by their adherents, but seriously distorted in the process. Everybody has seen, dimly, that a lot is at stake.

Moreover, although Darwin's own articulation of his theory was monumental, and its powers were immediately recognized by many of the scientists and other thinkers of his day, there really were large gaps in his theory that have only recently begun to be properly filled in. The biggest gap looks almost comical in retrospect. In all his brilliant musings, Darwin never hit upon the central concept, without which the theory of evolution is hopeless: the concept of a *gene*. Darwin had no proper *unit* of heredity, and so his account of the process of natural selection was plagued with entirely reasonable doubts about whether it would work. Darwin supposed that offspring would always exhibit a sort of blend or average of their parents' features. Wouldn't such "blending inheritance" always simply average out all differences, turning everything into uniform gray? How could diversity survive such relentless averaging? Darwin recognized the seriousness of this challenge, and neither he nor his many ardent supporters succeeded in responding with a description of a convincing and well-documented mechanism of heredity that could combine traits of parents while maintaining an underlying and unchanged identity. The idea they needed was right at hand, uncovered ("formulated" would be too strong) by the monk Gregor Mendel and published in a relatively obscure Austrian journal in 1865, but, in the best-savored irony in the history of science, it lay there unnoticed until its importance was appreciated (at first dimly) around 1900. Its triumphant establishment at the heart of the "Modern Synthesis" (in effect, the synthesis of Mendel and Darwin) was eventually made secure in the 1940s, thanks to the work of Theodosius Dobzhansky, Julian Huxley, Ernst Mayr, and others. It has taken another half-century to iron out most of the wrinkles of that new fabric.

The fundamental core of contemporary Darwinism, the theory of DNA-based reproduction and evolution, is now beyond dispute among scientists. It demonstrates its power every day, contributing crucially to the explanation of planet-sized facts of geology and meteorology, through middle-sized facts of ecology and agronomy, down to the latest microscopic facts of genetic engineering. It unifies all of biology and the history of our planet into a single grand story. Like Gulliver tied down in Lilliput, it is unbudgeable, not because of some one or two huge chains of argument that might—hope against hope—have weak links in them, but because it is securely tied by hundreds of thousands of threads of evidence anchoring it to virtually every other area of human knowledge. New discoveries may conceivably lead to dramatic, even "revolutionary" *shifts* in the Darwinian theory, but the hope that it will be "refuted" by some shattering breakthrough is about as reasonable as the hope that we will return to a geocentric vision and discard Copernicus.

Still, the theory is embroiled in remarkably hot-tempered controversy, and one of the reasons for this incandescence is that these debates about scientific matters are usually distorted by fears that the "wrong" answer would have intolerable moral implications. So great are these fears that they are carefully left unarticulated, displaced from attention by several layers of distracting rebuttal and counter-rebuttal. The disputants are forever changing the subject slightly, conveniently keeping the bogeys in the shadows. It is this misdirection that is mainly responsible for postponing the day when we can all live as comfortably with our new biological perspective as we do with the astronomical perspective Copernicus gave us.

Whenever Darwinism is the topic, the temperature rises, because more is at stake than just the empirical facts about how life on Earth evolved, or the correct logic of the theory that accounts for those facts. One of the precious things that is at stake is a vision of what it means to ask, and answer, the question "Why?" Darwin's new perspective turns several traditional assumptions upside down, undermining our standard ideas about what ought to count as satisfying answers to this ancient and inescapable question. Here science and philosophy get completely intertwined. Scientists sometimes deceive themselves into thinking that philosophical ideas are only, at best, decorations or parasitic commentaries on the hard, objective triumphs of science, and that they themselves are immune to the confusions that philosophers devote their lives to dissolving. But there is no such thing as philosophy-free science; there is only science whose philosophical baggage is taken on board without examination.

The Darwinian Revolution is both a scientific and a philosophical revolution, and neither revolution could have occurred without the other. As we shall see, it was the philosophical prejudices of the scientists, more than their lack of scientific evidence, that prevented them from seeing how the theory could actually work, but those philosophical prejudices that had to be overthrown were too deeply entrenched to be dislodged by mere philosophical brilliance. It took an irresistible parade of hard-won scientific facts to force thinkers to take seriously the weird new outlook that Darwin proposed. Those who are still ill-acquainted with that beautiful procession can be forgiven their continued allegiance to the pre-Darwinian ideas. And the battle is not yet over; even among the scientists, there are pockets of resistance.

Let me lay my cards on the table. If I were to give an award for the single best idea anyone has ever had, I'd give it to Darwin, ahead of Newton and Einstein and everyone else. In a single stroke, the idea of evolution by natural selection unifies the realm of life, meaning, and purpose with the realm of space and time, cause and effect, mechanism and physical law. But it is not just a wonderful scientific idea. It is a dangerous idea. My admiration for Darwin's magnificent idea is unbounded, but I, too, cherish many of the ideas and ideals that it *seems* to challenge, and want to protect them. For instance, I want to protect the campfire song, and what is beautiful and true in it, for my little grandson and his friends, and for their children when they grow up. There are many more magnificent ideas that are also jeopardized, it seems, by Darwin's idea, and they, too, may need protection. The only good way to do this—the only way that has a chance in the long run—is to cut through the smokescreens and look at the idea as unflinchingly, as dispassionately, as possible.

On this occasion, we are not going to settle for "There, there, it will all come out all right." Our examination will take a certain amount of nerve. Feelings may get hurt. Writers on evolution usually steer clear of this apparent clash between science and religion. Fools rush in, Alexander Pope said, where angels fear to tread. Do you want to follow me? Don't you really want to know what survives this confrontation? What if it turns out that the sweet vision—or a better one—survives intact, strengthened and deepened by the encounter? Wouldn't it be a shame to forgo the opportunity for a strengthened, renewed creed, settling instead for a fragile, sickbed faith that you mistakenly supposed must not be disturbed?

There is no future in a sacred myth. Why not? Because of our curiosity. Because, as the song reminds us, *we want to know why*. We may have outgrown the song's answer, but we will never outgrow the question. Whatever we hold precious, we cannot protect it from our curiosity, because being who we are, one of the things we deem precious is the truth. Our love of truth is surely a central element in the meaning we find in our lives. In any case, the idea that we might preserve meaning by kidding ourselves is a more pessimistic, more nihilistic idea than I for one can stomach. If that were the best that could be done, I would conclude that nothing mattered after all.

This book, then, is for those who agree that the only meaning of life worth caring about is one that can withstand our best efforts to examine it. Others are advised to close the book now and tiptoe away.

For those who stay, here is the plan. Part I of the book locates the Darwinian Revolution in the larger scheme of things, showing how it can transform the world-view of those who know its details. This first chapter sets out the background of philosophical ideas that dominated our thought before Darwin. Chapter 2 introduces Darwin's central idea in a somewhat new guise, as the idea of evolution as an *algorithmic process*, and clears up some common misunderstandings of it. Chapter 3 shows how this idea overturns the tradition encountered in chapter 1. Chapters 4 and 5 explore some of the striking—and unsettling—perspectives that the Darwinian way of thinking opens up.

Part II examines the challenges to Darwin's idea—to neo-Darwinism or the Modern Synthesis—that have arisen within biology itself, showing that contrary to what some of its opponents have declared, Darwin's idea survives these controversies not just intact but strengthened. Part III then shows what happens when the same thinking is extended to the species we care about most: *Homo sapiens*. Darwin himself fully recognized that this was going to be the sticking point for many people, and he did what he could to break the news gently. More than a century later, there are still those who want to dig a moat separating us from most if not all of the dreadful implications they think they see in Darwinism. Part III shows that this is an error of both fact and strategy; not only does Darwin's dangerous idea apply to us directly and at many levels, but the proper application of Darwinian thinking to human issues—of mind, language, knowledge, and ethics, for instance—illuminates them in ways that have always eluded the traditional approaches, recasting ancient problems and pointing to their solution. Finally, we can assess the bargain we get when we trade in pre-Darwinian for Darwinian thinking, identifying both its uses and abuses, and showing how what really matters to us—and ought to matter to us—shines through, transformed but enhanced by its passage through the Darwinian Revolution.

2. What, Where, When, Why—and How?

Our curiosity about things takes different forms, as Aristotle noted at the dawn of human science. His pioneering effort to classify them still makes a lot of sense. He identified four basic questions we might want answered about anything, and called their answers the four *aitia*, a truly untranslatable Greek term traditionally but awkwardly translated the four "causes."

(1) We may be curious about what something is made of, its matter or *material cause*.
(2) We may be curious about the form (or structure or shape) that that matter takes, its *formal cause*.
(3) We may be curious about its beginning, how it got started, or its *efficient cause*.
(4) We may be curious about its *purpose* or *goal* or *end* (as in "Do the ends justify the means?"), which Aristotle called its *telos*, sometimes translated in English, awkwardly, as "final cause."

It takes some pinching and shoving to make these four Aristotelian *aitia* line up as the answers to the standard English questions "what, where, when, and why." The fit is only fitfully good. Questions beginning with "why," however, do standardly ask for Aristotle's fourth "cause," the *telos* of a thing. Why this? we ask. What is it *for*? As the French say, what is its *raison d'être*, or reason for being? For hundreds of years, these "why" questions have been recognized as problematic by philosophers and scientists, so distinct that the topic they raise deserves a name: teleology.

A *teleological* explanation is one that explains the existence or occurrence of something by citing a goal or purpose that is served by the thing. Artifacts are the most obvious cases; the goal or purpose of an artifact is the function it was designed to serve by its creator. There is no controversy about the *telos* of a hammer: it is for hammering in and pulling out nails. The *telos* of more complicated artifacts, such as camcorders or tow trucks or CT scanners, is if anything more obvious. But even in simple cases, a problem can be seen to loom in the background:

"Why are you sawing that board?"
"To make a door."
"And what is the door for?"
"To secure my house."
"And why do you want a secure house?"
"So I can sleep nights."
"And why do you want to sleep nights?"
"Go run along and stop asking such silly questions."

This exchange reveals one of the troubles with teleology: where does it all stop? What *final* final cause can be cited to bring this hierarchy of reasons to a close? Aristotle had an answer: God, the Prime Mover, the *for-which* to end all *for-whiches*. The idea, which is taken up by the Christian, Jewish, and Islamic traditions, is that all *our* purposes are ultimately God's purposes. The idea is certainly natural and attractive. If we look at a pocket watch and wonder *why* it has a clear glass crystal on its face, the answer obviously harks back to the needs and desires of the users of watches, who want to tell time, by looking at the hands through the transparent, protective glass, and so forth. If it weren't for these facts about *us*, for whom the watch was created, there would be no explanation of the "why" of its crystal. If the universe was created by God, for God's purposes, then all the purposes we can find in it must ultimately be due to God's purposes. But what are God's purposes? That is something of a mystery.

One way of deflecting discomfort about that mystery is to switch the topic slightly. Instead of responding to the "why" question with a "because"-type answer (the sort of answer it seems to demand), people often substitute a "how" question for the "why" question, and attempt to answer it by telling a story about *how it came to be* that God created us and the rest of the universe, without dwelling overmuch on just why God might want to have done that. The "how" question does not get separate billing on Aristotle's list, but it was a popular question and answer long before Aristotle undertook his analysis. The answers to the biggest "how" questions are *cosmogonies*, stories about how the *cosmos*, the whole universe and all its denizens, came into existence. The book of Genesis is a cosmogony, but there are many others. Cosmologists exploring the hypothesis of the Big Bang, and speculating about black holes and superstrings, are present-day creators of cosmogonies. Not all ancient cosmogonies follow the pattern of an artifact-maker. Some involve a "world egg" laid in "the Deep" by one mythic bird or another, and some involve seeds' being sown and tended. Human imagination has only a few resources to draw upon when faced with such a mind-boggling question. One early creation myth speaks of a "self-existent Lord" who, "with a thought, created the waters, and deposited in them a seed which became a golden egg, in which egg he himself is born as Brahma, the progenitor of the worlds" (Muir 1972, vol. IV, p. 26).

And what's the point of all this egg-laying or seed-sowing or world-building? Or, for that matter, what's the point of the Big Bang? Today's cosmologists, like many of their predecessors throughout history, tell a diverting story, but prefer to sidestep the "why" question of teleology. Does the universe exist for any reason? Do reasons play any intelligible role in explanations of the cosmos? Could something exist for a reason without its being *somebody's* reason? Or are reasons—Aristotle's type (4) causes—only appropriate in explanations of the works and deeds of people or other rational agents? If God is not a person, a rational agent, an Intelligent Artificer, what possible sense could the biggest "why" question make? And if the biggest "why" question doesn't make any sense, how could any smaller, more parochial, "why" questions make sense?

One of Darwin's most fundamental contributions is showing us a new way to make sense of "why" questions. Like it or not, Darwin's idea offers one way—a clear, cogent, astonishingly versatile way—of dissolving these old conundrums. It takes some getting used to, and is often misapplied, even by its staunchest friends. Gradually exposing and clarifying this way of thinking is a central project of the present book. Darwinian thinking must be carefully distinguished from some oversimplified and all-too-popular impostors, and this will take us into some technicalities, but it is worth it. The prize is, for the first time, a stable system of explanation that does not go round and round in circles or spiral off in an infinite regress of mysteries. Some people would much prefer the infinite regress of mysteries, apparently, but in this day and age the cost is prohibitive: you have to get yourself deceived. You can either deceive yourself or let others do the dirty work, but there is no intellectually defensible way of rebuilding the mighty barriers to comprehension that Darwin smashed.

The first step to appreciating this aspect of Darwin's contribution is to see how the world looked before he inverted it. By looking through the eyes of two of his countrymen, John Locke and David Hume, we can get a clear vision of an alternative world-view—still very much with us in many quarters—that Darwin rendered obsolete.

3. LOCKE'S "PROOF" OF THE PRIMACY OF MIND

John Locke invented common sense, and only Englishmen have had it ever since!

—BERTRAND RUSSELL[2]

John Locke, a contemporary of "the incomparable Mr. Newton," was one of the founding fathers of British Empiricism, and, as befits an Empiricist, he was not much given to deductive arguments of the rationalist sort, but one of his uncharacteristic forays into "proof" deserves to be quoted in full, since it perfectly illustrates the blockade to imagination that was in place before the Darwinian Revolution. The argument may seem strange and stilted to modern minds, but bear with it—consider it a sign of how far we have come since then. Locke himself thought that he was just reminding people of something obvious! In this passage from his *Essay Concerning Human Understanding* (1690, IV, x, 10), Locke wanted to *prove* something that he thought all people knew in their hearts in any case: that "in the beginning" there was Mind. He began by asking himself what, if anything, was eternal:

> If, then, there must be something eternal, let us see what sort of Being it must be. And to that it is very obvious to Reason, that it must necessarily be a cogitative Being. For it is as impossible to conceive that ever bare incogitative Matter should produce a thinking intelligent Being, as that nothing should of itself produce Matter. . . .

Locke begins his proof by alluding to one of philosophy's most ancient and oft-used maxims, *Ex nihilo nihil fit*: nothing can come from nothing. Since this is to be a deductive argument, he must set his sights high: it is not just unlikely or implausible or hard to fathom but *impossible to conceive* that "bare incogitative Matter should produce a thinking intelligent Being." The argument proceeds by a series of mounting steps:

---

2. Gilbert Ryle recounted this typical bit of Russellian hyperbole to me. In spite of Ryle's own distinguished career as Waynflete Professor of Philosophy at Oxford, he and Russell had seldom met, he told me, in large measure because Russell steered clear of academic philosophy after the Second World War. Once, however, Ryle found himself sharing a compartment with Russell on a tedious train journey, and, trying desperately to make conversation with his world-famous fellow traveler, Ryle asked him why he thought Locke, who was neither as original nor as good a writer as Berkeley, Hume, or Reid, had been so much more influential than they in the English-speaking philosophical world. This had been his reply, and the beginning of the only good conversation, Ryle said, that he ever had with Russell.

> Let us suppose any parcel of Matter eternal, great or small, we shall find it, in itself, able to produce nothing.... Matter then, by its own strength, cannot produce in itself so much as Motion: the Motion it has, must also be from Eternity, or else be produced, and added to Matter by some other Being more powerful than Matter.... But let us suppose Motion eternal too: yet Matter, incogitative Matter and Motion, whatever changes it might produce of Figure and Bulk, could never produce Thought: Knowledge will still be as far beyond the power of Motion and Matter to produce, as Matter is beyond the power of nothing or nonentity to produce. And I appeal to everyone's own thoughts, whether he cannot as easily conceive Matter produced by nothing, as Thought produced by pure Matter, when before there was no such thing as Thought, or an intelligent Being existing....

It is interesting to note that Locke decides he may safely "appeal to everyone's own thoughts" to secure this "conclusion." He was sure that *his* "common sense" was truly *common* sense. Don't we see how obvious it is that whereas matter and motion could produce changes of "Figure and Bulk," they could *never* produce "Thought"? Wouldn't this rule out the prospect of robots—or at least robots that would claim to have genuine Thoughts among the motions in their material heads? Certainly in Locke's day—which was also Descartes's day—the very idea of Artificial Intelligence was so close to unthinkable that Locke could confidently expect unanimous endorsement of this appeal to his audience, an appeal that would risk hoots of derision today.[3] And as we shall see, the field of Artificial Intelligence is a quite direct descendant of Darwin's idea. Its birth, which was all but prophesied by Darwin himself, was attended by one of the first truly impressive demonstrations of the formal power of natural selection (Art Samuel's legendary checkers-playing program, which will be described in some detail later). And both evolution and AI inspire the same loathing in many people who should know better, as we shall see in later chapters. But back to Locke's conclusion:

> So if we will suppose nothing first, or eternal: Matter can never begin to be: If we suppose bare Matter, without Motion, eternal: Motion can never begin to be: If we suppose only Matter and Motion first, or eternal: Thought can never begin to be. For it is impossible to conceive that Matter either with or without Motion could have originally in and from itself Sense,

3. Descartes's inability to think of Thought as Matter in Motion is discussed at length in my book *Consciousness Explained* (1991a). John Haugeland's aptly titled book, *Artificial Intelligence: The Very Idea* (1985), is a fine introduction to the philosophical paths that make this idea thinkable after all.

Perception, and Knowledge, as is evident from hence, that then Sense, Perception, and Knowledge must be a property eternally inseparable from Matter and every particle of it.

So, if Locke is right, Mind must come first—or at least tied for first. It could not come into existence at some later date, as an effect of some confluence of more modest, mindless phenomena. This purports to be an entirely secular, logical—one might almost say mathematical—vindication of a central aspect of Judeo-Christian (and also Islamic) cosmogony: in the beginning was something with Mind—"a cogitative Being," as Locke says. The traditional idea that God is a rational, thinking agent, a Designer and Builder of the world, is here given the highest stamp of scientific approval: like a mathematical theorem, its denial is supposedly impossible to conceive.

And so it seemed to many brilliant and skeptical thinkers before Darwin. Almost a hundred years after Locke, another great British Empiricist, David Hume, confronted the issue again, in one of the masterpieces of Western philosophy, his *Dialogues Concerning Natural Religion* (1779).

4. HUME'S CLOSE ENCOUNTER

Natural religion, in Hume's day, meant a religion that was supported by the natural sciences, as opposed to a "revealed" religion, which would depend on revelation—on mystical experience or some other uncheckable source of conviction. If your only grounds for your religious belief is "God told me so in a dream," your religion is not natural religion. The distinction would not have made much sense before the dawn of modern science in the seventeenth century, when science created a new, and competitive, standard of evidence for all belief. It opened up the question:

Can you give us any *scientific* grounds for your religious beliefs?

Many religious thinkers, appreciating that the prestige of scientific thought was—other things being equal—a worthy aspiration, took up the challenge. It is hard to see why anybody would want to shun scientific confirmation of one's creed, if it were there to be had. The overwhelming favorite among purportedly scientific arguments for religious conclusions, then and now, was one version or another of the Argument from Design: among the effects we can objectively observe in the world, there are many that are not (cannot be, for various reasons) mere accidents; they must have been designed to be as they are, and there cannot be design without a Designer; therefore, a Designer, God, must exist (or have existed), as the source of all these wonderful effects.

Such an argument can be seen as an attempt at an alternate route to Locke's conclusion, a route that will take us through somewhat more empirical detail instead of relying so bluntly and directly on what is deemed inconceivable. The actual features of the observed designs may be analyzed, for instance, to secure the grounds for our appreciation of the wisdom of the Designer, and our conviction that mere chance could not be responsible for these marvels.

In Hume's *Dialogues*, three fictional characters pursue the debate with consummate wit and vigor. Cleanthes defends the Argument from Design, and gives it one of its most eloquent expressions.[4] Here is his opening statement of it:

> Look round the world: Contemplate the whole and every part of it: You will find it to be nothing but one great machine, subdivided into an infinite number of lesser machines, which again admit of subdivisions to a degree beyond what human senses and faculties can trace and explain. All these various machines, and even their most minute parts, are adjusted to each other with an accuracy which ravishes into admiration all men who have ever contemplated them. The curious adapting of means to ends, throughout all nature, resembles, exactly, though it much exceeds, the productions of human contrivance—of human design, thought, wisdom, and intelligence. Since therefore the effects resemble each other, we are led to infer, by all the rules of analogy, that the causes also resemble, and that the Author of Nature is somewhat similar to the mind of man, though possessed of much larger faculties, proportioned to the grandeur of the work which he has executed. By this argument *a posteriori*, and by this argument alone, do we prove at once the existence of a Deity and his similarity to human mind and intelligence. [Pt. II.]

Philo, a skeptical challenger to Cleanthes, elaborates the argument, setting it up for demolition. Anticipating Paley's famous example, Philo notes: "Throw several pieces of steel together, without shape or form; they will never arrange themselves so as to compose a watch."[5] He goes on: "Stone, and mortar, and wood, without an architect, never erect a house. But the

4. William Paley carried the Argument from Design into much greater biological detail in his 1803 book, *Natural Theology*, adding many ingenious flourishes. Paley's influential version was the actual inspiration and target of Darwin's rebuttal, but Hume's Cleanthes catches all of the argument's logical and rhetorical force.

5. Gjertsen points out that two millennia earlier, Cicero used the same example for the same purpose: "When you see a sundial or a water-clock, you see that it tells the time by design and not by chance. How then can you imagine that the universe as a whole is devoid of purpose and intelligence, when it embraces everything, including these artifacts themselves and their artificers?" (Gjertsen 1989, p. 199).

ideas in a human mind, we see, by an unknown, inexplicable economy, arrange themselves so as to form the plan of a watch or house. Experience, therefore, proves, that there is an original principle of order in mind, not in matter" (Pt. II).

Note that the Argument from Design depends on an inductive inference: where there's smoke, there's fire; and where there's design, there's mind. But this is a dubious inference, Philo observes: human intelligence is

> no more than one of the springs and principles of the universe, as well as heat or cold, attraction or repulsion, and a hundred others, which fall under daily observation.... But can a conclusion, with any propriety, be transferred from parts to the whole?... From observing the growth of a hair, can we learn any thing concerning the generation of a man?... What peculiar privilege has this little agitation of the brain which we call thought, that we must thus make it the model of the whole universe?... Admirable conclusion! Stone, wood, brick, iron, brass have not, at this time, in this minute globe of earth, an order or arrangement without human art and contrivance: Therefore the universe could not originally attain its order and arrangement, without something similar to human art. [Pt. II.]

Besides, Philo observes, if we put mind as the first cause, with its "unknown, inexplicable economy," this only postpones the problem:

> We are still obliged to mount higher, in order to find the cause of this cause, which you had assigned as satisfactory and conclusive.... How therefore shall we satisfy ourselves concerning the cause of that Being, whom you suppose the Author of nature, or, according to your system of anthropomorphism, the ideal world, into which you trace the material? Have we not the same reason to trace that ideal world into another ideal world, or new intelligent principle? But if we stop, and go no farther; why go so far? Why not stop at the material world? How can we satisfy ourselves without going on *in infinitum*? And after all, what satisfaction is there in that infinite progression? [Pt. IV.]

Cleanthes has no satisfactory responses to these rhetorical questions, and there is worse to come. Cleanthes insists that God's mind is *like the human*—and agrees when Philo adds "the liker the better." But, then, Philo presses on, is God's mind perfect, "free from every error, mistake, or incoherence in his undertakings" (Pt. V)? There is a rival hypothesis to rule out:

> And what surprise must we entertain, when we find him a stupid mechanic, who imitated others, and copied an art, which, through a long succession of ages, after multiplied trials, mistakes, corrections, deliberations, and controversies, had been gradually improving? Many worlds might have

> been botched and bungled, throughout an eternity, ere this system was struck out: Much labour lost: Many fruitless trials made: And a slow, but continued improvement carried on during infinite ages of world-making. [Pt. V.]

When Philo presents this fanciful alternative, with its breathtaking anticipations of Darwin's insight, he doesn't take it seriously except as a debating foil to Cleanthes' vision of an all-wise Artificer. Hume uses it only to make a point about what he saw as the limitations on our knowledge: "In such subjects, who can determine, where the truth; nay, who can conjecture where the probability, lies; amidst a great number of hypotheses which may be proposed, and a still greater number which may be imagined" (Pt. V).

Imagination runs riot, and, exploiting that fecundity, Philo ties Cleanthes up in knots, devising weird and comical variations on Cleanthes' own hypotheses, defying Cleanthes to show why his own version should be preferred. "Why may not several Deities combine in contriving and framing a world? . . . And why not become a perfect anthropomorphite? Why not assert the Deity or Deities to be corporeal, and to have eyes, a nose, mouth, ears, etc.?" (Pt. V). At one point, Philo anticipates the Gaia hypothesis: the universe

> bears a great resemblance to an animal or organized body, and seems actuated with a like principle of life and motion. A continual circulation of matter in it produces no disorder. . . . The world, therefore, I infer, is an animal, and the Deity is the SOUL of the world, actuating it and actuated by it. [Pt. VI.]

Or perhaps isn't the world really more like a vegetable than an animal?

> In like manner as a tree sheds its seed into the neighboring fields, and produces other trees; so the great vegetable, the world, or this planetary system, produces within itself certain seeds, which, being scattered into the surrounding chaos, vegetate into new worlds. A comet, for instance, is the seed of a world. . . . [Pt. VII.]

One more wild possibility for good measure:

> The Brahmins assert, that the world arose from an infinite spider, who spun this whole complicated mass from his bowels, and annihilates afterwards the whole or any part of it, by absorbing it again, and resolving it into his own essence. Here is a species of cosmogony, which appears to us ridiculous; because a spider is a little contemptible animal, whose operation we are never likely to take for a model of the whole universe. But still here is

> a new species of analogy, even in our globe. And were there a planet wholly inhabited by spiders (which is very possible), this inference would there appear as natural and irrefragable as that which in our planet ascribes the origin of all things to design and intelligence, as explained by Cleanthes. Why an orderly system may not be spun from the belly as well as from the brain, it will be difficult for him to give a satisfactory reason. [Pt. VII.]

Cleanthes resists these onslaughts gamely, but Philo shows fatal flaws in every version of the argument that Cleanthes can devise. At the very end of the *Dialogues*, however, Philo surprises us by agreeing with Cleanthes:

> ... the legitimate conclusion is that ... if we are not contented with calling the first and supreme cause a *God* or *Deity*, but desire to vary the expression, what can we call him but *Mind* or *Thought* to which he is justly supposed to bear a considerable resemblance? [Pt. XII.]

Philo is surely Hume's mouthpiece in the *Dialogues*. Why did Hume cave in? Out of fear of reprisal from the establishment? No. Hume knew he had shown that the Argument from Design was an irreparably flawed bridge between science and religion, and he arranged to have his *Dialogues* published after his death in 1776 precisely in order to save himself from persecution. He caved in because he *just couldn't imagine* any other explanation of the origin of the manifest design in nature. Hume could not see how the "curious adapting of means to ends, throughout all nature" could be due to chance—and if not chance, what?

What could possibly account for this high-quality design if not an intelligent God? Philo is one of the most ingenious and resourceful competitors in any philosophical debate, real or imaginary, and he makes some wonderful stabs in the dark, hunting for an alternative. In Part VIII, he dreams up some speculations that come tantalizingly close to scooping Darwin (and some more recent Darwinian elaborations) by nearly a century.

> Instead of supposing matter infinite, as Epicurus did, let us suppose it finite. A finite number of particles is only susceptible of finite transpositions: And it must happen, in an eternal duration, that every possible order or position must be tried an infinite number of times.... Is there a system, an order, an economy of things, by which matter can preserve that perpetual agitation, which seems essential to it, and yet maintain a constancy in the forms, which it produces? There certainly is such an economy: For this is actually the case with the present world. The continual motion of matter, therefore, in less than infinite transpositions, must produce this economy or order; and by its very nature, that order, when once established, supports itself, for many ages, if not to eternity. But wherever matter is so poised, arranged, and adjusted as to continue in perpetual motion, and yet pre-

> serve a constancy in the forms, its situation must, of necessity, have all the same appearance of art and contrivance which we observe at present.... A defect in any of these particulars destroys the form; and the matter, of which it is composed, is again set loose, and is thrown into irregular motions and fermentations, till it unite itself to some other regular form....
>
> Suppose ... that matter were thrown into any position, by a blind, unguided force; it is evident that this first position must in all probability be the most confused and most disorderly imaginable, without any resemblance to those works of human contrivance, which, along with a symmetry of parts, discover an adjustment of means to ends and a tendency to self-preservation.... Suppose, that the actuating force, whatever it be, still continues in matter.... Thus the universe goes on for many ages in a continued succession of chaos and disorder. But is it not possible that it may settle at last ... ? May we not hope for such a position, or rather be assured of it, from the eternal revolutions of unguided matter, and may not this account for all the appearing wisdom and contrivance which is in the universe?

Hmm, it seems that something like this might work ... but Hume couldn't quite take Philo's daring foray seriously. His final verdict: "A total suspense of judgment is here our only reasonable resource" (Pt. VIII). A few years before him, Denis Diderot had also written some speculations that tantalizingly foreshadowed Darwin: "I can maintain to you ... that monsters annihilated one another in succession; that all the defective combinations of matter have disappeared, and that there have only survived those in which the organization did not involve any important contradiction, and which could subsist by themselves and perpetuate themselves" (Diderot 1749). Cute ideas about evolution had been floating around for millennia, but, like most philosophical ideas, although they did seem to offer a solution of sorts to the problem at hand, they didn't promise to go any farther, to open up new investigations or generate surprising predictions that could be tested, or explain any facts they weren't expressly designed to explain. The evolution revolution had to wait until Charles Darwin saw how to weave an evolutionary hypothesis into an explanatory fabric composed of literally thousands of hard-won and often surprising facts about nature. Darwin neither invented the wonderful idea out of whole cloth all by himself, nor understood it in its entirety even when he had formulated it. But he did such a monumental job of clarifying the idea, and tying it down so it would never again float away, that he deserves the credit if anyone does. The next chapter reviews his basic accomplishment.

CHAPTER 1: *Before Darwin, a "Mind-first" view of the universe reigned unchallenged; an intelligent God was seen as the ultimate source of all Design, the ultimate answer to any chain of "Why?" questions. Even David*

Hume, who deftly exposed the insoluble problems with this vision, and had glimpses of the Darwinian alternative, could not see how to take it seriously.

CHAPTER 2: *Darwin, setting out to answer a relatively modest question about the origin of species, described a process he called natural selection, a mindless, purposeless, mechanical process. This turns out to be the seed of an answer to a much grander question: how does Design come into existence?*

CHAPTER TWO

An Idea Is Born

1. What Is So Special About Species?

Charles Darwin did not set out to concoct an antidote to John Locke's conceptual paralysis, or to pin down the grand cosmological alternative that had barely eluded Hume. Once his great idea occurred to him, he saw that it would indeed have these truly revolutionary consequences, but at the outset he was not trying to explain the meaning of life, or even its origin. His aim was slightly more modest: he wanted to explain the origin of *species*.

In his day, naturalists had amassed mountains of tantalizing facts about living things and had succeeded in systematizing these facts along several dimensions. Two great sources of wonder emerged from this work (Mayr 1982). First, there were all the discoveries about the *adaptations* of organisms that had enthralled Hume's Cleanthes: "All these various machines, and even their most minute parts, are adjusted to each other with an accuracy which ravishes into admiration all men who have ever contemplated them" (Pt. II). Second, there was the prolific *diversity* of living things—literally millions of different kinds of plants and animals. Why were there so many?

This diversity of design of organisms was as striking, in some regards, as their excellence of design, and even more striking were the patterns discernible within that diversity. Thousands of gradations and variations between organisms could be observed, but there were also huge gaps between them. There were birds and mammals that swam like fish, but none with gills; there were dogs of many sizes and shapes, but no dogcats or dogcows or feathered dogs. The patterns called out for classification, and by Darwin's time the work of the great taxonomists (who began by adopting and correcting Aristotle's ancient classifications) had created a detailed hierarchy of two kingdoms (plants and animals), divided into phyla, which divided into classes, which divided into orders, which divided into families, which divided into genera (the plural of "genus"), which divided into species.

Species could also be subdivided, of course, into subspecies or varieties—cocker spaniels and basset hounds are different varieties of a single species: dogs, or *Canis familiaris*.

How many different kinds of organisms were there? Since no two organisms are exactly alike—not even identical twins—there were as many different kinds of organisms as there were organisms, but it seemed obvious that the differences could be graded, sorted into minor and major, or *accidental* and *essential*. Thus Aristotle had taught, and this was one bit of philosophy that had permeated the thinking of just about everybody, from cardinals to chemists to costermongers. All things—not just living things—had two kinds of properties: essential properties, without which they wouldn't be the particular *kind* of thing they were, and accidental properties, which were free to vary within the kind. A lump of gold could change shape *ad lib* and still be gold; what made it gold were its essential properties, not its accidents. With each kind went an essence. Essences were definitive, and as such they were timeless, unchanging, and all-or-nothing. A thing couldn't be *rather* silver or *quasi*-gold or a *semi*-mammal.

Aristotle had developed his theory of essences as an improvement on Plato's theory of Ideas, according to which every earthly thing is a sort of imperfect copy or reflection of an ideal exemplar or Form that existed timelessly in the Platonic realm of Ideas, reigned over by God. This Platonic heaven of abstractions was not visible, of course, but was accessible to Mind through deductive thought. What geometers thought about, and proved theorems about, for instance, were the Forms of the circle and the triangle. Since there were also Forms for the eagle and the elephant, a deductive science of nature was also worth a try. But just as no earthly circle, no matter how carefully drawn with a compass, or thrown on a potter's wheel, could actually be one of the perfect circles of Euclidean geometry, so no actual eagle could perfectly manifest the essence of eaglehood, though every eagle strove to do so. Everything that existed had a divine specification, which captured its essence. The taxonomy of living things Darwin inherited was thus itself a direct descendant, via Aristotle, of Plato's essentialism. In fact, the word "species" was at one point a standard translation of Plato's Greek word for Form or Idea, *eidos*.

We post-Darwinians are so used to thinking in historical terms about the development of life forms that it takes a special effort to remind ourselves that in Darwin's day species of organisms were deemed to be as timeless as the perfect triangles and circles of Euclidean geometry. Their individual members came and went, but the species itself remained unchanged and unchangeable. This was part of a philosophical heritage, but it was not an idle or ill-motivated dogma. The triumphs of modern science, from Copernicus and Kepler, Descartes and Newton, had all involved the application of precise mathematics to the material world, and this apparently requires abstracting away from the grubby accidental properties of things to find their secret mathematical essences. It makes no difference what color or shape a thing is when it comes to the thing's obeying Newton's inverse-square law of gravitational attraction. All that matters is its mass. Similarly, alchemy had been succeeded by chemistry once chemists settled on their fundamental creed: There were a finite number of basic, *immutable* elements, such as carbon, oxygen, hydrogen, and iron. These might be mixed and united in endless combinations over time, but the fundamental building blocks were identifiable by their changeless essential properties.

The doctrine of essences looked like a powerful organizer of the world's phenomena in many areas, but was it true of every classification scheme one could devise? Were there *essential* differences between hills and mountains, snow and sleet, mansions and palaces, violins and violas? John Locke and others had developed elaborate doctrines distinguishing *real* essences from merely *nominal* essences; the latter were simply parasitic on the *names* or words we chose to use. You could set up any classification scheme you wanted; for instance, a kennel club could vote on a defining list of necessary conditions for a dog to be a genuine Ourkind Spaniel, but this would be a mere nominal essence, not a real essence. Real essences were discoverable by scientific investigation into the internal nature of things, where essence and accident could be distinguished according to principles. It was hard to say just what the *principled* principles were, but with chemistry and physics so handsomely falling into line, it seemed to stand to reason that there had to be defining marks of the real essences of living things as well.

From the perspective of this deliciously crisp and systematic vision of the hierarchy of living things, there were a considerable number of awkward and puzzling facts. These apparent exceptions were almost as troubling to naturalists as the discovery of a triangle whose angles didn't quite add up to 180 degrees would have been to a geometer. Although many of the taxonomic boundaries were sharp and apparently exceptionless, there were all manner of hard-to-classify intermediate creatures, who seemed to have portions of more than one essence. There were also the curious higher-order patterns of shared and unshared features: why should it be backbones rather than feathers that birds and fish shared, and why shouldn't *creature with eyes* or *carnivore* be as important a classifier as *warm-blooded creature*? Although the broad outlines and most of the specific rulings of taxonomy were undisputed (and remain so today, of course), there were heated controversies about the problem cases. Were all these lizards members of the same species, or of several different species? Which principle of classification should "count"? In Plato's famous image, which system "carved nature at the joints"?

Before Darwin, these controversies were fundamentally ill-formed, and could not yield a stable, well-motivated answer because there was no background theory of *why* one classification scheme would count as getting the joints right—the way things *really* were. Today bookstores face the same sort of ill-formed problem: how should the following categories be cross-organized: best-sellers, science fiction, horror, garden, biography, novels, collections, sports, illustrated books? If horror is a genus of fiction, then true tales of horror present a problem. Must all novels be fiction? Then the bookseller cannot honor Truman Capote's own description of *In Cold Blood* (1965) as a nonfiction novel, but the book doesn't sit comfortably amid either the biographies or the history books. In what section of the bookstore should the book you are reading be shelved? Obviously there is no one Right Way to categorize books—nominal essences are all we will ever find in this domain. But many naturalists were convinced on general principles that there were real essences to be found among the categories of their Natural System of living things. As Darwin put it, "They believe that it reveals the plan of the Creator; but unless it be specified whether order in time or space, or what else is meant by the plan of the Creator, it seems to me that nothing is thus added to our knowledge" (*Origin*, p. 413).

Problems in science are sometimes made easier by adding complications. The development of the science of geology and the discovery of fossils of manifestly extinct species gave the taxonomists further curiosities to confound them, but these curiosities were also the very pieces of the puzzle that enabled Darwin, working alongside hundreds of other scientists, to discover the key to its solution: species were *not* eternal and immutable; they had evolved over time. Unlike carbon atoms, which, for all one knew, had been around forever in exactly the form they now exhibited, species had births in time, could change over time, and could give birth to new species in turn. This idea itself was not new; many versions of it had been seriously discussed, going back to the ancient Greeks. But there was a powerful Platonic bias against it: essences were unchanging, and a thing couldn't change its essence, and new essences couldn't be born—except of course by God's command in episodes of Special Creation. Reptiles could no more *turn into* birds than copper could turn into gold.

It isn't easy today to sympathize with this conviction, but the effort can be helped along by a fantasy: consider what your attitude would be towards a theory that purported to show how the number 7 had once been an even number, long, long ago, and had gradually acquired its oddness through an arrangement whereby it exchanged some properties with the ancestors of the number 10 (which had once been a prime number). Utter nonsense, of course. Inconceivable. Darwin knew that a parallel attitude was deeply ingrained among his contemporaries, and that he would have to labor mightily to overcome it. Indeed, he more or less conceded that the elder authorities of his day would tend to be as immutable as the species they believed in, so in the conclusion of his book he went so far as to beseech the support of his younger readers: "Whoever is led to believe that species are mutable will do good service by conscientiously expressing his conviction; for only thus can the load of prejudice by which this subject is overwhelmed be removed" (*Origin*, p. 482).

Even today Darwin's overthrow of essentialism has not been completely assimilated. For instance, there is much discussion in philosophy these days about "natural kinds," an ancient term the philosopher W. V. O. Quine (1969) quite cautiously resurrected for limited use in distinguishing good scientific categories from bad ones. But in the writings of other philosophers, "natural kind" is often sheep's clothing for the wolf of real essence. The essentialist urge is still with us, and not always for bad reasons. Science does aspire to carve nature at its joints, and it often seems that we need essences, or something like essences, to do the job. On this one point, the two great kingdoms of philosophical thought, the Platonic and the Aristotelian, agree. But the Darwinian mutation, which at first seemed to be just a new way of thinking about kinds in biology, can spread to other phenomena and other disciplines, as we shall see. There are persistent problems both inside and outside biology that readily dissolve once we adopt the Darwinian perspective on what makes a thing the sort of thing it is, but the tradition-bound resistance to this idea persists.

2. Natural Selection—an Awful Stretcher

It is an awful stretcher to believe that a peacock's tail was thus formed; but, believing it, I believe in the same principle somewhat modified applied to man.

—Charles Darwin, letter quoted in Desmond and Moore 1991, p. 553

Darwin's project in *Origin* can be divided in two: to prove *that* modern species were revised descendants of earlier species—species had evolved—and to show *how* this process of "descent with modification" had occurred. If Darwin hadn't had a vision of a mechanism, natural selection, by which this well-nigh-inconceivable historical transformation could have been accomplished, he would probably not have had the motivation to assemble all the circumstantial evidence that it had actually occurred. Today we can readily enough imagine proving Darwin's first case—the brute historic fact of descent with modification—quite independently of any consideration of natural selection or indeed any other mechanism for bringing these brute events about, but for Darwin the idea of the mechanism was both the hunting license he needed, and an unwavering guide to the right questions to ask.[1]

The idea of natural selection was not itself a miraculously novel creation of Darwin's but, rather, the offspring of earlier ideas that had been vigorously discussed for years and even generations (for an excellent account of this intellectual history, see R. Richards 1987). Chief among these parent ideas was an insight Darwin gained from reflection on the 1798 *Essay on the Principle of Population* by Thomas Malthus, which argued that population explosion and famine were inevitable, given the excess fertility of human beings, unless drastic measures were taken. The grim Malthusian vision of the social and political forces that could act to check human overpopulation may have strongly flavored Darwin's thinking (and undoubtedly has flavored the shallow political attacks of many an anti-Darwinian), but the idea Darwin needed from Malthus is purely logical. It has nothing at all to do with political ideology, and can be expressed in very abstract and general terms.

Suppose a world in which organisms have many offspring. Since the offspring themselves will have many offspring, the population will grow and grow ("geometrically") until inevitably, sooner or later—surprisingly soon, in fact—it must grow too large for the available resources (of food, of space, of whatever the organisms need to survive long enough to reproduce). At that point, whenever it happens, not all organisms will have offspring. Many will die childless. It was Malthus who pointed out the mathematical inevitability of such a crunch in *any* population of long-term reproducers—people, animals, plants (or, for that matter, Martian clone-machines, not that such fanciful possibilities were discussed by Malthus). Those populations that reproduce at less than the replacement rate are headed for extinction unless they reverse the trend. Populations that maintain a stable population over long periods of time will do so by settling on a rate of overproduction of offspring that is balanced by the vicissitudes encountered. This is obvious, perhaps, for houseflies and other prodigious breeders, but Darwin drove the point home with a calculation of his own: "The elephant is reckoned to be the slowest breeder of all known animals, and I have taken some pains to estimate its probable minimum rate of natural increase: . . . at the end of the fifth century there would be alive fifteen million elephants, descended from the first pair" (*Origin*, p. 64).[2] Since elephants have been around for millions

---

1. This has often happened in science. For instance, for many years there was lots of evidence lying around in favor of the hypothesis that the continents have drifted—that Africa and South America were once adjacent and broke apart—but until the mechanisms of plate tectonics were conceived, it was hard to take the hypothesis seriously.

2. This sum as it appeared in the first edition is wrong, and when this was pointed out, Darwin revised his calculations for later editions, but the general principle is still unchallenged.

of years, we can be sure that only a fraction of the elephants born in any period have progeny of their own.

So the normal state of affairs for any sort of reproducers is one in which more offspring are produced in any one generation than will in turn reproduce in the next. In other words, it is almost always crunch time.[3] At such a crunch, which prospective parents will "win"? Will it be a fair lottery, in which every organism has an equal chance of being among the few that reproduce? In a political context, this is where invidious themes enter, about power, privilege, injustice, treachery, class warfare, and the like, but we can elevate the observation from its political birthplace and consider in the abstract, as Darwin did, what would—must—happen in nature. Darwin added two further logical points to the insight he had found in Malthus: the first was that at crunch time, if there was significant variation among the contestants, then any advantages enjoyed by any of the contestants would inevitably bias the sample that reproduced. However tiny the advantage in question, if it was actually an advantage (and thus not absolutely invisible to nature), it would tip the scales in favor of those who held it. The second was that *if* there was a "strong principle of inheritance"—if offspring tended to be more like their parents than like their parents' contemporaries—the biases created by advantages, however small, would become amplified over time, creating trends that could grow indefinitely. "More individuals are born than can possibly survive. A grain in the balance will determine which individual shall live and which shall die,—which variety or species shall increase in number, and which shall decrease, or finally become extinct" (*Origin*, p. 467).

What Darwin saw was that if one merely supposed these few general conditions to apply at crunch time—conditions for which he could supply ample evidence—the resulting process would *necessarily* lead in the direction of individuals in future generations who tended to be better equipped to deal with the problems of resource limitation that had been faced by the individuals of their parents' generation. This fundamental idea—Darwin's dangerous idea, the idea that generates so much insight, turmoil, confusion, anxiety—is thus actually quite simple. Darwin summarizes it in two long sentences at the end of chapter 4 of *Origin*:

> If during the long course of ages and under varying conditions of life, organic beings vary at all in the several parts of their organization, and I

---

3. A familiar example of Malthus' rule in action is the rapid expansion of yeast populations introduced into fresh bread dough or grape juice. Thanks to the feast of sugar and other nutrients, population explosions ensue that last for a few hours in the dough, or a few weeks in the juice, but soon the yeast populations hit the Malthusian ceiling, done in by their own voraciousness and the accumulation of their waste products—carbon dioxide (which forms the bubbles that make the bread rise, and the fizz in champagne) and alcohol being the two that we yeast-exploiters tend to value.

> think this cannot be disputed; if there be, owing to the high geometric powers of increase of each species, at some age, season, or year, a severe struggle for life, and this certainly cannot be disputed; then, considering the infinite complexity of the relations of all organic beings to each other and to their conditions of existence, causing an infinite diversity in structure, constitution, and habits, to be advantageous to them, I think it would be a most extraordinary fact if no variation ever had occurred useful to each being's own welfare, in the same way as so many variations have occurred useful to man. But if variations useful to any organic being do occur, assuredly individuals thus characterized will have the best chance of being preserved in the struggle for life; and from the strong principle of inheritance they will tend to produce offspring similarly characterized. This principle of preservation, I have called, for the sake of brevity, Natural Selection. [*Origin*, p. 127.]

This was Darwin's great idea, not the idea of evolution, but the idea of evolution *by natural selection*, an idea he himself could never formulate with sufficient rigor and detail to prove, though he presented a brilliant case for it. The next two sections will concentrate on curious and crucial features of this summary statement of Darwin's.

3. Did Darwin Explain the Origin of Species?

> *Darwin did wrestle brilliantly and triumphantly with the problem of adaptation, but he had limited success with the issue of diversity—even though he titled his book with reference to his relative failure: the origin of species.*
>
> —Stephen Jay Gould 1992a, p. 54

> *Thus the grand fact in natural history of the subordination of group under group, which, from its familiarity, does not always sufficiently strike us, is in my judgment fully explained.*
>
> —Charles Darwin, *Origin*, p. 413

Notice that Darwin's summary does not mention speciation at all. It is entirely about the adaptation of organisms, the *excellence* of their design, not the diversity. Moreover, on the face of it, this summary takes the diversity of species *as an assumption*: "the infinite [sic] complexity of the relations of all organic beings to each other and to their conditions of existence." What makes for this stupendous (if not actually infinite) complexity is the presence at one and the same time (and competing for the same living space) of so many different life forms, with so many different needs and strategies. Darwin doesn't even purport to offer an explanation of the origin of the *first* species, or of life itself; he begins in the middle, supposing many different species with many different talents already present, and claims that starting from such a mid-stage point, the process he has described will inevitably hone and diversify the talents of the species already existing. And will that process create still further species? The summary is silent on that score, but the book is not. In fact, Darwin saw his idea explaining both great sources of wonder in a single stroke. The generation of adaptations and the generation of diversity were different aspects of a single complex phenomenon, and the unifying insight, he claimed, was the principle of natural selection.

Natural selection would inevitably produce *adaptation*, as the summary makes clear, and under the right circumstances, he argued, accumulated adaptation would create speciation. Darwin knew full well that explaining variation is not explaining speciation. The animal-breeders he pumped so vigorously for their lore knew about how to breed *variety* within a single species, but had apparently never created a new *species*, and scoffed at the idea that their particular different breeds might have a common ancestor. "Ask, as I have asked, a celebrated raiser of Hereford cattle, whether his cattle might not have descended from longhorns, and he will laugh you to scorn." Why? Because "though they well know that each race varies slightly, for they win their prizes by selecting such slight differences, yet they ignore all general arguments and refuse to sum up in their minds slight differences accumulated during many successive generations" (*Origin*, p. 29).

The further diversification into species would occur, Darwin argued, because if there was a variety of heritable skills or equipment in a population (of a single species), these different skills or equipment would tend to have different payoffs for different subgroups of the population, and hence these subpopulations would tend to diverge, each one pursuing its favored sort of excellence, until eventually there would be a complete parting of the ways. Why, Darwin asked himself, would this divergence lead to separation or clumping of the variations instead of remaining a more or less continuous fan-out of slight differences? Simple geographical isolation was part of his answer; when a population got split by a major geological or climatic event, or by haphazard emigration to an isolated range such as an island, this discontinuity in the environment ought to become mirrored eventually in a discontinuity in the useful variations observable in the two populations. And once discontinuity got a foothold, it would be self-reinforcing, all the way to separation into distinct species. Another, rather different, idea of his was that in intraspecific infighting, a "winner take all" principle would tend to operate:

> For it should be remembered that the competition will generally be most severe between those forms which are most nearly related to each other in habits, constitution and structure. Hence all the intermediate forms

> between the earlier and later states, that is between the less and more improved state of a species, as well as the original parent-species itself, will generally tend to become extinct. [*Origin*, p. 121.]

He formulated a variety of other ingenious and plausible speculations on how and why the relentless culling of natural selection would actually create species boundaries, but they remain speculations to this day. It has taken a century of further work to replace Darwin's brilliant but inconclusive musings on the mechanisms of speciation with accounts that are to some degree demonstrable. Controversy about the mechanisms and principles of speciation still persists, so in one sense neither Darwin nor any subsequent Darwinian has explained the origin of species. As the geneticist Steve Jones (1993) has remarked, had Darwin published his masterpiece under its existing title today, "he would have been in trouble with the Trades Description Act because if there is one thing which *Origin of Species* is not about, it is the origin of species. Darwin knew nothing about genetics. Now we know a great deal, and although the way in which species begin is still a mystery, it is one with the details filled in."

But the fact of speciation itself is incontestable, as Darwin showed, building an irresistible case out of literally hundreds of carefully studied and closely argued instances. That is how species originate: by "descent with modification" from earlier species—not by Special Creation. So in another sense Darwin undeniably did explain the origin of species. Whatever the mechanisms are that operate, they manifestly begin with the emergence of variety within a species, and end, after modifications have accumulated, with the birth of a new, descendant species. What start as "well-marked varieties" turn gradually into "the doubtful category of subspecies; but we have only to suppose the steps in the process of modification to be more numerous or greater in amount, to convert these ... forms into well-defined species" (*Origin*, p. 120).

Notice that Darwin is careful to describe the eventual outcome as the creation of "well-defined" species. Eventually, he is saying, the divergence becomes so great that there is just no reason to deny that what we have are two different species, not merely two different varieties. But he declines to play the traditional game of declaring what the "essential" difference is:

> ... it will be seen that I look at the term species, as one arbitrarily given for the sake of convenience to a set of individuals closely resembling each other, and that it does not essentially differ from the term variety, which is given to less distinct and more fluctuating forms. [*Origin*, p. 52.]

One of the standard marks of species difference, as Darwin fully recognized, is reproductive isolation—there is no interbreeding. It is interbreeding that reunites the splitting groups, mixing their genes and "frustrating" the process of speciation. It is not that anything *wants* speciation to happen, of course (Dawkins 1986a, p. 237), but if the irreversible divorce that marks speciation is to happen, it must be preceded by a sort of trial separation period in which interbreeding ceases for one reason or another, so that the parting groups can move further apart. The criterion of reproductive isolation is vague at the edges. Do organisms belong to different species when they *can't* interbreed, or when they just *don't* interbreed? Wolves and coyotes and dogs are considered to be different species, and yet interbreeding does occur, and—unlike mules, the offspring of horse and donkey—their offspring are not in general sterile. Dachshunds and Irish wolfhounds are deemed to be of the same species, but unless their owners provide some distinctly unnatural arrangements, they are about as reproductively isolated as bats are from dolphins. The white-tailed deer in Maine don't in fact interbreed with the white-tailed deer in Massachusetts, since they don't travel that far, but they surely could if transported, and naturally they count as of the same species.

And finally—a true-life example seemingly made to order for philosophers—consider the herring gulls that live in the Northern Hemisphere, their range forming a broad ring around the North Pole.

> As we look at the herring gull, moving westwards from Great Britain to North America, we see gulls that are recognizably herring gulls, although they are a little different from the British form. We can follow them, as their appearance gradually changes, as far as Siberia. At about this point in the continuum, the gull looks more like the form that in Great Britain is called the lesser black-backed gull. From Siberia, across Russia, to northern Europe, the gull gradually changes to look more and more like the British lesser black-backed gull. Finally, in Europe, the ring is complete; the two geographically extreme forms meet, to form two perfectly good species: the herring and lesser black-backed gull can be both distinguished by their appearance and do not naturally interbreed. [Mark Ridley 1985, p. 5.]

"Well-defined" species certainly do exist—it is the purpose of Darwin's book to explain their origin—but he discourages us from trying to find a "principled" definition of the concept of a species. Varieties, Darwin keeps insisting, are just "incipient species," and what normally turns two varieties into two species is not the *presence* of something (a new essence for each group, for instance) but the *absence* of something: the intermediate cases, which used to be there—which were necessary stepping-stones, you might say—but have eventually gone extinct, leaving two groups that are *in fact* reproductively isolated as well as different in their characteristics.

Origin of Species presents an overwhelmingly persuasive case for Darwin's first thesis—the historical fact of evolution as the cause of the origin of species—and a tantalizing case in favor of his second thesis—that the fundamental mechanism responsible for "descent with modification" was natural selection.[4] Levelheaded readers of the book simply could no longer doubt that species had evolved over the eons, as Darwin said they had, but scrupulous skepticism about the power of his proposed mechanism of natural selection was harder to overcome. Intervening years have raised the confidence level for both theses, but not erased the difference (Ellegård [1958] provides a valuable account of this history). The evidence for evolution pours in, not only from geology, paleontology, biogeography, and anatomy (Darwin's chief sources), but of course from molecular biology and every other branch of the life sciences. To put it bluntly but fairly, anyone today who doubts that the variety of life on this planet was produced by a process of evolution is simply ignorant—inexcusably ignorant, in a world where three out of four people have learned to read and write. Doubts about the power of Darwin's idea of natural selection to explain this evolutionary process are still intellectually respectable, however, although the burden of proof for such skepticism has become immense, as we shall see.

So, although Darwin depended on his idea of the mechanism of natural selection to inspire and guide his research on evolution, the end result reversed the order of dependence: he showed so convincingly that species *had* to have evolved that he could then turn around and use this fact to support his more radical idea, natural selection. He had described a mechanism or process that, according to his arguments, *could* have produced all these effects. Skeptics were presented with a challenge: Could they show that his arguments were mistaken? Could they show how natural selection would be incapable of producing the effects?[5] Or could they even describe

4. As is often pointed out, Darwin didn't insist that natural selection explained everything: it was the "main but not exclusive means of modification" (*Origin*, p. 6).

5. It is sometimes suggested that Darwin's theory is systematically irrefutable (and hence scientifically vacuous), but Darwin was forthright about what sort of finding it would take to refute his theory. "Though nature grants vast periods of time for the work of natural selection, she does not grant an indefinite period" (*Origin*, p. 102), so, if the geological evidence mounted to show that not enough time had elapsed, his whole theory would be refuted. This still left a temporary loophole, for the theory wasn't formulatable in sufficiently rigorous detail to say just how many millions of years was the minimal amount required, but it was a temporary loophole that made sense, since at least some proposals about its size could be evaluated independently. (Kitcher [1985a, pp. 162–65], has a good discussion of the further subtleties of argument that kept Darwinian theory from being directly confirmed or disconfirmed.) Another famous instance: "If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down" (*Origin*, p. 189). Many have risen to this challenge, but, as we shall see in chapter 11, there are good reasons why they have not succeeded in their attempted demonstrations.

another process that might achieve these effects? What *else* could account for evolution, if not the mechanism he had described?

This challenge effectively turned Hume's predicament inside out. Hume caved in because he could not imagine how anything other than an Intelligent Artificer could be the cause of the adaptations that anyone could observe. Or, more accurately, Hume's Philo imagined several different alternatives, *but Hume had no way of taking these imaginings seriously.* Darwin described how a Nonintelligent Artificer could produce those adaptations over vast amounts of time, and proved that many of the intermediate stages that would be needed by that proposed process had indeed occurred. Now the challenge to imagination was reversed: given all the telltale signs of the historical process that Darwin uncovered—all the brushmarks of the artist, you might say—could anyone imagine how any process *other* than natural selection could have produced all these effects? So complete has this reversal of the burden of proof been that scientists often find themselves in something like the mirror image of Hume's predicament. When they are confronted with a *prima facie* powerful and undismissable objection to natural selection (we will consider the strongest cases in due course), they are driven to reason as follows: I cannot (yet) see how to refute this objection, or overcome this difficulty, but since I cannot imagine how anything other than natural selection could be the cause of the effects, I will have to assume that the objection is spurious; *somehow* natural selection must be sufficient to explain the effects.

Before anyone jumps on this and pronounces that I have just conceded that Darwinism is just as much an unprovable faith as natural religion, it should be borne in mind that there is a fundamental difference: having declared their allegiance to natural selection, these scientists have then proceeded to take on the burden of showing how the difficulties with their view could be overcome, and, time and time again, they have succeeded in meeting the challenge. In the process, Darwin's fundamental idea of natural selection has been articulated, expanded, clarified, quantified, and deepened in many ways, becoming stronger every time it overcame a challenge. With every success, the scientists' conviction grows that they must be on the right track. It is reasonable to believe that an idea that was ultimately false would surely have succumbed by now to such an unremitting campaign of attacks. That is not a conclusive proof, of course, just a mighty persuasive consideration. One of the goals of this book is to explain why the idea of natural selection appears to be a clear winner, even while there are unresolved controversies about how it can handle some phenomena.

4. NATURAL SELECTION AS AN ALGORITHMIC PROCESS

What limit can be put to this power, acting during long ages and rigidly scrutinising the whole constitution, structure, and habits of each creature,—favouring the good and rejecting the bad? I can see no limit to this power, in slowly and beautifully adapting each form to the most complex relations of life.

—CHARLES DARWIN, *Origin*, p. 469

The second point to notice in Darwin's summary is that he presents his principle as deducible by a formal argument—*if* the conditions are met, a certain outcome is *assured*.[6] Here is the summary again, with some key terms in boldface.

> **If**, during the long course of ages and under varying conditions of life, organic beings vary at all in the several parts of their organization, and I think this cannot be disputed; **if** there be, owing to the high geometric powers of increase of each species, at some age, season, or year, a severe struggle for life, and this certainly cannot be disputed; **then,** considering the infinite complexity of the relations of all organic beings to each other and to their conditions of existence, causing an infinite diversity in structure, constitution, and habits, to be advantageous to them, **I think it would be a most extraordinary fact if no variation ever had occurred useful to each being's own welfare,** in the same way as so many variations have occurred useful to man. But **if** variations useful to any organic being do occur, **assuredly** individuals thus characterized will have the best chance of being preserved in the struggle for life; and from the strong principle of inheritance they will tend to produce offspring similarly characterized. This principle of preservation, I have called, for the sake of brevity, Natural Selection. [*Origin,* p. 127 (facs. ed. of 1st ed.).]

The basic deductive argument is short and sweet, but Darwin himself described *Origin of Species* as "one long argument." That is because it

6. The ideal of a deductive (or "nomologico-deductive") science, modeled on Newtonian or Galilean physics, was quite standard until fairly recently in the philosophy of science, so it is not surprising that much effort has been devoted to devising and criticizing various axiomatizations of Darwin's theory—since it was presumed that in such a formalization lay scientific vindication. The idea, introduced in this section, that Darwin should be seen, rather, as postulating that evolution is an algorithmic process, permits us to do justice to the undeniable *a priori* flavor of Darwin's thinking without forcing it into the Procrustean (and obsolete) bed of the nomologico-deductive model. See Sober 1984a and Kitcher 1985a.

consists of two sorts of demonstrations: the logical demonstration that a certain *sort* of process would necessarily have a certain sort of outcome, and the empirical demonstration that the requisite conditions for that sort of process had in fact been met in nature. He bolsters up his logical demonstration with thought experiments—"imaginary instances" (*Origin*, p. 95)—that show *how* the meeting of these conditions *might* actually account for the effects he claimed to be explaining, but his whole argument extends to book length because he presents a wealth of hard-won empirical detail to convince the reader that these conditions have been met over and over again.

Stephen Jay Gould (1985) gives us a fine glimpse of the importance of this feature of Darwin's argument in an anecdote about Patrick Matthew, a Scottish naturalist who as a matter of curious historical fact had scooped Darwin's account of natural selection by many years—in an appendix to his 1831 book, *Naval Timber and Arboriculture*. In the wake of Darwin's ascent to fame, Matthew published a letter (in *Gardeners' Chronicle*![7]) proclaiming his priority, which Darwin graciously conceded, excusing his ignorance by noting the obscurity of Matthew's choice of venue. Responding to Darwin's published apology, Matthew wrote:

> To me the conception of this law of Nature came intuitively as a self-evident fact, almost without an effort of concentrated thought. Mr. Darwin here seems to have more merit in the discovery than I have had—to me it did not appear a discovery. He seems to have worked it out by inductive reason, slowly and with due caution to have made his way synthetically from fact to fact onwards; while with me it was by a general glance at the scheme of Nature that I estimated this select production of species as an a priori recognizable fact—an axiom, requiring only to be pointed out to be admitted by unprejudiced minds of sufficient grasp. [Quoted in Gould 1985, pp. 345–46.]

Unprejudiced minds may well resist a new idea out of sound conservatism, however. Deductive arguments are notoriously treacherous; what seems to "stand to reason" can be betrayed by an overlooked detail. Darwin appreciated that only a relentlessly detailed survey of the evidence for the historical processes he was postulating would—or should—persuade scientists to abandon their traditional convictions and take on his revolutionary vision, even if it was in fact "deducible from first principles."

7. *Gardeners' Chronicle*, April 7, 1860. See Hardin 1964 for more details.

From the outset, there were those who viewed Darwin's novel mixture of detailed naturalism and abstract reasoning about processes as a dubious and inviable hybrid. It had a tremendous air of plausibility, but so do many get-rich-quick schemes that turn out to be empty tricks. Compare it to the following stock-market principle: Buy Low, Sell High. This is guaranteed to make you wealthy. You cannot fail to get rich *if* you follow this advice. Why doesn't it work? It does work—for everybody who is fortunate enough to act according to it, but, alas, there is no way of determining that the conditions are met until it is too late to act on them. Darwin was offering a skeptical world what we might call a get-rich-*slow* scheme, a scheme for creating Design out of Chaos without the aid of Mind.

The theoretical power of Darwin's abstract scheme was due to several features that Darwin quite firmly identified, and appreciated better than many of his supporters, but lacked the terminology to describe explicitly. Today we could capture these features under a single term. Darwin had discovered the power of an *algorithm*. An algorithm is a certain sort of formal process that can be counted on—logically—to yield a certain sort of result whenever it is "run" or instantiated. Algorithms are not new, and were not new in Darwin's day. Many familiar arithmetic procedures, such as long division or balancing your checkbook, are algorithms, and so are the decision procedures for playing perfect tic-tac-toe, and for putting a list of words into alphabetical order. What is relatively new—permitting us valuable hindsight on Darwin's discovery—is the theoretical reflection by mathematicians and logicians on the nature and power of algorithms in general, a twentieth-century development which led to the birth of the computer, which has led in turn, of course, to a much deeper and more lively understanding of the powers of algorithms in general.

The term *algorithm* descends, via Latin (*algorismus*) to early English (*algorisme* and, mistakenly therefrom, *algorithm*), from the name of a Persian mathematician, Mûusâ al-Khowârizm, whose book on arithmetical procedures, written about 835 A.D., was translated into Latin in the twelfth century by Adelard of Bath or Robert of Chester. The idea that an algorithm is a foolproof and somehow "mechanical" procedure has been present for centuries, but it was the pioneering work of Alan Turing, Kurt Gödel, and Alonzo Church in the 1930s that more or less fixed our current understanding of the term. Three key features of algorithms will be important to us, and each is somewhat difficult to define. Each, moreover, has given rise to confusions (and anxieties) that continue to beset our thinking about Darwin's revolutionary discovery, so we will have to revisit and reconsider these introductory characterizations several times before we are through:

(1) *substrate neutrality:* The procedure for long division works equally well with pencil or pen, paper or parchment, neon lights or skywrit-

ing, using any symbol system you like. The power of the procedure is due to its *logical* structure, not the causal powers of the materials used in the instantiation, just so long as those causal powers permit the prescribed steps to be followed exactly.

(2) *underlying mindlessness:* Although the overall design of the procedure may be brilliant, or yield brilliant results, each constituent step, as well as the transition between steps, is utterly simple. How simple? Simple enough for a dutiful idiot to perform—or for a straightforward mechanical device to perform. The standard textbook analogy notes that algorithms are *recipes* of sorts, designed to be followed by *novice* cooks. A recipe book written for great chefs might include the phrase "Poach the fish in a suitable wine until almost done," but an algorithm for the same process might begin, "Choose a white wine that says 'dry' on the label; take a corkscrew and open the bottle; pour an inch of wine in the bottom of a pan; turn the burner under the pan on high; . . . "—a tedious breakdown of the process into dead-simple steps, requiring no wise decisions or delicate judgments or intuitions on the part of the recipe-reader.

(3) *guaranteed results:* Whatever it is that an algorithm does, it always does it, if it is executed without misstep. An algorithm is a foolproof recipe.

It is easy to see how these features made the computer possible. *Every computer program is an algorithm*, ultimately composed of simple steps that can be executed with stupendous reliability by one simple mechanism or another. Electronic circuits are the usual choice, but the power of computers owes nothing (save speed) to the causal peculiarities of electrons darting about on silicon chips. The very same algorithms can be performed (even faster) by devices shunting photons in glass fibers, or (much, much slower) by teams of people using paper and pencil. And as we shall see, the capacity of computers to run algorithms with tremendous speed and reliability is now permitting theoreticians to explore Darwin's dangerous idea in ways heretofore impossible, with fascinating results.

What Darwin discovered was not really *one* algorithm but, rather, a large class of related algorithms that he had no clear way to distinguish. We can now reformulate his fundamental idea as follows:

> Life on Earth has been generated over billions of years in a single branching tree—the Tree of Life—by one algorithmic process or another.

What this claim means will become clear gradually, as we sort through the various ways people have tried to express it. In some versions it is utterly vacuous and uninformative; in others it is manifestly false. In between lie the versions that really do explain the origin of species and promise to explain much else besides. These versions are becoming clearer all the time, thanks as much to the determined criticisms of those who frankly hate the idea of evolution as an algorithm, as to the rebuttals of those who love it.

5. PROCESSES AS ALGORITHMS

When theorists think of algorithms, they often have in mind kinds of algorithms with properties that are *not* shared by the algorithms that will concern us. When mathematicians think about algorithms, for instance, they usually have in mind algorithms that can be proven to compute particular mathematical functions of interest to them. (Long division is a homely example. A procedure for breaking down a huge number into its prime factors is one that attracts attention in the exotic world of cryptography.) But the algorithms that will concern us have nothing particular to do with the number system or other mathematical objects; they are algorithms for sorting, winnowing, and building things.[8]

Because most mathematical discussions of algorithms focus on their guaranteed or mathematically provable powers, people sometimes make the elementary mistake of thinking that a process that makes use of chance or randomness is not an algorithm. But even long division makes good use of randomness!

$$\begin{array}{r} 7? \\ 47\overline{)326574} \end{array}$$

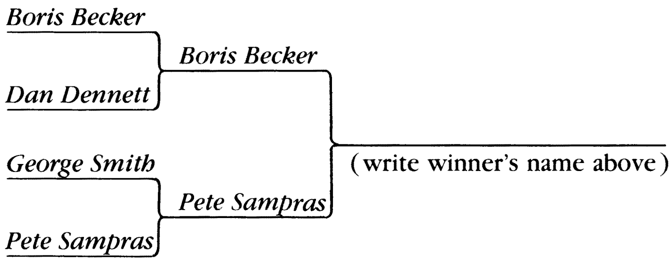
Does the divisor go into the dividend six or seven or eight times? Who knows? Who cares? You don't have to know; you don't have to have any wit or discernment to do long division. The algorithm directs you just to choose a digit—at random, if you like—and check out the result. If the chosen number turns out to be too small, increase it by one and start over; if too large, decrease it. The good thing about long division is that it always works

---

8. Computer scientists sometimes restrict the term *algorithm* to programs that can be proven to *terminate*—that have no infinite loops in them, for instance. But this special sense, valuable as it is for some mathematical purposes, is not of much use to us. Indeed, few of the computer programs in daily use around the world would qualify as algorithms in this restricted sense; most are designed to cycle indefinitely, patiently waiting for instructions (including the instruction to terminate, without which they keep on going). Their subroutines, however, are algorithms in this strict sense—except where undetected "bugs" lurk that can cause the program to "hang."

eventually, even if you are maximally stupid in making your first choice, in which case it just takes a little longer. Achieving success on hard tasks in spite of utter stupidity is what makes computers seem magical—how could something as mindless as a machine do something as smart as that? Not surprisingly, then, the tactic of finessing ignorance by randomly generating a candidate and then testing it out mechanically is a ubiquitous feature of interesting algorithms. Not only does it not interfere with their provable powers as algorithms; it is often the key to their power. (See Dennett 1984, pp. 149–52, on the particularly interesting powers of Michael Rabin's random algorithms.)

We can begin zeroing in on the phylum of evolutionary algorithms by considering everyday algorithms that share important properties with them. Darwin draws our attention to repeated waves of competition and selection, so consider the standard algorithm for organizing an elimination tournament, such as a tennis tournament, which eventually culminates with quarter-finals, semi-finals, and then a final, determining the solitary winner.



Notice that this procedure meets the three conditions. It is the same procedure whether drawn in chalk on a blackboard, or updated in a computer file, or—a weird possibility—not written down anywhere, but simply enforced by building a huge fan of fenced-off tennis courts each with two entrance gates and a single exit gate leading the winner to the court where the next match is to be played. (The losers are shot and buried where they fall.) It doesn't take a genius to march the contestants through the drill, filling in the blanks at the end of each match (or identifying and shooting the losers). And it always works.

But what, exactly, does this algorithm *do*? It takes as input a set of competitors and guarantees to terminate by identifying a single winner. But what is a winner? It all depends on the competition. Suppose the tournament in question is not tennis but coin-tossing. One player tosses and the other calls; the winner advances. The winner of this tournament will be that single player who has won *n* consecutive coin-tosses without a loss, depending on how many rounds it takes to complete the tournament.

There is something strange and trivial about this tournament, but what is it? The winner does have a rather remarkable property. How often have you ever met anyone who just won, say, ten consecutive coin-tosses without a loss? Probably never. The odds against there being such a person might seem enormous, and in the normal course of events, they surely are. If some gambler offered you ten-to-one odds that he could produce someone who before your very eyes would proceed to win ten consecutive coin-tosses using a fair coin, you might be inclined to think this a good bet. If so, you had better hope the gambler doesn't have 1,024 accomplices (they don't have to cheat—they play fair and square). For that is all it takes (2^{10} competitors) to form a ten-round tournament. The gambler wouldn't have a clue, as the tournament started, which person would end up being the exhibit A that would guarantee his winning the wager, but the tournament algorithm is sure to produce such a person in short order—it is a sucker bet with a surefire win for the gambler. (I am not responsible for any injuries you may sustain if you attempt to get rich by putting this tidbit of practical philosophy into use.)

Any elimination tournament produces a winner, who "automatically" has whatever property was required to advance through the rounds, but, as the coin-tossing tournament demonstrates, the property in question *may* be "merely historical"—a trivial fact about the competitor's past history that has no bearing at all on his or her future prospects. Suppose, for instance, the United Nations were to decide that all future international conflicts would be settled by a coin-toss to which each nation sends a representative (if more than one nation is involved, it will have to be some sort of tournament—it might be a "round robin," which is a different algorithm). Whom should we designate as our national representative? The best coin-toss caller in the land, obviously. Suppose we organized every man, woman, and child in the U.S.A. into a giant elimination tournament. Somebody would have to win, and that person would have just won twenty-eight consecutive coin-tosses without a loss! This would be an irrefutable historical fact about that person, but since calling a coin-toss is just a matter of luck, there is absolutely no reason to believe that the winner of such a tournament would do any better in international competition than somebody else who lost in an earlier round of the tournament. Chance has no memory. A person who holds the winning lottery ticket has certainly *been* lucky, and, thanks to the millions she has just won, she may never need to be lucky again—which is just as well, since there is no reason to think she is more likely than anyone else to win the lottery a second time, or to win the next coin-toss she calls. (Failing to appreciate the fact that chance has no memory is known as the Gambler's Fallacy; it is surprisingly popular—so popular that I should probably stress that it *is* a fallacy, beyond any doubt or controversy.)

In contrast to tournaments of pure luck, like the coin-toss tournament, there are tournaments of skill, like tennis tournaments. Here there *is* reason to believe that the players in the later rounds would do better *again* if they played the players who lost in the early rounds. There is reason to believe—but no guarantee—that the winner of such a tournament is the best player of them all, not just today but tomorrow. Yet, though any well-run tournament is guaranteed to produce a winner, there is no guarantee that a tournament of skill will identify the best player as the winner in any nontrivial sense. That's why we sometimes say, in the opening ceremonies, "May the best man win!"—because it is not guaranteed by the procedure. The best player—the one who is best by "engineering" standards (has the most reliable backhand, fastest serve, most stamina, etc.)—may have an off day, or sprain his ankle, or get hit by lightning. Then, trivially, he may be bested in competition by a player who is not really as good as he is. But nobody would bother organizing or entering tournaments of skill if it weren't the case that *in the long run*, tournaments of skill are won by the best players. *That* is guaranteed by the very definition of a fair tournament of skill; if there were no probability greater than half that the better players would win each round, it would be a tournament of luck, not of skill.

Skill and luck intermingle naturally and inevitably in any real competition, but their ratios may vary widely. A tennis tournament played on very bumpy courts would raise the luck ratio, as would an innovation in which the players were required to play Russian roulette with a loaded revolver before continuing after the first set. But even in such a luck-ridden contest, more of the better players would *tend*, statistically, to get to the late rounds. The power of a tournament to "discriminate" skill differences in the long run may be diminished by haphazard catastrophe, but it is not in general reduced to zero. This fact, which is as true of evolutionary algorithms in nature as of elimination tournaments in sports, is sometimes overlooked by commentators on evolution.

Skill, in contrast to luck, is *projectable*; in the same or similar circumstances, it can be counted on to give repeat performances. This relativity to circumstances shows us another way in which a tournament might be weird. What if the conditions of competition kept changing (like the croquet game in *Alice in Wonderland*)? If you play tennis the first round, chess in the second round, golf in the third round, and billiards in the fourth round, there is no reason to suppose the eventual winner will be particularly good, compared with the whole field, in *any* of these endeavors—all the good golfers may lose in the chess round and never get a chance to demonstrate their prowess, and even if luck plays no role in the fourth-round billiards final, the winner might turn out to be the second-*worst* billiards player in the whole field. Thus there has to be some measure of uniformity of the conditions of competition for there to be any *interesting* outcome to a tournament.

But does a tournament—or any algorithm—have to do something interesting? No. The algorithms we tend to talk about almost always do something interesting—that's why they attract our attention. But a procedure doesn't fail to be an algorithm just because it is of no conceivable use or value to anyone. Consider a variation on the elimination-tournament algorithm in which the *losers* of the semi-finals play in the finals. This is a stupid rule, destroying the *point* of the whole tournament, but the tournament would still be an algorithm. Algorithms don't have to have points or purposes. In addition to all the useful algorithms for alphabetizing lists of words, there are kazillions of algorithms for reliably *mis*alphabetizing words, and they work perfectly every time (as if anyone would care). Just as there is an algorithm (many, actually) for finding the square root of any number, so there are algorithms for finding the square root of any number except 18 or 703. Some algorithms do things so boringly irregular and pointless that there is no succinct way of saying what they are *for*. They just do what they do, and they do it every time.

We can now expose perhaps the most common misunderstanding of Darwinism: the idea that Darwin showed that evolution by natural selection is a procedure *for* producing Us. Ever since Darwin proposed his theory, people have often misguidedly tried to interpret it as showing that we are the destination, the goal, the point of all that winnowing and competition, and our arrival on the scene was guaranteed by the mere holding of the tournament. This confusion has been fostered by evolution's friends and foes alike, and it is parallel to the confusion of the coin-toss tournament winner who basks in the misconsidered glory of the idea that since the tournament had to have a winner, and since he is the winner, the tournament had to produce him as the winner. Evolution can be an algorithm, and evolution can have produced us by an algorithmic process, without its being true that evolution is an algorithm for producing us. The main conclusion of Stephen Jay Gould's *Wonderful Life: The Burgess Shale and the Nature of History* (1989a) is that if we were to "wind the tape of life back" and play it again and again, the likelihood is infinitesimal of *Us* being the product on any other run through the evolutionary mill. This is undoubtedly true (if by "Us" we mean the particular variety of *Homo sapiens* we are: hairless and upright, with five fingers on each of two hands, speaking English and French and playing tennis and chess). Evolution is not a process that was designed to produce us, but it does not follow from this that evolution is not an algorithmic process that has in fact produced us. (Chapter 10 will explore this issue in more detail.)

Evolutionary algorithms are manifestly interesting algorithms—interesting to us, at least—not because what they are guaranteed to do is interesting to us, but because what they are guaranteed to *tend* to do is interesting to us. They are like tournaments of skill in this regard. The power of an algorithm to yield something of interest or value is not at all limited to what the algorithm can be mathematically proven to yield in a foolproof way, and this is especially true of evolutionary algorithms. Most of the controversies about Darwinism, as we shall see, boil down to disagreements about just how powerful certain postulated evolutionary processes are—could they actually do all this or all that in the time available? These are typically investigations into what an evolutionary algorithm *might* produce, or *could* produce, or is *likely* to produce, and only indirectly into what such an algorithm would *inevitably* produce. Darwin himself sets the stage in the wording of his summary: his idea is a claim about what "assuredly" the process of natural selection will "tend" to yield.

All algorithms are guaranteed to do whatever they do, but it need not be anything interesting; some algorithms are further guaranteed to tend (with probability p) to do something—which may or may not be interesting. But if what an algorithm is guaranteed to do doesn't have to be "interesting" in any way, how are we going to distinguish algorithms from other processes? Won't *any* process be an algorithm? Is the surf pounding on the beach an algorithmic process? Is the sun baking the clay of a dried-up riverbed an algorithmic process? The answer is that there may be features of these processes that *are* best appreciated if we consider them as algorithms! Consider, for instance, the question of why the grains of sand on a beach are so uniform in size. This is due to a natural sorting process that occurs thanks to the repetitive launching of the grains by the surf—alphabetical order on a grand scale, you might say. The pattern of cracks that appear in the sun-baked clay may be best explained by looking at chains of events that are not unlike the successive rounds in a tournament.

Or consider the process of annealing a piece of metal to temper it. What could be a more physical, less "computational" process than that? The blacksmith repeatedly heats the metal and then lets it cool, and somehow in the process it becomes much stronger. How? What kind of an explanation can we give for this magical transformation? Does the heat create special toughness atoms that coat the surface? Or does it suck subatomic glue out of the atmosphere that binds all the iron atoms together? No, nothing like that happens. The right level of explanation is the algorithmic level: As the metal cools from its molten state, the solidification starts in many different spots at the same time, creating crystals that grow together until the whole is solid. But the first time this happens, the arrangement of the individual crystal structures is suboptimal—weakly held together, and with lots of internal stresses and strains. Heating it up again—but not all the way to melting—partially breaks down these structures, so that, when they are permitted to cool the next time, the broken-up bits will adhere to the still-solid bits in a different arrangement. It can be proven mathematically that these rearrangements will tend to get better and better, approaching the optimum or strongest total structure, provided the regime of heating and cooling has the right parameters. So powerful is this optimization procedure that it has been used as the inspiration for an entirely general problem-solving technique in computer science—"simulated annealing," which has nothing to do with metals or heat, but is just a way of getting a computer program to build, disassemble, and rebuild a data structure (such as another program), over and over, blindly groping towards a better—indeed, an optimal—version (Kirkpatrick, Gelatt and Vecchi 1983). This was one of the major insights leading to the development of "Boltzmann machines" and "Hopfield nets" and the other constraint-satisfaction schemes that are the basis for the Connectionist or "neural-net" architectures in Artificial Intelligence. (For overviews, see Smolensky 1983, Rumelhart 1989, Churchland and Sejnowski 1992, and, on a philosophical level, Dennett 1987a, Paul Churchland 1989.)

If you want a deep understanding of how annealing works in metallurgy, you have to learn the physics of all the forces operating at the atomic level, of course, but notice that the basic idea of how annealing works (and particularly *why* it *works*) can be lifted clear of those details—after all, I just explained it in simple lay terms (and I don't know the physics!). The explanation of annealing can be put in *substrate-neutral* terminology: we should expect optimization of a certain sort to occur in any "material" that has components that get put together by a certain sort of building process and that can be disassembled in a sequenced way by changing a single global parameter, etc. That is what is common to the processes going on in the glowing steel bar and the humming supercomputer.

Darwin's ideas about the powers of natural selection can also be lifted out of their home base in biology. Indeed, as we have already noted, Darwin himself had few inklings (and what inklings he had turned out to be wrong) about how the microscopic processes of genetic inheritance were accomplished. Not knowing any of the details about the physical substrate, he could nevertheless discern that if certain conditions were somehow met, certain effects would be wrought. This substrate neutrality has been crucial in permitting the basic Darwinian insights to float like a cork on the waves of subsequent research and controversy, for what has happened since Darwin has a curious flip-flop in it. Darwin, as we noted in the preceding chapter, never hit upon the utterly necessary idea of a gene, but along came Mendel's concept to provide just the right structure for making mathematical sense out of heredity (and solving Darwin's nasty problem of blending inheritance). And then, when DNA was identified as the actual physical vehicle of the genes, it looked at first (and still looks to many participants) as if Mendel's genes could be simply *identified* as particular hunks of DNA. But then complexities began to emerge; the more scientists have learned about the actual molecular biology of DNA and its role in reproduction, the clearer it becomes that the Mendelian story is at best a vast oversimplification. Some would go so far as to say that we have recently learned that there really *aren't* any Mendelian genes! Having climbed Mendel's ladder, we must now throw it away. But of course no one wants to throw away such a valuable tool, still proving itself daily in hundreds of scientific and medical contexts. The solution is to bump Mendel up a level, and declare that he, like Darwin, captured an *abstract* truth about inheritance. We may, if we like, talk of *virtual genes*, considering them to have their reality distributed around in the concrete materials of the DNA. (There is much to be said in favor of this option, which I will discuss further in chapters 5 and 12.)

But then, to return to the question raised above, are there any limits at all on what may be considered an algorithmic process? I guess the answer is No; if you wanted to, you could treat any process at the abstract level as an algorithmic process. So what? Only some processes yield interesting results when you do treat them as algorithms, but we don't have to try to define "algorithm" in such a way as to include only the *interesting* ones (a tall philosophical order!). The problem will take care of itself, since nobody will waste time examining the algorithms that aren't interesting for one reason or another. It all depends on what needs explaining. If what strikes you as puzzling is the uniformity of the sand grains or the strength of the blade, an algorithmic explanation is what will satisfy your curiosity—and it will be the truth. Other interesting features of the same phenomena, or the processes that created them, might not yield to an algorithmic treatment.

Here, then, is Darwin's dangerous idea: the algorithmic level *is* the level that best accounts for the speed of the antelope, the wing of the eagle, the shape of the orchid, the diversity of species, and all the other occasions for wonder in the world of nature. It is hard to believe that something as mindless and mechanical as an algorithm could produce such wonderful things. No matter how impressive the products of an algorithm, the underlying process always consists of nothing but a set of individually mindless steps succeeding each other without the help of any intelligent supervision; they are "automatic" by definition: the workings of an automaton. They feed on each other, or on blind chance—coin-flips, if you like—and on nothing else. Most algorithms we are familiar with have rather modest products: they do long division or alphabetize lists or figure out the income of the Average Taxpayer. Fancier algorithms produce the dazzling computer-animated graphics we see every day on television, transforming faces, creating herds of imaginary ice-skating polar bears, simulating whole virtual worlds of entities never seen or imagined before. But the actual biosphere is much fancier still, by many orders of magnitude. Can it really be the outcome of nothing but a cascade of algorithmic processes feeding on chance? And if so, who designed that cascade? Nobody. It is itself the product of a blind, algorithmic process. As Darwin himself put it, in a letter to the geologist Charles Lyell shortly after