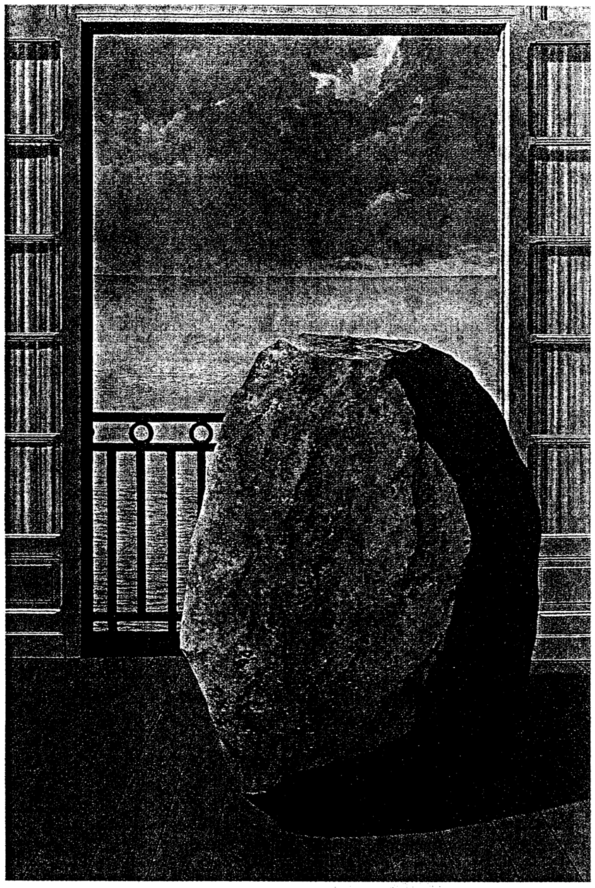
THE UNKNOWN AND THE UNKNOWABLE

Exploring the boundaries of scientific knowledge

BY JOSEPH F. TRAUB



René Magritte, The Invisible World, *1954*

IMPOSSIBILITY: THE LIMITS OF SCIENCE AND THE SCIENCE OF LIMITS
by John D. Barrow
Oxford University Press, 1998
279 pages; $25.00

WHAT REMAINS TO BE DISCOVERED
by John Maddox
Martin Kessler Books, The Free Press, 1998
434 pages; $26.00

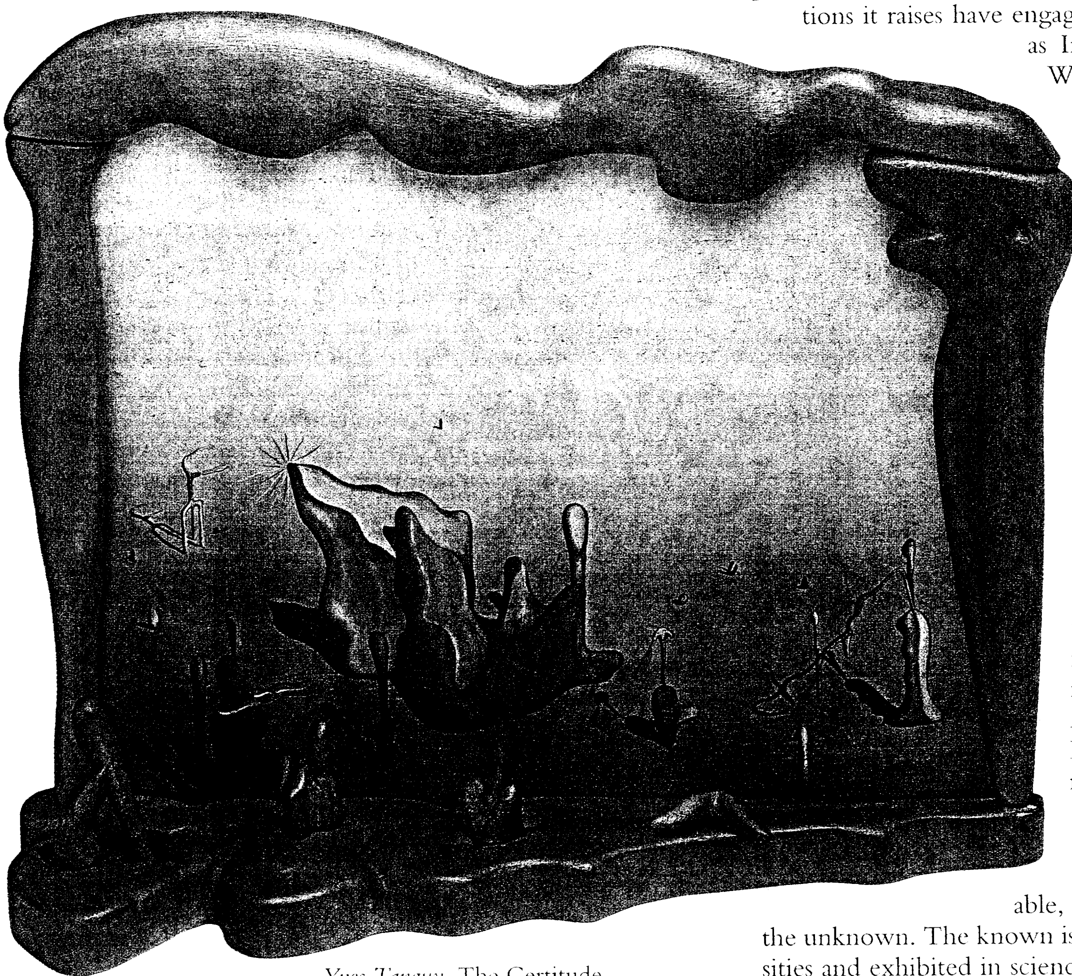
IN THE EARLY 1990S I MENTIONED TO A SENIOR European scientist that I was interested in understanding the intrinsic limits to scientific knowledge. He replied that such limits had been established by Kurt Gödel. That interpretation of Gödel's fundamental contribution to logic seems to be a commonly held belief, but it is simply not so.

It is easy to understand why one might come to think that Gödel's result would have implications for science. In 1931 the Czech logician, then a young professor at the University of Vienna, published a scholarly paper summarizing his doctoral investigations into the properties of formal logical systems. Such systems are made up of some finite number of statements called axioms, which are accepted as true and "self-evident," as well as some equally unobjectionable rules for deriving one statement from another, called rules of inference. Arithmetic furnishes some examples: the axiom of arithmetic that makes, say, 2 + 0 equal to 2 (and likewise, for the addition of zero to any natural number); and the logical rule that enables one to infer, from the proposition *if p then q* and the proposition *p,* that *q.* Gödel showed that if a system as rich as arithmetic is consistent, it cannot be complete. That is, there are statements within arithmetic that cannot be proved true or false within the system; arithmetic is undecidable.

When Gödel announced his result, its impact on logic and mathematics was sensational and profound. It solved, in one brilliant flash, the so-called decision problem of the great nineteenth-century German mathematician David Hilbert: Devise a mechanical procedure that could determine whether or not any given mathematical statement was true. (Gödel's shocking answer was, it could not be done.) Gödel's result also doomed, in one fatal thrust, the fondest hopes of turn-of-the-century logicians: that axiom systems, modeled on Euclid's geometry, could "capture" all the truths of arithmetic. Even more, Gödel's theorem seemed, at least to many people, to impose a fundamental limit on human understanding. For if understanding itself was a matter of reducing a proposition to a theorem provable from certain "self-evident" truths, then what Gödel had done was to plumb one of the intrinsic limits to understanding.

Now what does this have to do with science? It is essential to keep in mind that Gödel's theorem is about the formal manipulation of symbols; that is, it is about mathematics. Science certainly uses mathematics, but science is also very different from mathematics. Science is about understanding the universe and everything in it. Examples of scientific questions abound: How do children acquire language? Will human activities cause major global changes, and what will be the effects of those changes on the ocean levels, on agriculture and on biodiversity? How do physical processes in the brain give rise to subjective experience—that is, how do such processes explain consciousness? Can the healthy, active lives of people be prolonged by, say, a factor of two or three? How did life originate on earth? Will the universe expand forever, or will it collapse? Is there life elsewhere in the universe? Is it intelligent?

Note that, at least on the face of it, there are no mathematical models that formalize the relevant aspects of the world, within which those questions can be asked. Without such models, the threat allegedly posed to science by Gödel's theorem cannot even get off the ground. Yet mathematical models, of course, are indispensable to the physical sciences: Newton's laws of mechanics provide a good example. Is it possible, then, to up the ante from Gödel's analysis of mathematics and prove impossibility results in science? That is, can one



Yves Tanguy, The Certitude of the Never Seen, *1933*

establish the unknowable in science, the bounds of possible scientific knowledge? The two books under review, John Maddox's *What Remains to Be Discovered* and John Barrow's *Impossibility,* probe many of the issues relevant to that question. For the reader, they complement each other splendidly, and, taken together, they offer an intellectual journey to the very heart of the scientific enterprise.

Investigating the limits of science should not be confused with the intent of the 1996 book by the writer John Horgan, *The End of Science.* When Horgan sent me the manuscript for comments, I told him that I totally disagreed with his thesis: that science has made such extraordinary progress that the golden age is over and only the mopping up is left. At that time, incidentally, the manuscript was titled *The Ends of Science,* an ambiguous phrase that, to my mind, is far more interesting. Apparently the publisher settled on *The End of Science,* in hopes of reaping the same success that has attended a slew of "The End of You-Name-It" books.

Horgan writes very well indeed, but I was astonished by the amount of attention his book received in the general media. Its message about science is basically pessimistic, whereas the scientists with whom I am in touch are vitally excited by their work. There is more to be done than ever, and we cannot wait to get on with it.

The investigation of the unknowable has long been the province of philosophy and epistemology, and the questions it raises have engaged such penetrating thinkers as Immanuel Kant and Ludwig Wittgenstein. But what is unknowable to one generation is another generation's mere technical challenge: to Aristotelians in the Middle Ages, the regions beyond the moon were celestial spheres reaching to heaven, as unapproachable as whatever might have come *before* the big bang seems to cosmologists today. A more constructive viewpoint recognizes that what is known or can be known about the world can change with time. It makes sense, then, to think about the unknowable as one (somewhat adjustable) region on a map of what Ralph E. Gomory, the president of the Alfred P. Sloan Foundation in New York City and the former senior vice president of science and technology at IBM, calls the tripartite division of knowledge. To the unknowable, Gomory adds the known and the unknown. The known is taught at schools and universities and exhibited in science museums. The unknown is the frontier, a territory that may someday become known and so is not, in principle, unknowable.

SCIENTISTS ARE FASCINATED BY THE UNKNOWN, but for the most part they have been content to leave the contemplation of the unknowable to philosophers. That may be changing. For one thing, the intellectual climate seems right for tackling such questions. Gödel's success has led to other impossibility results in mathematics, as well as in theoretical computer science. Furthermore, the study of hard problems—problems not strictly impossible in Gödel's sense, but whose difficulty is thought to grow exponentially fast—has led to the highly fruitful idea of classifying problems by their computational complexity. The difficulty of a problem can be measured by the cost of the computer resources—say, the length of run time—needed to solve it by the fastest method possible.

Although those ideas from mathematics and computer science cannot be directly applied to science, the modes of thought might be transferable. For example, suppose you could list all the formal models that capture the essence of a scientific question. If you could prove that all the models are undecidable—that, by analogy with the algorithms for solving a hard problem, none of the models is logically simple enough to be decidable—then the answer to that scientific question would be unknowable.

Second, and perhaps more important, moving the effort of distinguishing between the unknown and the unknowable from philosophy to science could lead to a great enrichment of science. Why should a scientist pay any attention to what cannot be known? Partly, because what is allegedly unknowable may help articulate what is (merely) unknown. Does cosmology really have nothing to say, even in principle, about conditions before the big bang? Or is that a legitimate inquiry into an unknown but knowable aspect of the universe? A clear perspective on the boundaries of the unknowable makes it less likely that people will become discouraged, thereby permitting the kingdom of the unknowable illegitimately to annex parts of the kingdom of the knowable unknown.

Barrow eloquently makes the case for such a perspective:

> It would be easy to write such a scientific success story. But we have another tale to tell: one that tells not of the known but of the unknown; of things impossible; of limits and barriers which cannot be crossed. Perhaps this sounds a little perverse. Surely there is little enough to say about the unknown without dragging in the unknowable? But the impossible is a powerful and persistent notion . . . [though] its positive role has escaped the critics' attention. Our goal is to uncover some of the limits of science: to see how our minds' awareness of the impossible gives a new perspective on reality.

For all those reasons, the promise of gaining real insight by studying the map of what is known, unknown and unknowable in science has given rise to a groundswell of interest. Recent workshops have brought together leading physicists, economists, cognitive scientists, biologists, computer scientists and mathematicians who have a strong interest in identifying the unknowable in their own fields. In the past four years, four workshops have been held at the Santa Fe Institute in New Mexico, sponsored by the Alfred P. Sloan Foundation, and a fifth workshop was held at Abisko, the Arctic research station of the Royal Swedish Academy of Sciences. Barrow, whose book explores the varieties of the unknowable, was a participant at one of the Santa Fe Institute workshops and at Abisko. Maddox's book is best read as a wall chart of the known and the unknown.

As editor in chief of the prestigious scientific journal *Nature* for nearly twenty-three years, Maddox is eminently qualified for the millennial task he has set himself. He begins by imagining how his book might have been written a century ago. The hypothetical nineteenth-century Maddox might have asked: What is space made of? What does energy have to do with matter? What is life? But in 1900 neither he nor anyone else could have foreseen the emergence of quantum mechanics, and it is highly unlikely that anyone would have recognized how important the question of gravity was to be for the science of the twentieth century. Nor could nineteenth-century Maddox have foreseen that DNA molecules made up of just four chemical units would be revealed as the font of the fantastic variety with which the living world abounds.

Maddox's text is divided into three parts: "Matter," "Life" and "Our World." "Matter" explores the origins of the universe and the elementary particles of matter, as well as the prospects for what physicists call a theory of everything. "Life" also turns to origins, as well as to such topics as the role of

IF PEOPLE BECOME DISCOURAGED, ***the kingdom of the unknowable may illegitimately annex parts of the knowable unknown.***

the selfish gene and the human genome project. "Our World" concludes the book with expositions on the nature of consciousness, mathematics, and Maddox's thoughts on avoiding future calamities. In each chapter Maddox surveys what is known and concludes with questions about the unknown.

Maddox energizes that simple outline by regaling the reader with stories. For example, to make the point that adaptation through natural selection can lead to extinction, he relates the tale of the dodo, a large, flightless bird that, along with two related species, once made its home on islands in the Indian Ocean. The ancestors of the dodo could fly: how else could the birds have populated three widely spaced islands? But natural selection later favored individuals that gave up flying, thereby saving huge metabolic costs. How could natural selection have anticipated that a time would come when flight would once again become an advantage—such as when, in the sixteenth century, European sailors began hunting the dodo and its relatives for food?

For Maddox, expanding the scope of scientific knowledge is an immensely practical matter. Converting the unknown into the known offers the best means people have for coping with the uncertainties of the future—and for preparing to avoid its most potent threats. For those of us already concerned about nuclear proliferation, the destruction of biodiversity, overpopulation, the possibility that global climate change will trigger a new ice age, or the arrival of tropical diseases in temperate climates, Maddox gives us some more things to worry about.

One calamity could transpire in the relatively near future, within a century or two. The western Antarctic ice shelf may be unstable and could therefore become detached as a result of global warming. Maddox writes:

> This mass of ice, reaching for several hundreds of kilometers out to sea from the solid continent, . . . supports hundreds of cubic kilometers of ice. If this block of ice should become detached from Antarctica, perhaps because melting ice lubricates its junction with the rock beneath, the whole mass could become the

world's largest iceberg. In that case, there would be an increase of some five meters (sixteen feet) in global sea level. The map of the world would change quickly.

A second potential hazard could lie hundreds of generations in the future: the human genome could become unstable. That is, enough defects could accumulate in the genome that further survival of the species would be threatened. As Maddox points out, there is ample precedent for the process, recorded among the remains of extinct plants and animals. Unlike the Antarctic catastrophe, however, genetic instability, it seems to me, does not pose any insurmountable problems. The human species has enough time to deal with genetic problems as they come along.

Maddox lets readers have his views on certain issues in no uncertain terms. On a number of those points, however, I find him unconvincing. About the fears regarding recent developments in genetics, both within the research community and outside it, he writes dismissively:

> Merely the availability of prenatal diagnosis suggests to some that communities in which the techniques are already available are about to practice eugenics of the kind ignorantly advocated in Germany in the 1930s. Others fear the imminence of the "designer baby"—an individual equipped with superior faculties, both intellectual and physical. Both fears are groundless. There are important ethical questions to ask, but these are not them.

Perhaps Maddox has good reasons for thinking those fears to be groundless. The reader, though, would have been better served if he had provided them. Merely dismissing the fears will not make them go away.

A SECOND MATTER ON WHICH I FIND MADDOX unconvincing is his account of Gödel's theorem and other, related results, and their implications for what computers and mathematicians can and cannot do. One of the strongest impossibility results was proved just five years after Gödel's, by the English polymath Alan M. Turing. Turing devised a pencil-and-paper scheme that incorporated just a few simple rules. But with his scheme, he asserted, one could carry out any mechanical procedure, or algorithm, of the kind envisioned by Hilbert for grinding out all possible true mathematical theorems. The scheme has become known as a Turing machine, and it is abstractly equivalent to any digital computer, no matter how powerful, in the sense that it can perform any computation the latter machine can be programmed to accomplish.

Some procedures on a Turing machine—or, equivalently, on the latest-model PC—lead quickly through a few machine cycles and then halt. "Find the square root of 3 (to five decimal places) and print the answer," is a good example. Other procedures, some written in just a few characters of programming code, will never halt: the standard jocular example comes from the instructions on a bottle of shampoo: "Lather, rinse, repeat." But some computer programs fall into neither category. They do not seem to halt, but there is no clear reason to think they will go on forever, either. Turing proved that, for an arbitrary string of programming code, it is impossible to tell in advance whether the procedures it specifies on the machine will halt or not. In particular, one could never know in advance whether any mechanical procedure for listing all true mathematical theorems would ever come to an end. In Turing's hands, Hilbert's decision problem leads at once to an undecidable question.

Maddox and many others infer that since computers cannot decide whether a mathematical statement is true or false, such questions cannot be relegated to machines. But that inference is not valid. One might as well say that such questions should not be relegated to human mathematicians either. They are equally incapable of determining whether arbitrary statements about, say, arithmetic, are true or false. The question is not a matter of the superiority of human or machine.

Reflection on how Gödel's and Turing's impossibility results might be relevant to human creativity has led to a host of fascinating questions not addressed by Maddox. Many mathematicians, for instance, believe they create by a process of intuition—but just what is intuition? Does it enable mathematicians to spot truths that cannot be reached through dogged, axiomatic argument? Can one create a mathematical model of intuition, one that captures the essence of the human mind? And how do such questions tie into the quest for a scientific exploration of consciousness? The very fact that an analysis of the impossible profoundly enriches such questions is compelling evidence that science has much to gain from this line of thought.

IF THE ANTARCTIC *ice shelf breaks away, major cities will be inundated.*

THAT POTENTIAL FOR THE ENRICHMENT OF science is what Barrow has set out to explore in *Impossibility*. A professor of astronomy at the University of Sussex in England, Barrow writes with great scholarship, elegance and wit. He classifies the limits to science according to their source: some limits are imposed by the nature of the brain, some by cosmology, some are technological, and some arise from Gödel's theorem and from computational theory.

Concerning the brain, he writes:

> Our minds were not designed with science in mind, nor did evolution primarily fit them for that purpose. . . . On the face of it, there is no reason why we should possess the conceptual ability to make sense of the way the Universe works. It would require a coincidence of cosmic proportions if the Universe were complicated enough to give rise to life, yet simple enough for one species to understand its deepest structure after just a few hundred years of serious scientific investigation.

It could be, for instance, that we humans believe Occam's razor, which holds that the best theories are the simplest, because those are the only theories our puny brains can comprehend. Yet I think it possible that one of humankind's greatest inventions, the computer, could be a tool for transcending the puny-brain limitation. Needless to say, such a view is highly speculative.

What of Barrow's cosmological and technological limits? They are relatively practical constraints, the kinds of limits that are imposed by the immense universe in which we happen to live, and by the trifling resources we are ever likely to command. In Barrow's view many of the great questions about the beginning, the end and the structure of the universe are unanswerable in that practical sense. Because the speed of light is finite, for instance, the universe is partitioned into realms that are out of causal contact with one another. Furthermore, if the theory of the inflationary universe is valid, information about the structure of the visible universe before inflation took place has been wiped clean—forever beyond scientific ken.

Technology will also determine the limits to what we can eventually discover. Barrow points out that scientists can make precise predictions about the universe under conditions that we cannot even remotely approach via direct experiment. What really takes place, for instance, when matter is heated to temperatures more than fifteen orders of magnitude greater than any achieved on earth? But such limits are fluid. New technology is unpredictable, and there is always the possibility that it will enable some earlier limit to be sidestepped (in this case, through astronomical observations).

Barrow offers an extended and subtle discussion of the possible implications of Gödel's theorem for the limits to science. As I pointed out earlier, Gödel showed that if a formal system is as rich as arithmetic, it must be incomplete. But, as Barrow notes, simply removing the multiplication operation from arithmetic gives a smaller system known as Presburger arithmetic. Gödel showed that Presburger arithmetic is complete. It is possible, Barrow observes, that the "inner logic of the physical universe" is rooted in a simpler logic than full arithmetic. In that case, Gödelian incompleteness might not apply.

The same conclusion holds for the conjectures and theories of computational complexity, and their potential for defining the limits of science. Although I greatly admire Barrow's book—and here, I must immediately confess that I agreed to write an endorsement for the book after reading it in manuscript form—I am not entirely happy with his discussion of computational complexity.

René Magritte, Clairvoyance, *1936*

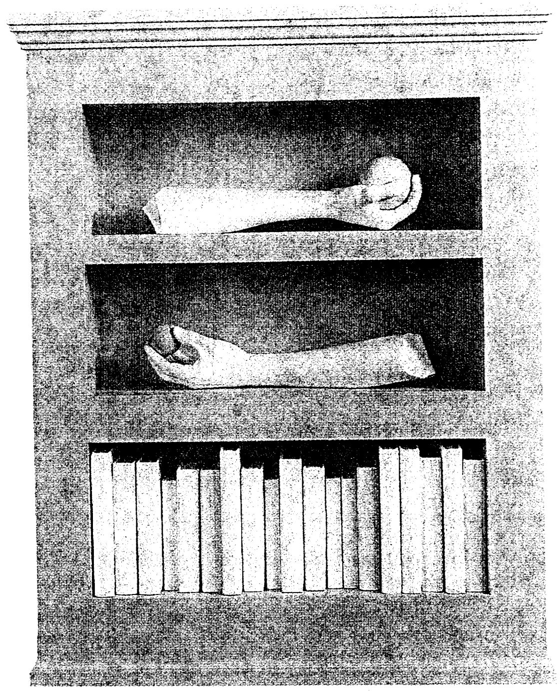
THE COMPUTATIONAL COMPLEXITY OF A PROBlem is a measure of the minimum amount of computational resources needed to solve it. Even if a mathematical problem is decidable, it might be computationally intractable. Barrow describes the traveling salesman problem (TSP) to illustrate the idea. Suppose a salesman must make calls in some set of cities. How does he plan his itinerary so that he visits each of them just once, minimizes the total distance he travels, and returns at the end of his journey to the city where he started? It turns out that as the number of cities on the list grows, the problem becomes increasingly difficult; the number of route combinations grows exponentially with the number of cities in the salesman's tour. No solution more efficient than simply enumerating all the possible combinations of travel routes is known (though no one has proved such a solution does not exist). Hence, though the computational complexity of TSP is unknown, it is thought to grow exponentially with the number of cities. TSP is said to be computationally intractable.

Computer scientists have made a remarkable discovery about TSP. It turns out that numerous problems, some of them of great practical importance, all possess essentially the same computational complexity as TSP. They are all tractable or all intractable, and the common belief among experts is that they are all intractable. Barrow writes that in practice it is extremely hard to prove that any given problem is intractable. "At present," he remarks, to substantiate his point, "there are no more than about a thousand problems" in the same complexity class as TSP. In my view, Barrow has it backward. I find it surprising that so many problems have been proven to have the same computational complexity as TSP.

Another shortcoming of Barrow's account of computational complexity is that it is confined to combinatorial problems, which involve discrete variables (the number of cities in the salesman's tour) rather than continuous variables. As I noted, their complexity is unknown; one must settle instead for statements such as "problem *a* is equivalent to problem *b*," or "it is conjectured that problem *x* is more complex than problem *y*." Mathematical models in the sciences, however, often involve continuous functions of real variables. Differential equations involving distance and time provide a typical example. The computational complexity of the numerical solution of such problems is studied in the field of information-based complexity, and it is often known. There is no need to settle for conjectures.

WHAT, THEN, IS THE CONNECTION between the conjectures and results of computational complexity theory, and science? The protein-folding problem of molecular biochemistry provides an enlightening example. The problem is easy to state: Given a linear sequence of amino acids, into what three-dimensional configuration will the sequence fold? The implications of the answer for biology are manifold, because in protein chemistry the function of a molecule is almost entirely determined by its shape. Thus the answer to the protein-folding problem has been called the Holy Grail of molecular biology.

In nature the folding takes place quickly—in about a second. Yet the known models of the process are so complex that they cannot be simulated even on the most powerful supercomputer. Aviezri S. Fraenkel, a mathematician at the Weizmann Institute of Science in Rehovot, Israel, has shown that one formulation of the problem is just as complex as TSP. Hence protein folding "should" become exponentially harder as the length of the chain of amino acids grows. Nature should not take a second; it should take perhaps millions of years.


Richard E. Prince, Venus and Mars, *1998*

How can such dissonances be explained? The possibilities recall the relation between Gödel's theorem and scientific limits. For example, there may be other mathematical models of protein folding that fit the observed behavior, yet are computationally tractable. Or the conjecture that TSP is intractable may turn out to be false.

Perhaps evolution has selected for proteins that fold easily. Alternatively, nature might be undertaking a massive number of parallel computations, though the mechanisms one might propose for such activity remain highly speculative.

What emerges from such an analysis is a possible experiment: How does the time that nature takes to fold a protein depend on the length of the sequence of amino acids? The time dependence is not exponential, but is it superlinear, sublinear or even constant—that is, independent of the length of the amino acid strand? Experimental measurements might help to construct a predictive theory.

SO REFLECT JUST A MOMENT ABOUT THE NEW LOGic. It shows that there are dissonances between complexity theory, supercomputer simulation and nature. Those dissonances led to a proposal for an experiment that could help constrain the search for improved models: in short, to normal science. Perhaps that is a small success, or just the potential for a small success. But I think it begins to make the case that, from time to time, scientists can step back from their hand-to-hand struggles within their disciplines and profitably explore a bigger picture. The parsing of the unknown, the intractable, the unknowable, so appealing to the popular imagination, and so entertainingly and persuasively carried out in these two books, is a worthy mission for science itself. ●

JOSEPH F. TRAUB is the Edwin Howard Armstrong professor of computer science at Columbia University. Cambridge University Press has just published his most recent book, COMPLEXITY AND INFORMATION.