

## What Brings a World into Being?

David Berlinski

SINCE THEIR inception in the 17th century, the modern sciences have been given over to a majestic vision: there is nothing in nature but atoms and the void. This is hardly a new thought, of course; in the ancient world, it received its most memorable expression in Lucretius' *On the Nature of Things*. But it has been given contemporary resonance in theories—like general relativity and quantum mechanics—of terrifying (and inexplicable) power. If brought to a successful conclusion, the trajectory of this search would yield a single theory that would subsume all other theories and, in its scope and purity, would be our only necessary intellectual edifice.

In science, as in politics, the imperial destiny drives hard. If the effort to subordinate all aspects of experience to a single set of *laws* has often proved inconclusive, the scientific enterprise has also been involved in the search for universal *ideas*. One such idea is information.

Like energy, indeed, information has become ubiquitous as a commodity and, like energy, inescapable as an idea. The thesis that the human mind is nothing more than an information-processing device is now widely regarded as a fact. "Viewed at the most abstract level," the science writer George Johnson remarked recently in the *New York Times*, "both brains

and computers operate *the same way* by translating phenomena—sounds, images, and so forth—into a code that can be stored and manipulated" (emphasis added). More generally, the evolutionary biologist Richard Dawkins has argued that life is itself fundamentally a river of information, an idea that has in large part also motivated the successful effort to decipher the human genome. Information is even said to encompass the elementary particles. "All the quarks and electrons in the cosmic wilds," Johnson writes, "are exchanging information each time they interact."

These assertions convey a current of intellectual optimism that it would be foolish to dismiss. Surely an idea capable of engaging so many distinct experiences must be immensely attractive. But it seems only yesterday that other compelling ideas urged their claims: chaos and nonlinear dynamics, catastrophe theory, game theory, evolutionary entropy, and various notions of complexity and self-organization.

The history of science resembles a collection of ghosts remembering that once they too were gods. With respect to information, a note of caution may well be in order if only because a note of caution is always in order.

### II

IF INFORMATION casts a cold white light on the workings of the mind in general, it should certainly shed a little on the workings of language in particular.

DAVID BERLINSKI is the author of *A Tour of the Calculus*, *The Advent of the Algorithm*, and, most recently, *Newton's Gift* (Free Press). His articles in COMMENTARY include "Was There a Big Bang?" (February 1998) and "The Deniable Darwin" (June 1996).

The words and sentences of Herman Melville's *Moby-Dick*, to take a suggestive example, have the power to bring a world into being. The beginning of the process is in plain sight. There are words on the printed page, and they make up a discrete, one-dimensional, linear progression. Discrete—there are no words between words (as there are fractions between fractions); one-dimensional—each word might well be specified by a single number; linear—as far as words go, it is one thing after another. The end of the process is in sight as well: a richly organized, continuous, three- (or four-) dimensional universe. Although that universe is imaginary, it is recognizably contiguous to our own.

Bringing a world into being is an act of creation. But bringing a world into being is also an activity that suggests, from the point of view of the sciences, that immemorial progression in which causes evoke various effects: connections achieved between material objects, or between the grand mathematical abstractions necessary to explain their behavior.

And therein lies a problem. Words are, indeed, material objects, or linked as abstractions to material objects. And as material objects, they have an inherent power to influence other material objects. But no informal account of what words do as material objects seems quite sufficient to explain what they do in provoking certain experiences and so in creating certain worlds.

In the case of *Moby-Dick*, the chemical composition of words on the printed page, their refractive index, their weight, their mass, and ultimately their nature as a swarm of elementary particles—all this surely plays *some* role in getting the reader sympathetically to see Captain Ahab and imaginatively suffer his fate. The relevant causal pathways pass from the printed ink to our eyes, a river of light then serving to staple the shape of various words to our tingling retinal nerves; thereafter our nervous system obligingly passes on those shapes in the form of various complicated electrical signals. This is completely a physical process, one that begins with physical causes and ends with physical effects.

And yet the experience of reading begins where those physical effects end. It is, after all, an experience, and the world that it reveals is imaginary. If purely physical causes are capable of creating imaginary worlds, it is not by means of any modality known to the physical sciences.

Just how *do* one set of discrete objects, subject to the constraints of a single dimension, give rise to a universe organized in completely different ways and according to completely different principles?

IT IS here that information makes its entrance. The human brain, the linguist Steven Pinker has argued in *How the Mind Works*, is a physical object existing among other physical objects. Ordinary causes in the world at large evoke their ordinary effects within the brain's complicated folds and creases. But the brain is, *also*, an information-processing device, an instrument designed by evolution for higher things.

It is the brain's capacity to process information that, writes Pinker, allows human beings to "see, think, feel, choose, and act." Reading is a special case of seeing, one in which information radiates from the printed page and thereafter transforms itself variously into various worlds.

So much for what information does—clearly, almost everything of interest. But what is it, and how does it manage to do what it does? Pinker's definition, although informal, is brisk and to the point. Information, he writes, "is a correlation between two things that is produced by a lawful process." Circles in a tree stump carry information about the tree's age; lines in the human face carry information about the injuries of time.

Words on the page also contain or express information, and as carriers of information they convey the stuff from one place to another, piggy-backed, as it were, on a stream of physical causes and their effects.

Why not? The digital computer is a device that brilliantly compels a variety of discrete artifacts to scuttle along various causal pathways, ultimately exploiting pulsed signals in order to get one thing to act upon another. But in addition to their physical properties, the symbols flawlessly manipulated by a digital computer are capable of carrying and so conveying information, transforming one information-rich stream, such as a data base of proper names, into another information-rich stream, such as those same names arranged in alphabetical order.

The human mind does as much, Pinker argues; indeed, what it does, it does in the same way. Just as the computer transforms one information stream into another, the human mind transforms one source of information—words on the printed page—into another—a world in which whalers pursue whales and the fog lowers itself ominously over the spreading sea.

Thus Pinker; thus almost everyone.

THE THEORY that gives the concept of information almost all of its content was created by the late mathematician Claude Shannon in 1948 and 1949. In it, the rich variety of human intercourse dwindles and disappears, replaced by an ide-

alized system in which an information *source* sends signals to an information *sink* by means of a communication channel (such as a telephone line).

Communication, Shannon realized, gains traction on the real world by means of the firing pistons of tension and release. From far away, where the system has its source, messages are selected and then sent, one after the other—perhaps by means of binary digits. In the simplest possible set-up, symbols are limited to a single digit: 1, say. A binary digit may occupy one of two states (*on* or *off*). We who are tensed at the system's sink are uncertain whether 1 will erupt into phosphorescent life or the screen will remain blank. Let us assume that each outcome is equally likely. The signal is sent—and then received. Uncertainty collapses into blessed relief, the binary digit 1 emerging in a swarm of pixels. The exercise has conveyed one unit, or bit, of information. And with the definition of a unit in place, information has been added to the list of properties that are interesting because they are *measurable*.

The development of Shannon's theory proceeds toward certain deep theorems about coding channels, noise, and error-reduction. But the details pertinent to this discussion proceed in another direction altogether, where they promptly encounter a roadblock.

"Frequently," Shannon observes, "messages have *meaning*: that is, they refer to or are correlated according to some system with certain physical or conceptual entities." Indeed. Witness *Moby-Dick*, which is *about* a large white whale. For Shannon, however, these "semantic aspects of communication" take place in some other room, not the one where his theory holds court; they are, he writes with a touch of asperity, "irrelevant to the engineering problem." The significance of communication lies only with the fact that "the actual message is one selected from a set of possible messages"—*this* signal, and not some other.

Shannon's strictures are crucial. They have, however, frequently proved difficult to grasp. Thus, in explaining Shannon's theory, Richard Dawkins writes that "the sentence, 'It rained in Oxford every day this week,' carries relatively little information, because the receiver is not surprised by it. On the other hand, 'It rained in the Sahara every day this week' would be a message with high information content, well worth paying extra to send."

But this is to confuse a signal with what it signifies. Whether I am surprised by the sentence "It rained in the Sahara desert every day this week" depends only on my assessment of the source sending

the signal. Shannon's theory makes no judgments whatsoever about the subjects treated by various signals and so establishes no connection whatsoever to events in the real world. It is entirely possible that whatever the weather in Oxford *or* the Sahara may be, a given source might send both sentences with equal probability. In that case, they would convey precisely the same information.

The roadblock now comes into view. Under ordinary circumstances, reading serves the end of placing one man's thoughts in contact with another man's mind. On being told that whales are not fish, Melville's readers have learned something *about* whales and so about fish. Their uncertainty, and so their intellectual tension, has its antecedent roots in facts about the world beyond the symbols they habitually encounter. For most English speakers, the Japanese translation of *Moby-Dick*, although conveying precisely the same information as the English version, remains unreadable and thus unavailing as a guide to the universe created by the book in English.

What we who have conceived an interest in reading have required is some idea of how the words and sentences of *Moby-Dick* compel a world into creation. And about this, Shannon's theory says nothing.

For readers, it is the connections that are crucial, for it is those connections themselves—the specific correlations between the words in Melville's novel and the world of large fish and demented whalers—that function as the load-bearing structures. Just how, then, *are* such connections established?

Apparently they just are.

### III

IF, IN reading, every reader embodies a paradox, it is a paradox that in living he exemplifies as well. "Next to the brain," George Johnson remarks, "the most obvious biological information-processor is the genetic machinery of the cell."

The essential narrative is by now familiar. All living creatures divide themselves into their material constituents *and* an animating system of instruction and information. The plan is in effect wherever life is in command: both the reader and the bacterial cell are expressions of an ancestral text, their brief appearance on the stage serving in the grand scheme of things simply to convey its throbbing voice from one generation to another.

Within the compass of the cell itself, there are two molecular classes: the proteins, and the nucle-

ic acids (DNA and RNA). Proteins have a precise three-dimensional shape, and resemble tight tensed knots. Their essential structure is nonetheless linear; when denatured and then stretched, the complicated jumble of a functional protein gracefully reveals a single filament, a kind of strand, punctuated by various amino acids, one after another.

DNA, on the other hand, is a double-stranded molecule, the two strands turned as a helix. Within the cell, DNA is wound in spools and so has its own complicated three-dimensional shape; but like the proteins, it also has an essentially linear nature. The elementary constituents of DNA are the four nucleotides, abbreviated as A, C, G, and T. The two strands of DNA are fastened to one another by means of struts, almost as if the strands were separate halves of a single ladder, and the struts gain purchase on these strands by virtue of the fact that certain nucleotides are attracted to one another by means of chemical affinities.

The structure of DNA as a double helix endows one molecule with two secrets. In replicating itself, the cell cleaves its double-stranded DNA. Each strand then reconstitutes itself by means of the same chemical affinities that held together the original strands. When replication has been concluded, there are two double-stranded DNA molecules where formerly there was only one, thus allowing life on the cellular level to pass from one generation to the next.

But if DNA is inherently capable of reproducing itself, it is also inherently capable of conveying the linear order of its nucleotides to the cell's amino acids. In these respects, DNA functions as a template or pattern. The mechanism is astonishingly complex, requiring intermediaries and a host of specialized enzymes to act in concert. But whatever the details, the central dogma of molecular biology is straight as an arrow. The order of nucleotides within DNA is read by the cell and then expressed in its proteins.

Read by the cell? Apparently so. The metaphor is inescapable, and so hardly a metaphor. As the DNA is read, proteins form in its wake, charged with carrying on the turbulent affairs of the cell itself. It was an imaginary reader, nose deep in Melville's great novel, who suggested the distinction between what words do as material causes and what they achieve as symbols. The same distinction recurs in biology. Like words upon the printed page, DNA functions in any number of causal pathways, the tic of its triplets inducing certain biochemical changes and suppressing others.

And this prompts what lawyers call a leading question. We quite know what DNA is: it is a macromolecule and so a material object. We quite know what it achieves: apparently everything. Are the two sides of this equation in balance?

The cell is, after all, a living system. It partakes of all the mysteries of life. The bacterium *escherichia coli*, for example, contains roughly 2,000 separate proteins, and every one of them is mad with purpose and busy beyond belief. Eucaryotic cells, which contain a nucleus, are more complicated still. Chemicals cross the cell membrane on a tight schedule, consult with other chemicals, undertake their work, and are then capped in cylinders, degraded and unceremoniously ejected from the cell. Dozens of separate biochemical systems act independently, their coordination finely orchestrated by various signaling systems. Enzymes prompt chemical reactions to commence and, work done, cause them to stop as well. The cell moves forward in time, functional in its nature, continuous in its operations.

Explaining all this by appealing to the causal powers of a single molecule involves a disturbing division of attention, rather as if a cathedral were seen suddenly to rise from the head of a carrot. Nonetheless, many biologists, on seeing the carrot, are persuaded that they *can* discern the steps leading to the cathedral. Their claim is often presented as a fact in the textbooks. The difficulty is just that, while the carrot—DNA, when all is said and done—remains in plain sight, subsequent steps leading to the cathedral would seem either to empty in a computational wilderness or to gutter out in an endless series of inconclusive causal pathways.

**F**IRST, THE computational wilderness. Proteins appear in living systems in a variety of three-dimensional shapes. Their configuration is crucial to their function and so to the role they play in the cell. The beginning of a causal process is once again in plain sight—the linear order expressed by a protein's amino acids. And so, too, is the end—a specific three-dimensional shape. It is the mechanism in the middle that is baffling.

Within the cell, most proteins fold themselves into their proper configuration within seconds. Folding commences as the protein itself is being formed, the head of an amino-acid chain apparently knowing its own tail. Some proteins fold entirely on their own; others require molecular chaperones to block certain intermediate configurations and encourage others. Just how a protein manages

to organize itself in space, using only the sequence of its own amino acids, remains a mystery, perhaps the deepest in computational biology.

Mathematicians and computer scientists have endeavored to develop powerful algorithms in order to predict the three-dimensional configuration of a given protein. The most successful of these algorithms gobble the computer's time and waste prodigally its power. To little effect. Protein-folding remains a mystery.

Just recently, IBM announced the formation of a new division, intended to supply computational assistance to the biological community. A supercomputer named Blue Gene is under development. Operating at processing speeds 100 times faster than existing supercomputers, the monster will be dedicated largely to the problem of protein folding.

The size of the project is a nice measure of the depth of our ignorance. The slime mold has been slithering since time immemorial, its proteins folding themselves for just as long. No one believes that the slime mold accomplishes this by means of supercomputing firepower. The cell is not obviously an algorithm, and a simulation, needless to say, is not obviously an explanation. Whatever else the cell may be doing, it is not using Monte Carlo methods or consulting genetic algorithms in order to fold its proteins into their proper shape. The requisite steps are chemical. No other causal modality is available to the cell.

If these chemical steps were understood, simulations would be easy to execute. The scope of the research efforts devoted to simulation suggest that the opposite is the case: simulations are difficult to achieve, and the requisite chemical steps are poorly understood.

**I**F COMPUTATIONS are for the moment intractable, every analysis of the relevant causal pathways is for the moment inconclusive.

As they are unfolding, proteins trigger an "unfolding protein response," one that alerts an "intracellular signaling system" of things to come. It is this system that in turn "senses" when unfolded proteins accumulate. The signal sent, the signaling system responds by activating the transcription of still other genes that provide assistance to the protein struggling to find its correct three-dimensional shape. Each step in the causal analysis suggests another to come.

But no matter the causal pathways initiated by DNA, some *overall* feature of living systems seems stubbornly to lie beyond their reach. Signaling systems must themselves be regulated, their activities

timed. If unfolding proteins require chaperones, these must make their appearance in the proper place; their formation requires energy, and so, too, do their degradation and ejection from the cell. Like the organism of which it is a part, the cell has striking global properties. It is *alive*.

Our own experience with complex dynamical systems, such as armies in action (or integrated microchips), suggests that in this regard command and coordination are crucial. The cell requires what one biologist has called a "supreme controlling and coordinating power." But if there is such a supreme system, biologists have not found it. The analysis of living systems is, to be sure, a science still in its infancy. My point, however, is otherwise, and it is general.

Considered strictly as a material object, DNA falls under the descriptive powers of biochemistry, its causal pathways bounded by chemical principles. Chemical actions are combinatorial in nature, and *local* in their effect. Chemicals affect chemicals within the cell by means of various weak affinities. There is no action at a distance. The various chemical affinities are essentially arrangements in which molecules exchange their parts irenically or like seaweed fronds drift close and then hold fast.

But command, control, and coordination, if achieved by the cell, would represent a phenomenon incompatible with its chemical activities. A "supreme controlling and coordinating power" would require a device receiving signals from every part of the cell and sending its own universally understood signals in turn. It would require, as well, a universal clock, one that keeps time globally, and a universal memory, one that operates throughout the cell. There is no trace of these items within the cell.

Absent these items, it follows that the cell quite plainly has the ability to organize itself *from itself*, its constituents bringing order out of chaos on their own, like a very intricate ballet achieved without a choreographer. And what holds for the cell must hold as well for the creatures of which cells are a part. One biologist has chosen to explain a mystery by describing it as a fact. "Organisms," he writes, "from daisies to humans, are naturally endowed with a remarkable property, an ability to *make themselves*."

Naturally endowed?

Just recently, the biologist Evelyn Fox Keller has tentatively endorsed this view. The system of control and coordination that animates the cell, she observes in *The Century of the Gene*, "consists of, and lives in, the interactive complex made up of genom-

ic structures and the vast network of cellular machinery in which those structures are embedded." This may well be so. It is also unprecedented in our experience.

We have no insight into such systems. No mathematical theory predicts their existence or explains their properties. How, then, *do* a variety of purely local chemical reactions manage to achieve an overall and global mode of functioning?

INFORMATION NOW makes its second appearance as an analytic tool. DNA is a molecule—that much is certain. But it is also, molecular biologists often affirm, a library, a blueprint, a code, a program, or an algorithm, and as such it is quivering with information that it is just dying to be put to good use. As a molecule, DNA does what molecules do; but in its secondary incarnation as *something else*, DNA achieves command of the cell and controls its development.

A dialogue first encountered on the level of matter (DNA as a molecule and nothing more) now reappears on the level of metaphor (DNA as an information source). Once again we know what DNA is like, and we know what it does: apparently everything. And the question recurs: are the sides of this equation in balance?

Unfortunately, we do not know and cannot tell.

Richard Dawkins illustrates what is at issue by means of a thought experiment. "We have an intuitive sense," he writes,

that a lobster, say, is more complex (more "advanced," some might even say more "highly evolved") than another animal, perhaps a millipede. Can we *measure* something in order to confirm or deny our intuition? Without literally turning it into bits, we can make an approximate estimation of the *information contents* [emphasis added] of the two bodies as follows. Imagine writing a book describing the lobster. Now write another book describing the millipede down to the same level of detail. Divide the word-count in one book by the word-count in the other, and you have an approximate estimate of the relative information content of lobster and millipede.

These statements have the happy effect of enforcing an impression of quantitative discipline on what until now have been a series of disorderly concepts. Things are being measured, and that is always a good sign. The comparison of one book to another makes sense, of course. Books are made up of words the way computer programs are made up

of binary digits, and words and binary digits may both be counted.

It is the connection outward, from these books or programs to the creatures they describe, that remains problematic. What level of detail is required? In the case of a lobster, a very short book comprising the words "Yo, a lobster" is clearly not what Dawkins has in mind. But adding detail to a description—and thus length—is an exercise without end; descriptions by their very nature form an infinitely descending series.

Information is entirely a static concept, and we know of no laws of nature that would tie it to other quantitative properties. Still, if we cannot answer the question precisely, then perhaps it might be answered partially by saying that we have reached the right level of descriptive detail when the information in the book—that is, the lobster's DNA—is roughly of the same order of magnitude as the information latent in everything that a lobster is and does. This would at least tell us that the job at hand—constructing a lobster—is doable insofar as information plays a role in getting *anything* done.

Some biologists, including John Maynard Smith, have indeed argued that the information latent in a lobster's DNA *must* be commensurate with the information latent in the lobster itself. How otherwise could the lobster get on with the business at hand? But this easy response assumes precisely what is at issue, namely that it is *by means of information* that the lobster gets going in the first place. Skeptics such as ourselves require a direct measurement, a comparison between the information resident in the lobster's DNA and the information resident in the lobster itself. Nothing less will do.

DNA is a linear string. So far, so good. And strings are well-defined objects. There is thus no problem in principle in measuring their information. It is there for the asking and reckoned in bits. But what on earth is one to count in the case of the lobster? A lobster is not discrete; it is not made of linear symbols; and it occupies three or four dimensions and not one. Two measurements are thus needed, but only one is obviously forthcoming.

The unhappy fact is that we have in general no noncircular way of specifying the information in any three- (or four-) dimensional object except by an appeal to the information by which it is generated. But this appeal makes literal sense only when strings or items like strings are in concourse. The request for a direct comparison between what the lobster has to go on—its DNA—and what it is—a living lobster—ends with only one measurement in place, the other left dangling like a mountain climber's rope.

We are thus returned to our original question: how do symbols—words, strings, DNA—bring a world into being?

Apparently they just do.

#### IV

ONE MIGHT hope that in one discipline, at least, the situation might be different. Within the austere confines of mathematical physics, where a few pregnant symbols command the flux of space and time, information as an idea might come into its own at last.

The laws of physics have a peculiar role to play in the economy of the sciences, one that goes beyond anything observed in psychology or biology. They lie at the bottom of the grand scheme, comprising principles that are not only fundamental but irreducible. They must provide an explanation for the behavior of matter in all of its modes, and so they must explain the *emergence* as well as the organization of material objects. If not, then plainly they would not explain the behavior of matter in all of its modes, and, in particular, they would not explain its *existence*.

This requirement has initiated a curious contemporary exercise. Current cosmology suggests that the universe began with a big bang, erupting from nothing whatsoever 15 billion years ago. Plainly, the creation of something from nothing cannot be explained in terms of the behavior of material objects. This circumstance has prompted some physicists to assign a causative role to the laws of physics *themselves*.

The inference, indeed, is inescapable. For what else is there? "It is hard to resist the impression," writes the physicist Paul Davies, "of something—some influence capable of transcending space-time and the confinements of relativistic causality—possessing an overview of the entire cosmos at the instant of its creation, and manipulating all the causally disconnected parts to go bang with almost exactly the same vigor at the same time."

More than one philosopher has drawn a correlative conclusion: that, in this regard, the fundamen-

tal laws of physics enjoy attributes traditionally assigned to a deity. They are, in the words of Mary Hesse, "universal and eternal, comprehensive without exception (omnipotent), independent of knowledge (absolute), and encompassing all possible knowledge (omniscient)."

If *this* is so, the fundamental laws of physics cannot themselves be construed in material terms. They lie beyond the system of causal influences that they explain. And in this sense, the information resident in those causal laws is richer—it is more abundant—than the information resident in the universe itself. Having composed one book describing the universe to the last detail, a physicist, on subtracting that book from the fundamental laws of physics, would rest with a positive remainder, the additional information being whatever is needed to bring the universe into existence.

WE ARE now at the very limits of the plausible. Contemporary cosmology is a subject as speculative as scholastic theology, and physicists who find themselves irresistibly drawn to the very largest of its intellectual issues are ruefully aware that they have disengaged themselves from any evidential tether, however loose. Nevertheless, these flights of fancy serve a very useful purpose. In the image of the laws of nature zestfully wrestling a universe into existence, one sees a peculiarly naked form of information—naked because it has been severed from every possibility of a material connection. Stripped of its connection to a world that does not yet exist, the information latent in the laws of physics is nonetheless capable of *doing* something, by bringing the universe into being.

A novel brings a world into creation; a complicated molecule an organism. But these are the low taverns of thought. It is only when information is assigned the power to bring something into existence from *nothing whatsoever* that its essentially magical nature is revealed. And contemplating magic on this scale prompts a final question. Just how did the information latent in the fundamental laws of physics unfold itself to become a world?

Apparently it just did.