4

The Demonology of Information

In the beginning, the main route connecting Santa Fe to the rest of the known universe was the Camino Real, the royal highway that ran up from Mexico City, meeting the Rio Grande at El Paso and following it northward through Albuquerque, Santa Fe, and on to the hinterlands of New Spain. Today the American portion of the Camino Real has been replaced by Interstate 25, but the scenery along the route remains pretty much the same. Those who fly into Albuquerque International Airport and drive north for a scientific conference in Santa Fe or Los Alamos pass through a stark landscape very much like that the Spanish conquistadores saw.

To the east, as one leaves the suburban sprawl of the Albuquerque metropolitan area, the Sandia Mountains rise nearly six thousand feet above the already mile-high terrain, exposing a rocky facade so fractured and so sheer it looks as though half the mountain has been sliced away. In a sense that is what happened. The Sandias are an example of what geologists call a fault-block mountain. Like the Sangre de Cristos they were squeezed from the earth when two continental plates collided, but in the case of the Sandias, one side collapsed; instead of a slope, the western face of the mountain is a bare, almost vertical expanse of steep granite walls. The most prominent of these is the Shield, so formidible, the guidebooks say, that some of its more onerous ascents can take days of hard climbing, the nights spent roped to the cliff like a tent worm, trying to fall asleep on vertical ground.

To the west, beyond a line of dormant volcanoes, one can barely see Mount Taylor, a jagged blue bump on the horizon that was named after General Zachary Taylor, after he took this land from the Mexicans in the War of 1846. The Mexicans, and the Spanish before them, called the mountain Cebolleta, "Little Onion." They took it from the Navajos, who still call it Turquoise Mountain and consider it the southern border of their universe and the home of Monster Slayer, one of the legendary Hero Twins who fought against the evils of the earth. Drive west from Albuquerque on Interstate 40, old Route 66, and just before Grants, a mining town that in better times billed itself as the Uranium Capital of the United States, you cross over the petrified bubbles of the Malpais ("Bad Land") lava flow; the Navajos say it is the dried blood of Ye-itsa, one of Monster Slayer's victims. Ye-itsa's head can be found to the north in the form of an old volcanic plug with sloping shoulders that the Spanish named Cabezón Peak. (*Cabeza* means head, and a *cabezón* is one that is particularly big and ugly.) Ye-itsa's bones (the geologists say they are petrified trees) lie as far east as Albuquerque. Though Ye-itsa was killed and turned to stone, some of the other monsters survived, the legend goes. Demons called hunger, greed, filth, and old age still stalk the land.

The pueblo Indians included the Navajos among the monsters and still remember the stories of their raids on the adobe villages that lie between Albuquerque and Santa Fe along the Rio Grande—Sandia Pueblo, Santa Ana, Santo Domingo, San Felipe, Cochiti, little worlds with their own languages and, like the Tewa pueblos to the north, their own quartets of magic mountains marking off their personal universes. The landscape on this part of the journey is like nothing else on earth. Far to the west the Jemez Mountains reach toward the river with fingers of lava, hardened into the black, flat mesas that, to use another metaphor, look like frozen breakers of stone. The turnoff to San Felipe, a hive of adobe houses shaded with cottonwoods hunched against the base of one of the larger mesas, marks the halfway point of the drive to Santa Fe. A few miles later, just after the highway crosses the dusty arroyo known as the Galisteo River, a steep volcanic wall looms into view. The Spanish called it La

Bajada, "The Descent," though when one is driving up from Albuquerque it is quite the opposite, an eight-hundred-foot rise that divides the lower country of southern New Mexico from the highlands of the north. Until this point the highway has been cutting across what the Spanish cartographers called Río Abajo, "Lower River," the part of the northern kingdom that lay closest to Mexico City. In those days of horses and wagons, La Bajada was known for the treachery of its hairpin turns—the price one paid for entering another realm: Río Arriba, "Upper River," the vast, barely explored region that extended north of La Bajada and then off the top of the maps.

It is fitting that La Bajada was named from the Río Arribans' point of view. Sitting in their perch seven thousand feet above sea level, the people of Santa Fe and beyond literally and sometimes figuratively looked down on their neighbors in Río Abajo. Except for a gradual rise to reach the top of the La Bajada hump, it was a two-thousand-foot slide from Santa Fe to Albuquerque. The change wasn't simply one of geography. La Bajada was, and is, a psychological and a cultural divide. Though southern New Mexico has its share of mountains, it is largely a flat, desert land whose subtle beauty requires the heart and eye of a connoisseur. There is nothing subtle about the topography of northern New Mexico. Once you ascend La Bajada, with the Sangre de Cristos looming straight in front of you and the carved symmetrical volcanoes rising from either side of the highway, Río Arriba opens up all around. You know you are in another country, where even the light seems changed.

It wasn't easy, going between the land of the familiar and the land of the strange. Wagon drivers coming down La Bajada often had to brace their wheels with rocks to keep from succumbing to the force called gravity. Cars heading the other way sometimes had to back up the hill, reverse gear providing them with more leverage, as their boiling radiators protested against the heat. Today the endless turns have been straightened into a more gradual ascent; cars and trucks barely slow down as they surmount the divide. But they are still bound by the same laws of physics that held sway in the conquistadores' time. Then and now, it takes energy to cross the divide.

In May 1989, some three dozen scientists, mostly from the United States but a few from as far away as Germany, Britain, France, Israel, and Japan, flew into Albuquerque and boarded rental cars and shuttle buses for the journey up the Camino Real. Skirting the edges of the pueblo universes, they ascended La Bajada, arriving in Santa Fe for a conference sponsored by the Santa Fe Institute and held in the spectacular setting provided by St. John's College, which sits at a confluence of arroyos that cut through the foothills of the Sangre de Cristo Mountains. Hike three miles up the canyons from St. John's and you reach Atalaya Peak. *Atalaya* means "watchtower," and if you stand on its heights and look down on Santa Fe and across to Los Alamos you will be seeing what may be the world's largest concentration of scientists (granted, there aren't many) working in a new field called the physics of information, which sits at the boundary where mind and nature, subject and object, seem to collide.

In some ways, St. John's seemed an incongruous setting for a conference on so revolutionary a subject as information and physics. The school is known for its classical curriculum: students learn physics by starting with the pre-Socratics, then moving on to the more recent ideas of Plato and Aristotle. The physicists and mathematicians were coming to St. John's to discuss ideas at the very edge of twentieth-century science. They were responding to a manifesto with the intriguing title "Complexity, Entropy, and the Physics of Information," which had been dispatched by Wojciech H. Zurek, a Polish-born physicist who works at the Los Alamos National Laboratory's Theoretical Astrophysics Group.

In building a tower of abstraction, one must start with a foundation, those things that are taken as given: mass, energy, space, time. Everything else can then be defined in terms of these fundamentals. But gradually over the last half century some scientists—and Zurek was among the most adamant—had come to believe that another basic ingredient was necessary to make sense of the universe: information. "The specter of information is haunting the sciences," his manifesto began. There is a "border territory," he believed, where information, physics, complexity, quantum theory, and computation meet. So in another way, St. John's wasn't so strange a setting for the conference after all. What Zurek and his colleagues had in mind was a return to basics, a rethinking of reality's pillars as thorough as any undertaken by Thales, who thought all was made of water, or Heraclitus, who thought all was made of fire.

Most of us are used to thinking of information as secondary, not fundamental, something that is made from matter and energy. Whether we are thinking of petroglyphs carved in a cliff or the electromagnetic waves beaming from the transmitters on Sandia Crest, information seems like an artifact, a human invention. We impose pattern on matter and energy

[redacted]

and use it to signal our fellow humans. Though information is used to describe the universe, it is not commonly thought of as being part of the universe itself. But to many of those at the Santa Fe conference, the world just didn't make sense unless information was admitted into the pantheon, on an equal footing with mass and energy. A few went so far as to argue that information may be the most fundamental of all; that mass and energy could somehow be derived from information.

There was, first of all, the mysterious connection that seemed to exist between information, energy, and entropy, the amount of disorder in a system. We learn in school that, left on its own, any closed system becomes more and more disorderly; its entropy increases. It is because of this fact, embodied in the second law of thermodynamics, that neat geological strata become gnarled into formless Precambrian rock. The planar geometry of an adobe village melts until it is barely distinguishable from the surrounding hills. Along the way, pattern is washed away; information is lost. Information can be thought of as a measurement of distinctions, the simplest being 1 or 0, the presence or absence of a certain quality. By this measure, there is more information in something that is orderly than in a homogeneous, undifferentiated mess.

On the other hand, by gathering and processing information, we can *create* order—we can take the matter and energy of our world and arrange it into songs, civilizations, fragile eddies in the entropic tide. Using our powers as information processors, we can find unlikely structures that already exist—water trapped in a mountain lake, carbon molecules strung in a volatile chain, protons and neutrons stacked into a precarious nuclear sphere. And then we simply let them follow the path of least resistance. As they topple and move down the hill from order to disorder, we can extract work by harnessing the entropic flow. The nucleus disintegrates, the bonds of the carbon atoms break, the water flows from its pool to the formless sea. Entropy increases, information is lost, but the energy released in the process can be tapped to build new structures, to create information, though all our creations must eventually succumb to the second law.

No wonder the mind craves patterns. It is the ability to find order in the world that allows us to make use of its resources. For many scientists this would be reason enough to believe that information is fundamental. But, going beyond the laws of thermodynamics, some believe information plays an even deeper role. According to some interpretations of quantum theory put together by Zurek and his circle, without information there would be no resources to exploit and no one to exploit them; there would be nothing that resembled what we call the real world. The mathematics used to describe the subatomic realm tells us that, left to its own devices, an electron lacks the very attributes that we, on our macroscopic plateau, consider the very hallmark of existence—a definite position in time and space. It exists, we are told, as a probability wave, a superposition of all the possible trajectories that takes on substance only when it is measured, when, as it is often put, an observer collapses the probability wave. How this transformation occurs is one of the deepest mysteries of physics, the so-called measurement problem: How does the rock-solid classical world, in which things occupy definite positions in space and time, crystallize from the quantum haze? In the past, quantum theory has often been embraced by those who would elevate subjectivity over objectivity, championing a mystical world view in which consciousness brings the universe into being. By making information fundamental, Zurek and some of his colleagues hoped to demystify quantum theory. For what is an observation but a gathering of information? And if information is fundamental, it exists as surely as does matter and energy, without the need of conscious beings. The quantum wave might collapse not because it was beheld by a mind but simply because information flowed from one place to another in the subatomic realm.

Of course, it is easy to be fooled by our own metaphors, becoming so dazzled by the concepts we invent that we can see the world only through their glare. In the nineteenth century, entropy and the laws of thermodynamics were invented to deepen our understanding of the steam engine and make it as efficient as nature would allow. Any closed system, sealed off from its environment, would inevitably march from order to disorder. Soon scientists and philosophers were applying these new mental tools to the universe itself, declaring that, as the most closed of closed systems—what could possibly be outside of it?—the universe was marching inevitably toward thermodynamic death, a state of equilibrium, lifeless, unstructured, random. In the twentieth century, information theory was invented to help engineers make electronic communications channels as efficient as possible. And before we knew it, people were speaking of information as real, a few going so far as to imagine that we live in a universe of computation, created from the shuffling of bits.

One of the challenges implicit in Zurek's manifesto was to find new ways to think about whether computation—and therefore information—is natural or artificial. The computers we have built over the years

have been crafted from macroscopic parts: first gears, then vacuum tubes, then transistors, and now chips inscribed with thousands of transistors that get smaller and more densely packed every year. We stamp our designs on nature's designs; circuitry onto silicon lattices. But the finer the blueprints of our artifices, the more they begin to clash with the physics underneath. Quantum randomness scrambles our neat choreography of 1s and 0s. But perhaps as engineers reach tinier and tinier scales they can somehow exploit the natural behavior of atoms to make their machines more efficient, bridging the divide between the circuitry we design and the "circuitry" of nature. An atom with an electron that could be in one of two states might naturally be thought of as a register containing a 1 or a 0. How thin can we make this gap between the laws of computation and the laws of physics? Where will the shrinking bottom out? If computation can take place only down to a certain scale, requiring components made up of many, many molecules, then perhaps information is simply an artifice, secondary to the laws of physics, a pattern imposed by people as they struggle to describe the world. But if single molecules or even atoms can be said to somehow process information, then maybe computation is as fundamental as what we think of as the laws of physics. Like mass and energy, information would be irreducible, at the roots of creation.

For many of the people who gathered in Santa Fe to talk about information, thermodynamics, and quantum theory, this would be the first of many visits to northern New Mexico. Another conference followed a year later, this one at the Santa Fe Institute, which was then housed in an old convent among the galleries and adobe houses on Canyon Road. In a way, though, the first conference never really ended. Over the years, the physics of information group Zurek started at the Santa Fe Institute has attracted a changing cast of visitors. Rolf Landauer and Charles Bennett, two of the first people to make a connection between physics and information, visit often from the IBM Thomas J. Watson Research Center in New York. At his retreat in Tesuque, a rural village that provides refuge for those who find even Santa Fe's slow pace too frenetic, Murray Gell-Mann and his frequent guest James Hartle, of the University of California at Santa Barbara, try to use information to make sense of quantum cosmology, in which the whole universe can be thought of as a quantum probability wave.

As one listened to these scientists' lectures, read their papers, and spoke to them privately, at dinner or in hikes through the mountains, it was hard not to be struck by hints of an even deeper purpose to their travails. The physicists at Santa Fe were not simply doing science. In this land where so many people see the universe in so many different ways, they were examining the very nature of the scientific enterprise, this curious drive we have for gathering bits and weaving them into pictures of the world.

In fact, to some of the visitors making the drive up La Bajada to discuss their ideas with colleagues in Los Alamos and Santa Fe, it has become natural to think of information as the fuel that, quite literally, takes them over the divide. During one of the Santa Fe Institute conferences, Charles Bennett of IBM declared that given a long enough memory tape—a blank string to be filled with 1s and 0s—he would have all the energy he needed to drive from Albuquerque to Santa Fe. Several years later, at a conference in Dallas called "The Symbiosis of Physics and Information," Bennett said he couldn't recall making the statement, but that it was not one with which he would disagree. "It's what I believe," he said. "It definitely sounds like something I would say."

To all but the handful of initiates, it sounds impenetrably mysterious, this notion that information and energy could be somehow intertwined. To understand what Zurek, Bennett, Landauer, and their colleagues have in mind, one must become submerged in a way of thinking and carving up the world that has its origins in the late nineteenth century, when James Maxwell tried to pick open a loophole in what was thought to be an unassailable universal law. In 1871, several years after inventing the equations braiding together electricity and magnetism, Maxwell publicly introduced, in his book *Theory of Heat,* an imaginary imp, later to be dubbed Maxwell's demon, that seemed to have the ability to outthink the second law of thermodynamics.

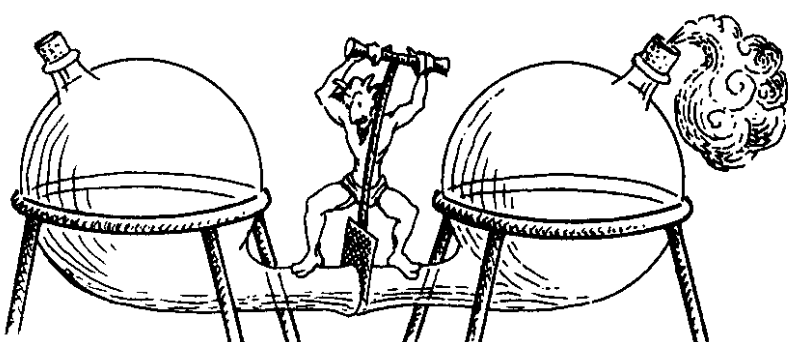
In the age of the computer it is hard to imagine how something as prosaic as the steam engine could have done so much to shape nineteenth-century thought. Someday, perhaps, our own preoccupation with the digital computer will seem just as quaint. In contemplating how to get Robert Fulton's engine to mesh as closely as possible with the laws of nature, squeezing the maximum amount of work from the steam, Sadi Carnot, a French army engineer, concluded that even with his utmost efforts, he could never hope to reach an efficiency of 100 percent. In transforming the energy of the steam into the energy needed to turn a wheel, some

would inevitably, irreversibly, leak away. This truth was expressed in the form of the two laws of thermodynamics. The first law can be taken as the good news: it declares that energy is indeed conserved, that it can be neither created nor destroyed, but simply changed from one form to another. The second law, however, tells us that whenever energy is put to use it is degraded: the potential energy of water stored behind a dam turns to kinetic energy and then to electricity as it rushes down the spillway and turns the turbine blades of a generator. In the end the accounts must balance: the energy coming out must equal the energy that went in. But not all of the energy of the water can be converted into electricity. Some is dissipated in the form of heat—the friction of water molecules bumping into air molecules and into each other, the friction of the imperfect bearings on the turbine blades, the resistance of the electricity in the wires. The energy of the wasted heat is still somewhere in the environment, in the form of randomly vibrating molecules. We can imagine ways to recapture some of this random motion and channel it back into the system. But it can never be completely recovered. If it weren't for this loss, we could use a generator to power a motor and then use the motor to turn the generator and have a perpetual motion machine.

Rudolph Clausius, in Berlin, was so struck by this inevitable change from useful to useless energy that he gave it a name: entropy. Water above a dam, steam compressed in a chamber, a spring wound tight, a battery with its negative charges sequestered from its positive charges—all are in highly structured states and are said to have low entropy. As they do work they become randomized. Viewed this way, entropy is a measure of disorder, and what the second law is telling us is that the march toward randomness is inevitable. One can reduce entropy (water can be pumped back uphill; a dead battery can be recharged, its homogenized negative and positive ions divided between the two poles again), but only by expending energy. And this produces more entropy. Our refrigerators freeze shapeless water into the crystalline lattices called ice, but as they do so, heat, the random vibration of molecules, is exported into the room. In the long run, entropy always wins. Pockets of order must be paid for with larger pockets of disorder, and the system as a whole—the universe—increases in entropy. We are fortunate in finding around us huge stores of potential energy, clocksprings already wound—food, fossil fuels, rivers, uranium. By letting them flow down the energy hill, we can run our civilization. For now, the whole system is continuously recharged by the sun. But eventually it too must run down.

There is also a third law of thermodynamics, which insists that it is impossible to reach absolute zero, the temperature at which all molecular motion would cease. Thus there will always be heat in the world, the energy of these randomly moving molecules. But according to the second law, it always takes more work to harness this scattered, ubiquitous motion than we can possibly gain from the attempt. Otherwise our cars and our appliances—for that matter, trees, animals, anything that requires power—could run by themselves, fueled by nothing more than this bottomless sea of vibrations.

It was quite a radical move when, in a thought experiment, Maxwell tried to devise a way to break the second law, to show that if a creature were clever enough it could create energy out of thin air. He began by imagining a vessel divided into two chambers, connected by a small tubular passage. Suppose you place a barrier in the passage and then fill one chamber with a hot gas. Remove the barrier and this initially ordered system, with all the heat on one side and none on the other, will quickly move to a state of equilibrium, with both sides filled with a gas at a lower, uniform temperature, a homogeneous expanse of randomly moving molecules. Place a paddle wheel or a piston in the passageway and this rush to disorder lets us do work. But once the system is in thermal equilibrium, with entropy at a maximum, the second law tells us that there is no way to extract any more work from the gas. We would have to pump it back into one chamber, and that would require energy.



But, Maxwell wondered, why couldn't you instead place a small, intelligent being in the middle of the contraption to observe the movements of the molecules and manipulate the valve so that the faster ones congregated in chamber A, while the slower ones stayed in chamber B?

When a molecule came speeding from chamber B toward chamber A, this "very observant and neat-fingered" being would open the valve and let it through. It would close the valve if it saw a fast molecule about to escape from A to B. Merely by the exercise of its wits, Maxwell contended, the demon would cause the temperature in A to exceed that in B; it could build up a potential and use it to do work. With intelligence, it seemed, one could overcome entropy.

Maxwell wasn't seriously interested in building a perpetual motion machine. Where, after all, was one going to find one of these uncomplaining little slaves? His purpose was to show that, unlike the laws science had proposed in the past, the second law is not absolute but statistical; the best we can say is that it works the overwhelming majority of the time. Even in a system without a demon and a valve, there is a tiny but real chance that the fast molecules would just happen to congregate on one side and the slow molecules on the other. But the vast likelihood is that the temperature would even out: for each chance fluctuation that put a fast molecule in the left chamber, another fluctuation could be expected to put a fast molecule in the right chamber.

Still, it *could* happen. Individual gas molecules don't know about the second law and the one-way flow toward entropy; they simply obey the laws of mechanics. Nothing in Newton's writ bans the possibility that after the valve was opened and the gas flowed from chamber A to chamber B, a majority of the molecules might reverse course and flow back into the first chamber. The system would be restored to its original state, providing us with work for free. But the chance of this happening is so remote that we would be better off waiting around for a swarm of fireflies to spell out messages in the sky.

The moral, Maxwell once wrote, is that the second law of thermodynamics "has the same degree of truth as the statement that if you throw a tumblerful of water into the sea, you cannot get the same tumblerful of water out again." But again, it could happen, either by chance or the good graces of an observant, nimble-fingered demon gathering the molecules up again and putting them back into the glass. Entropy, it seemed, was a measure of ignorance; it depended on the observer.

The demon was a fantasy. But between its perfect perception and nimbleness and our myopic clumsiness, one can imagine a continuum of creatures endowed with different powers. Nature is like a text and these various beings will vary in their ability to decipher its code, or even to suspect that there are patterns there to divine. In an article he wrote for the 1878 edition of the *Encyclopaedia Britannica,* Maxwell compared the situation with trying to read a notebook written in the owner's personal shorthand: "A memorandum-book does not, provided it is neatly written, appear confused to an illiterate person, or to the owner who understands it thoroughly, but to any other person able to read it appears to be inextricably confused. Similarly the notion of dissipated energy would not occur to a being who could not turn any of the energies of nature to his own account, or to one who could trace the motion of every molecule and seize it at the right moment. It is only to a being in the intermediate stage, who can lay hold of some forms of energy while others elude his grasp, that energy appears to be passing inevitably from the available to the dissipated state." The implication was that entropy existed for moderately intelligent creatures like people but not for demons or dogs—that order and disorder were in the eye of the beholder.

If the second law was indeed statistical, then the best way to treat it was with the mathematics of probability. Seen this way, systems tend to move from ordered (unlikely) to disordered (likely) states because there are vastly more disordered ones. Try to imagine the countless ways in which gas molecules could arrange themselves in a closed vessel. In a tiny number of these configurations, the molecules will appear bunched into one corner or another, or sequestered in various-shaped blobs; in a precious few cases, they might arrange themselves in spheres or cubes. According to what statisticians call the ergodic hypothesis, the gas will eventually visit every one of its possible arrangements as its molecules wander randomly through the chamber; one is no more likely than the other. But in the vast, vast majority of possible arrangements, the molecules will form what appears to our myopic eyes as a featureless mix uniformly occupying the container.

In Maxwell's engine, we begin with the molecules forced into an unlikely arrangement, all occupying chamber A. When we open the door they rush to assume one of the vastly more likely arrangements in which the gas is uniformly distributed throughout both chambers. Another way to say it is that by opening the door we give the gas more "degrees of freedom," twice as much room to roam.

Probability could also be used to explain the inevitable sucking away of energy through friction and other forms of dissipation. The environment, after all, represents a huge, essentially infinite number of degrees of freedom. If we shatter Maxwell's vessel, the molecules will escape, fanning out through a labyrinth so vast and convoluted that they can

never find their way back again. And so it is with the heat produced by an engine or any kind of machine. When these vibrating molecules are allowed to bump up against the molecules in the open wilderness of the air, the energy follows the path of least resistance, flowing irretrievably into the great beyond.

Among Victorian intellectuals, the demon elicited two extreme reactions, both reaching far beyond anything Maxwell seems to have intended. To those who took comfort in the objectivist creed, that it was possible to stand outside creation and see it whole, the notion that entropy was subjective was seriously disturbing. Could the second law really be no more than an anthropomorphic effect caused by our myopia and clumsiness and the fact that we are so much larger than molecules? Others found Maxwell's thought experiment liberating and declared that intelligence was a force that could somehow overcome the constraints of physical law, a solace against the gloomy idea of a universe doomed to increasing entropy. By precisely monitoring and manipulating molecules, a creature could (theoretically at least) outwit the second law. There was something special about life and mind that eluded the cold equations of the physicists. Or so some people wanted to believe.

In an attempt to dispel such wishful thinking, some scientists tried to lobotomize the demon by showing that Maxwell's paradox would arise even when mind was removed from the mix. One didn't need to actively sort the molecules. The second law could be overcome, they argued, with nothing more than a one-way door; it would passively swing open when a molecule traveling in one direction collided with it but would stay closed when it was struck by a molecule coming the other way. Eventually more molecules would accumulate on one side of the door than the other: energy for free. Maxwell himself found this convincing—after all, his only intention was to show that the second law was statistical, not to elevate intelligence to the supernatural realm. "I do not see why even intelligence might not be dispensed with and the thing made self-acting," Maxwell wrote, adding later: "This reduces the demon to a valve. As such value him. Call him no more a demon but a valve."

But as it turned out, Maxwell conceded the point too easily. In 1912 the Polish physicist Maryan Smoluchowski showed that a trapdoor tiny enough to serve as an automatic demon would absorb heat and vibrate so wildly that it would be completely ineffective. But, he allowed, "such a device might, perhaps, function regularly if it were appropriately operated by intelligent beings." Could thought overcome entropy after all?

The argument lay in this murky realm until 1929, when it was taken up by Leo Szilard, the Hungarian-born physicist who would later be so instrumental in the founding of the Manhattan Project. The title of his paper, "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings," sounds like another attempt to elevate mind over matter. But actually Szilard's intent was to demystify the demon by replacing the ethereal notion of mind with the more concrete notion of information processing. By doing so he set off a chain of arguments and counterarguments that can be traced sixty years later to Charles Bennett's pronouncement about using information to get up and over La Bajada.

To crystallize his argument, Szilard reduced Maxwell's apparatus to its simplest possible form: a chamber with a single gas molecule wandering randomly inside. First the demon would insert a movable partition in the middle of the chamber. Then it would determine which side the molecule was on, left or right. By hooking up a rope and pulley to the proper side of the partition, the demon could use it as a piston. As the molecule pushed against the barrier, it would pull the rope, turn the pulley, and lift a weight. Potential energy would now be stored in the weight hovering above the ground. By dropping it on a piezoelectric crystal, which generates electricity when it is squeezed, or by using it to pull a belt attached to the armature of a generator, the demon could do work. Then it could decouple the weight and remove the piston from the chamber. With the system back in its original state, the demon could repeat the process, seemingly creating work from nothing more than its ability to perceive which side of the partition the molecule was on.

So far this sounds like just another version of Maxwell's tale. But Szilard reached quite a different conclusion. His breakthrough was to realize that the demon's measurement, determining whether the molecule was on the left or right side, entailed making a binary record, recording what we now would call a bit of information, 1 or 0, left side or right. And making this measurement, Szilard suggested, inevitably consumed a certain amount of energy—enough to ensure that the second law was not violated. In rigging a Maxwellian demon one was traditionally allowed to assume things like perfectly frictionless pistons, the justification being that there is nothing in the laws of physics that would keep one from approaching this ideal as closely as technology and cleverness allowed. But Szilard proposed what amounted to an underlying limit. The very act of

gathering information, he implied, must always dissipate at least enough energy to offset any gain in work and ensure that perpetual motion was impossible.

In the old demon arguments, mind had been looked upon as something separate from the material world; it was an essence with powers of its own. Long before the beginning of information theory and computer science, Szilard focused the argument by showing that, in this simple case at least, intelligence could be thought of as processing bits. And processing bits expended energy. In building our intellectual cathedrals, we might think of ourselves as detached observers, but our powers are finite, our observations rooted in the physical world.

As some would later put it, Szilard showed that there is no such thing as an immaculate perception. This idea was further explored in a paper published in 1951 by the Frenchman Léon Brillouin, who proposed that there would be no way for a demon to sort molecules without seeing them—using a flashlight to bounce photons from the molecules to its eyes. One could imagine making the beam weaker and weaker, but eventually, Brillouin argued, you would bottom out at a minimum intensity. The chamber is, after all, filled with vibrating molecules. A signal that was too weak would be indistinguishable from the surrounding noise. In the same year, the physicist Dennis Gabor calculated that as the light beam was made weaker and weaker, it would, for reasons of quantum mechanical uncertainty, become harder and harder to focus. Again, the implication was that processing information required a minimum amount of energy.

Though this idea was fated to undergo an important modification, in a fundamental sense Szilard, Brillouin, and Gabor were on the right track. Information was becoming less ethereal, a choice between two states of a physical system: molecule on the left or molecule on the right. In retrospect we can see that Szilard showed it was indeed true that the demon could be replaced by a machine, but only if the machine was an information processor, a computer of some kind. And computers must be plugged into the wall. The work the electronic demon gained by processing information and lowering entropy would be offset by the kilowatt-hours it consumed; the entropy is not eliminated but rather exported into the environment—the heat produced by the generating station and the resistance in the wires.

What Szilard had intuited about information and entropy Claude

Shannon made more solid and precise. Shannon, who worked for Bell Laboratories, was studying the best way to encode signals so that they could be transmitted without becoming hopelessly garbled by the random molecular vibrations called noise. Though Shannon was dealing with telephone lines, not heat engines, the general issues were the same as those for thermodynamics—how, in a universe ruled by an inexorable tendency toward disorder, do you preserve structure amid randomness? The fruit of Shannon's investigation was a pair of papers, published in 1948, in which he derived a mathematical expression for the amount of information in a signal. As it turned out, the expression was essentially the same as the one derived in the previous century for entropy.

In retrospect, this connection is not so surprising. Chaitin and Kolmogorov later showed that random, incompressible numbers have a higher algorithmic information content than orderly compressible numbers—it takes longer computer programs to spit out the random numbers. However, many of Shannon's followers found it more intuitively satisfying to put a minus sign in front of the expression for information, making it the opposite of entropy. A highly ordered, low-entropy system can be said to contain a high level of information—distinctions that can be encoded with bits. All the gas is confined in chamber A, none in chamber B. But open the valve and the information content decreases as the entropy rises. What can one say about a random, homogeneous mix of molecules? It is featureless, with no distinctions to be made. Since orderly systems are less probable than disorderly ones, Shannon's measure of information is sometimes called statistical information (as opposed to Kolmogorov and Chaitin's algorithmic information). It is also sometimes called negentropy, the opposite of disorder.

Since, in the general scheme of things, an orderly, high-information system is a rare device, information is also sometimes called a measurement of the degree of surprise. Hiking up a mountain, we are startled to look down and find a perfect arrowhead at our feet: it has a higher statistical information content than a rough piece of granite. If we lose the trail, we look for a marker—stones piled up to form a cairn. The more stones in the pile, the less likely it is that they fell that way by chance. Of course, the amount of structure in a system lies in the eye of the beholder—remember Maxwell's story about the memo book, or think of the labyrinthine molecular world open to those who could see past the surface of a piece of rock. Thus Shannon's new information theory reinforced the notion that there was something subjective about entropy and order.

Though many scientists were intrigued by the resemblance between the mathematical expression for entropy and the one Shannon derived for information, not everyone liked the idea of introducing this slippery concept as one of the atoms of creation. Shannon himself was skeptical of the interpretations. It was one thing for engineers to introduce a concept called information for analyzing man-made systems, but quite another to claim that it was an important part of the physical world. If the cost of saving the second law required accepting that there was a subjective element to our perceptions of randomness and order, the very basis by which we carve up the universe, then many nonbelievers wanted no part of the trade-off.

Their challenge was to show that there wasn't really any information floating around in Szilard's engine, that the reason it was incapable of generating perpetual motion was not the cost of processing information but more mundane considerations, like the thermal vibrations that had defeated Smoluchowski's trapdoor. By replacing the demon with ingenious arrangements of sensors and electromechanical devices to engage the gears and pulleys, they tried to design an automated version of Szilard's single-molecule engine in which no binary decision—left or right, 1 or 0—need be made. But like debunkers of magic acts, their opponents were able to show time after time that there was something hiding behind the curtain; information was lurking in the cracks of the machines. Once one had tried on Szilard's newly ground eyeglasses it was hard not to see bits everywhere. If, at the end of a cycle of one of the automated machines, a weight was left dangling on either the left or right side of the piston, or if a lever was flipped one way or the other, this was considered information, 1 or 0. The machine, no matter how crude or lumbering, had a memory—it stored a bit that represented the state that the molecule had previously assumed. To repeat the cycle and keep the engine turning, someone or something would have to reset the machine. And how this was done depended on which of two states it was in, which weight was suspended, which way the lever leaned. Information would have to be gathered, a decision made. Implicit in this was an idea that would not fully emerge until Charles Bennett entered the picture in 1973: that it was not the actual gathering of information but its erasure—resetting the apparatus—that necessarily dissipated energy and saved the second law. And that, at last, is where the notion of driving up La Bajada fueled by a memory tape comes into the story.

In 1961, Rolf Landauer of IBM set out to do for the digital computer what Carnot had done for the steam engine: plumb its thermodynamic depths. The second law showed that nature sets limits on how efficient a heat engine can be. Steam engines could never convert 100 percent of their heat into energy because some was dissipated irrevocably into the environment as heat. When work is performed, a minimum amount of energy must always be irreversibly lost. Everything in the saga of Maxwell's demon suggests that the same might be true for the labor we call computation. It was left for Landauer to clinch the argument.

Punch 2 + 2 into a calculator, press the "equals" button, and the display says 4. But if you find a calculator someone has left on a desk and *it* says 4, you have no way of knowing where the number came from. Did someone punch in 2 + 2, 3 + 1, 1 + 1 + 1 + 1, or perhaps 9 − 5, or 1,239,477 − 1,239,473? There are an infinite number of calculations that can yield this same answer. Such a computation is irreversible. You can't go from 4 back to 2 + 2. The expression 2 + 2 contains more information than the expression 4—a surplus that is lost when you complete the computation.

Where does the information go? Landauer showed that it is dissipated into the environment as heat, and is as difficult to gather up again as the friction generated by a turbine or the molecules in a glass of water dumped into the sea. His argument went something like this: Recall the apparatus in Maxwell's original thought experiment, with two chambers connected by a valve. Once the gas has been loosed from the confines of the left-hand chamber, so it is free to fill the whole vessel, thermodynamics tells us that it takes energy to squeeze it back into the first chamber again. We are taking a system that now has many more degrees of freedom—all the ways the molecules might be arranged throughout the entire container—and squeezing it back into a system with many fewer degrees. The same is true for the calculator. Information has to be represented by physical states, whether voltages in a wire or positions of beads on an abacus. In an electronic calculator, 2 + 2 is represented by a string of 1s and 0s each held by a transistor that is either on or off. Each of these memory cells then has two degrees of freedom; it can represent either a 1 or a 0. To erase it, Landauer figured, the two degrees would have to be squeezed back into one: a memory cell that could only be empty. And, as with the gas, that would require a minimum amount of energy. As a digital computer churns through a long series of calculations, clearing registers so they can be filled again and again, the machine throws away information, shedding heat into the environment.

In the computers we build, the energy lost from clearing memory registers is insignificant compared with the energy consumed by resistance in the connections, the filament that lights the video display, or the motor that turns the disk drive. Still, these losses are dependent on the technology used; in theory they can be made as slight as we wish. But nature seems to put a limit on how cheaply we can erase bits. Below a certain level, the loss cannot be reduced. Information, Landauer argued, indeed is physical.

But that was not the end of the story. A little more than a decade later, Landauer's colleague Charles Bennett was struck by one of those questions that seem both simple and profound: What if you don't erase? Imagine that each time a computer made a calculation, it saved the intermediate result. As the machine ran through a chain of computations, it would accumulate a tape of its history. It would come up with an answer without having thrown away information. Aha, you might think. This is where it must pay the thermodynamic cost: the tape of all those intermediate steps must be erased to make room for more. But no. Bennett showed that the machine could be reset simply by running the tape backward, retracing its computational history until it was in its original state. Computation is merely the converting of an input (the question) to an output (the answer) according to a set of rules. Usually this is a one-way flow—given 4, we can't uniquely infer 2 + 2. But with a reversible computer, we have the extra information needed to convert the output back into the input. And why should going from output to input require any more energy than going from input to output? Of course, a huge disadvantage would be that in reversing the computation the answer would be lost. But, Bennett pointed out, before we kicked the machine into reverse, we could copy the answer onto a blank tape. And copying, unlike erasure, does not incur a minimum energy cost.

While such a machine would dissipate energy through electrical resistance, whirring disk drives, and glowing video screens, the actual act of computation could be done with no minimum energy cost. Though erasing information requires an amount of energy below which it is impossible to go, computation can otherwise unfold using an arbitrarily tiny amount of work. In fact, Bennett designed a hypothetical computer ered by nothing more than Brownian motion, the natural thermal vibration of molecules. While this boundless thermal reservoir could not be used to power a perpetual motion machine, as some fans of Maxwell's demon had hoped, it apparently could be tapped to perform computations, as long as you were provided with one of these carefully designed, reversible computers.

A few scientists, such as Edward Fredkin of Boston University, believe that the possibility of reversible computation implies that information is more fundamental than matter and energy, unconstrained by the second law. He envisions a hidden layer beneath what is currently taken as the laws of physics, where the shuffling of bits somehow gives rise to the world we see. The implication, of course, is that reality is some kind of simulation. The question of what is running the simulation or why is left as an exercise for the reader. Fredkin has called for an effort to recast the laws of physics in the form of algorithms for this hypothetical machine, carving up the world in an entirely different manner. But little work has been done in this direction.

Most scientists in the small world of information physics take Bennett's work as an amplification of Landauer's principle rather than a contradiction. Bennett strengthened the notion that it is not the gathering but the erasure of information that necessarily dissipates energy. The demon could make each measurement expending an arbitrarily small amount of energy. But before it acted—opening or closing the trapdoor—it would have to store the result of the measurement in its memory. The second law would exact its toll when the bits were erased, for that would require at least as much work as was generated by the engine. Another way to look at it is that as the demon is lowering the entropy of the gas, creating a more orderly arrangement of molecules, it is funneling all that randomness into its memory, scrambling its brain. Or, if you subdivide the system a little differently, the memory can be considered part of the environment. So, once again, creating order in one place requires exporting entropy to another.

But again—and here is what Bennett was thinking of that day in Santa Fe—what if you don't erase? A memory can be made as gigantic as you like. Imagine it as a long tape. The demon could just keep filling it with bits and postpone erasure indefinitely. Hook the demon and its engine to a set of wheels and you could drive down I-25, up and over La Bajada, and back to the Albuquerque airport, spewing out an exhaust of 1s and 0s all the way.

There is nothing mystical about this subtly different way of carving up the physical realm. Like anything, Bennett's engine is mired in the laws of thermodynamics. One doesn't find blank memory tape sitting around in the world waiting to be used as fuel. Its state, all 0s, is highly improbable, as unlikely as a gas that sits only on the left side of a container. It takes work to create this order. Then this work can be exploited by letting the tape run down the information hill, from orderliness to randomness, just as the energy behind a dam or stored between the poles of a battery runs down the entropy hill.

Viewed this way, there is no reason why a battery cannot be thought of as a memory tape. It begins in a blank, orderly state (positive charges at one pole, negative charges at the other) and is randomized as it runs down. Once the battery is spent, the rearrangement of its molecules is a memory—a history of its use. This record is erased by recharging. The positive and negative charges are sequestered again, the order is restored, the system reset. But the battery charger dissipates heat, transferring the randomness to the environment, filling it with jumbled bits. In fact, we can think of the universe as a memory tape, blank and structured. As all this order turns into entropy, the Universal Memory Tape is filled with random bits. But it can never be erased. There is nothing to erase it, nowhere to export the randomness. You cannot reset the universe. The randomness just keeps on accumulating. And that is the information version of the second law.

Maxwell saw entropy as purely subjective and concluded that the more intelligent a creature, the more work it could extract from a source of fuel. Szilard took the first step toward demystifying this notion. But the trade-off was that he had to elevate information—something most of us think of as subjective and man-made—to the objective realm.

Though intelligence doesn't allow us to overcome the second law, it remains true that creatures with more acute senses and more powerful brains will see pattern where others see randomness. How can we have a science if every observer, depending on its abilities, looks at the same system and perceives a different entropy, a different order? Zurek showed that we can get around this problem if we take up our ontological scalpel and slice the world yet another way, into two kinds of entropy.

Zurek's division is based on the notion that there are two kinds of information. First there is Shannon's information, measuring how improbable a structure is: an intricately patterned system is highly unlikely, and so it can be said to contain a lot of information. But there is also Chaitin and Kolmogorov's way to look at the situation, using algorithmic information. From this opposite perspective, a highly patterned, compressible system takes fewer bits to describe than a random one. If all the gas molecules are in chamber A, our demon's memory tape might say 111111111111111111111111 . . . And this could be reduced to a simple algorithm: "repeat '1' ten billion times." On the other hand, once we open the valve and let all the molecules into the whole vessel so that they are distributed at random, the only way to describe the system's state would be to specify where each molecule is. By definition, this random string of 1s and 0s cannot be reduced to a shorter algorithm. It is incompressible.

Zurek proposed that we define what he calls physical entropy as consisting of two reciprocal quantities: the ignorance of the observer, measured by Shannon's statistical entropy, and the randomness of the object being observed, measured by its algorithmic entropy—the smallest number of bits it would take to record it in memory. The beauty of this explanation is that, during measurements, the ignorance of the measurer decreases as its memory tape gets longer—so their sum, the physical entropy, remains the same.

Different demons with different acuities may disagree over how much order or entropy there is in a system. The more precisely something is measured, the less random and entropic it might seem. But the more precise the measurements, the more scrambled the demon's brain will become. Taken together, the two kinds of entropy balance out, so that from the point of view of an outsider looking in, the physical entropy of the whole system remains the same.

The important lesson to take away from all this is that the measurer must always be included as part of the system. The result is nothing less than a law of conservation of information, which Zurek would like to see stand alongside the conservation of energy and the conservation of momentum as the pillars on which science stands.

There is a final coda to the story of the demon. Looking back at the march of ideas that began with Maxwell, we can see that the key to exorcising the demon was to replace it with an information processor, a juggler of bits. But then measuring molecules and deciding when to open trapdoors doesn't require much intelligence. In all the thought experiments, the demons could be made from a few dollars' worth of electronic parts.

But, Zurek began to wonder, what if the demon was *really* smart? Suppose, for example, that in observing the position of the molecules, its memory tape recorded a string that included this sequence: 100100100100100100100100100. Now with some simple pattern-recognition algorithms, the demon could notice the regularities. The first string could be replaced by an algorithm that said "repeat '100' nine times." Or suppose part of the string included the sequence 11111111111111111, which could be replaced by "repeat '1' seventeen times." The algorithms are much shorter than the strings. After it made its measurements, a smart enough demon (using a reversible computer so that no energy was consumed) could compress its memory. Then it would have fewer characters to erase. Less energy would be dissipated. The demon would pay less than the full thermodynamic price. Purely random sequences couldn't be compressed. But if there were pockets of regularity, the string could be squeezed shorter. (This, in essence, is what a program like DiskDoubler does when it compresses the information on a computer's hard disk so that it takes up less room.) Some sequences that appeared random might in fact contain hidden order. If the demon was really smart—and here we are encroaching on the realm of artificial intelligence—it might notice that part of the memory string represented the decimal expansion of the square root of 2 or of pi.

Was there something transcendental about intelligence after all? Could a demon this intelligent get a slight edge on the second law, where dumber demons had failed? Zurek suspected not.

Chaitin showed that it is impossible to prove whether or not a particular number is random, whether a compression is the most concise—it is always possible that there is more order that can be squeezed out. The demon could never know if it had made its memory tape as compact as possible. But the demon wouldn't have to find the most concise description. Anything that significantly reduced the number of bits it had to erase would save energy. The question was whether compression could save so many bits that a very intelligent demon got more work out of the system than it put in. But once again, the second law was rescued. Zurek showed that Shannon's information theory and the laws of computation put a limit on compressibility. The most concise description of a message still must contain at least as much information as was in the original. The demon's memory tape can be thought of as the receiver of a message whose source is the pattern of the gas molecules. It is this very information that allows the demon to do work, manipulating the molecules to lower entropy. But while the "message" might be compressible, whatever is squeezed out is redundant information that does not contribute to the demon's success. In the end, when the demon erases its memory, it cannot throw away any less information than was required to make the engine run. Compressing the memory tape will make the demon more efficient, and if one was lucky enough to stumble upon the most compact description, then the engine would reach maximum efficiency, expending just as much energy as it generated. But Zurek showed that the best it can ever do is break even. It can generate only as much energy as it takes to erase its memory.

It may strain our intuition to think of batteries as information stores and memory tapes as fuel. We think of matter and energy as fundamentals—we can feel the heft of a rock or the jolt of electricity. Information seems subjective. Yet why should what we know through our bodies be more fundamental than what we know through our brains? In the end, we only know about matter and energy through the signals sent by our senses—our eyes, ears, noses, the receptors in our skin. It all comes down to information. And yet what is this information but matter and energy—charged ions carrying electrochemical signals through our nervous systems. Landauer and Bennett showed the limitations that physics puts on computation; Zurek showed that laws of computation—the limits of compressibility—have implications for physics. And so the circle turns.

As one of the world's premier demonologists, Zurek sometimes finds himself identifying with Maxwell's little creature. To the demon, the gas in the chamber is the universe; its quest to find hidden orders is like the scientific quest to find universal laws. We decrease our ignorance by measuring, but only at the cost of this informational exhaust. If there were little order in the universe, if it were in equilibrium like the gas, then we would simply be funneling the randomness intact to our memory tape—the library of scientific knowledge. We would be no better off than the demon, measuring and measuring but never getting ahead. We could gather bits and bits of data, but we couldn't compress them into more compact forms, the succinct statements we call universal laws. Science would be reduced to cataloguing every fact about every particle. The universe, like a random string, would be its own shortest description.

Of course, a completely random universe wouldn't have information gatherers at all. There would be no structure. Our very existence stands as proof that the universe we live in is far from being in equilibrium. There is order to exploit, compressions to be made. And so, as Zurek says, it pays to measure.